A Model of Focusing in Economic Choice[*]

Botond Kőszegi
UC-Berkeley

Adam Szeidl
UC-Berkeley, NBER, CEPR

July 22, 2011

Abstract

We present a generally applicable theory of focusing based on the hypothesis that a person focuses more on, and hence overweights, attributes in which her options differ more. Our model predicts that the decisionmaker is too prone to choose options with concentrated advantages relative to alternatives, but maximizes utility when the advantages and disadvantages of alternatives are equally concentrated. 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, the decisionmaker often exhibits a form of time inconsistency. She is present-biased when the costs of current misbehavior are distributed over many future dates (such as in harmful consumption), but "future-biased" when the benefit of many periods' effort is concentrated in a single goal (such as in career advancement). In a market setting, a profit-maximizing firm selling to a consumer with focus-dependent behavior chooses a product with one core attribute, and splits its price into multiple 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.

JEL Codes: D03, D40, D91

1   Introduction

People often focus disproportionately on, and hence overweight, certain 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, and therefore be too likely to take the job (Kahneman, Krueger, Schkade, Schwarz and Stone 2006). Similarly, a person comparing the quality of life in California and the Midwest may focus more on climate than on the many determinants of life satisfaction in which the two regions are similar, and hence be too likely to believe that California is the better place to live (Schkade and Kahneman 1998).

Motivated in part by the above examples and related evidence, we develop a model of focusing based on the idea that a person focuses more on attributes in which her options differ more. We demonstrate the economic relevance of this determinant of focus in two applications, individuals' choice over time and firms' product design and pricing behavior, showing that the model yields a unified explanation for seemingly conflicting evidence and intuitions, and also generates new economic insights. Furthermore, because our model makes a prediction in any choice problem in which a classical model would, it provides a simple way to incorporate this type of focusing into other economic applications.

Section 2 presents our framework, which attempts to meet a key challenge for portable economic models of focusing and attention—to define focus without appealing too much to "outside" information that is difficult to observe or not relevant in most economic settings. We model choices from a finite set $C \subset \mathbb{R}^K$ of K-dimensional deterministic consumption vectors, where each dimension represents an "attribute." The *consumption utility* from a choice $c = (c_1, \ldots, c_K)$—identical to welfare in our model—is $U(c) = \sum_{k=1}^{K} u_k(c)$. But instead of consumption utility, the decisionmaker acts to maximize *focus-weighted utility* $\sum_{k=1}^{K} g_k u_k(c)$, with her focus drawn disproportionately to 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.[1] Because there

---

[1] While our theory, applications, and results are different, our basic assumption that greater differences attract focus parallels some key assumptions in the framework for intransitive preferences by Tversky (1969), the models for choice under uncertainty by Bordalo, Gennaioli and Shleifer (2010) and Gabaix (2011), and the analysis of context effects by Bordalo (2011). We discuss these related approaches, as well as direct evidence for our model, in more detail in Section 2 below.

may be some available options (such as extremely unattractive options) that do not affect focus, we allow C to be different from the agent's entire choice set, thinking of it as a "consideration set" of reasonable options. We extend our basic setup to choice over time by assuming that a person thinks of her consideration 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 identify some central implications of our model. Most importantly, because the decisionmaker focuses too much on a few large advantages relative to many small disadvantages, she exhibits a "bias towards concentration": she is too prone to choose options whose advantages are concentrated in fewer attributes. This prediction drives many of our model's implications in specific applications. In "balanced choices," however—decisions in which the advantages and disadvantages of one option relative to another are equally concentrated—the agent maximizes consumption utility. This observation helps identify the types of situations in specific applications in which choices reflect welfare. In addition, the same observation allows us to elicit consumption utility and other model ingredients necessary for behavioral and welfare predictions from choice behavior, so that our model satisfies the spirit of the revealed-preference criterion for economic theories.

Turning to applications, in Section 4 we explore the agent's behavior in intertemporal choices. 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 relative costs and benefits depends not on their temporal placement, but on their *concentration*. Furthermore, because the concentration of relative 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.

Consider first the class of intertemporal decisions where, such as in the case of exercise, a larger sacrifice today results in small per-period benefits for many periods in the future. When deciding whether to exercise on each individual day, a person focuses too little on the many small gains relative to the one big cost, so she tends to exercise too little. But when considering the entire sequence of decisions ex ante, the person focuses on the large fitness gains she will enjoy

every day if she exercises regularly, so she is prone to commit to doing so. This class of decisions fits many types of choice situations, including consumption-savings decisions and lifestyle choices, that have been invoked as prime examples for present bias in the literature. Hence, our model can be thought of as providing a focus-based foundation for present bias that complements the hyperbolic-discounting-based mechanism of Laibson (1997) and O'Donoghue and Rabin (1999b), while also predicting that—it being a more balanced choice—the ex-ante preference toward more future-oriented behavior better reflects welfare.

By virtue of providing a foundation for present bias, our theory also makes predictions about when present bias and time inconsistency are likely to be stronger or weaker—or might be reversed. To illustrate such "future bias," next consider the class of intertemporal decisions where, such as in the case of writing a book, a sequence of sacrifices leads to a single large goal. When deciding ex ante whether to commit to the project, the decisionmaker focuses too much on the finished book, and hence she is too prone to commit. But when deciding whether to push along the project on any given day ex post, the person weighs effort that day against just a marginally better book, so she is less willing to work. In this case, while predicting the same kind of present-oriented time inconsistency in effort choices as existing models of present bias, our theory deviates from these models in saying that—it being more balanced—the ex-*post* choice better reflects welfare. This class of choice problems 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, for example). Moreover, since the diffuse incremental costs associated with each *additional* achievement look small relative to the single benefit, we predict that the decisionmaker often adds more future-oriented challenges—writing an editorial, becoming a journal editor, and so on—that she would not agree to ex ante. Hence, for such commitment choices, she exhibits the opposite time inconsistency than that implied by existing models.

In Section 5, we explore how firms design and price products when facing consumers with focus-dependent choice, thereby contributing to a recent literature (reviewed in Spiegler 2011) that analyzes traditionally marketing questions using economics methods. Because a consumer dislikes

concentrated disadvantages, a firm has an incentive to split the price of a product into multiple components the consumer evaluates as separate attributes. This prediction may explain why many retailers (e.g. airlines and mortgage brokers) artificially split prices, and provides one reason for the popularity of financing for purchases. In addition, because a consumer likes concentrated advantages, a firm has an incentive to concentrate product value on a single attribute. This implication seems consistent with marketing analyses of successful brand positioning and consumer "value propositions" (Aaker 1991, Anderson, Narus and van Rossum 2006, and others). Moreover, because the consumer focuses too much on the concentrated value relative to the dispersed price, we predict that a firm often has an inefficiently strong incentive to increase that value.

We also analyze firms' incentives in choosing which value attribute to concentrate on. As examplified by the association of Volvo with safety and of BMW with performance, firms often try to position their products 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, firms also frequently attempt to copy the key attribute of a competitor. Our model predicts that these two strategies are optimal for different kinds of firms. For a firm that produces higher overall qualities and hence determines consumers' focus, being especially strong on and thereby drawing more attention to the competitor's weak attribute is optimal. For a firm that produces lower overall qualities and hence does not affect consumer focus, being relatively strong on and thereby taking advantage of the attention consumers already pay to the competitor's strong attribute is optimal.

Beyond its ability to make predictions in the applications we explore in this paper, a key advantage of our model is its portability to new economic domains. A recipe for translating a deterministic classical model into one with focus-dependent choice is to (i) specify the relevant attributes in the given setting; (ii) take the utility function from the classical model; and (iii) equate the consideration set with the choice set of the classical model.[2] Nevertheless, application of our model is not entirely mechanical. While specification of the attributes seems fairly straightforward in many applications, it might sometimes require non-trivial modeling judgment. And due to the

---

[2] A similar recipe works for applying our model to a new setting for which no appropriate classical model exists: specify (i) the relevant attributes; (ii) the utility function; and (iii) the consideration set. Note that in such an application, a classical model would also have to make assumptions corresponding to (ii) and (iii).

potential effect of unchosen options on behavior, the recipe might require adjustment when the choice set is not an appropriate approximation of the consideration set. We discuss further issues with our model, as well as other potential applications, in Section 6.

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; we are currently developing and extension to risky choice (Kőszegi and Szeidl 2011, hopefully), and we discuss modeling issues that arise in the context of that extension in the conclusion.

2.1 Focus-Weighted Utility

We model choices from a finite set $C \subset \mathbb{R}^K$ of K-dimensional consumption vectors, where each dimension represents an "attribute." The decisionmaker's *consumption utility* and welfare if she chooses option $c = (c_1, \ldots, c_K) \in C$ is $U(c) = \sum_{k=1}^{K} u_k(c_k)$, and we will often represent an option in C by its vector $(u_1(c_1), \ldots, u_K(c_K))$ of consumption utilities rather than by its vector of consumption levels.[3] But instead of consumption utility, 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.

The major challenge for any broadly 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

---

[3] Following classical economic models, in our formulation consumption utility depends only on the agent's consumption level. For our model to make correct predictions, however, in some situations it might be necessary to start from a richer notion of consumption utility, incorporating for instance reference-dependent utility or social preferences.

are highlighted in red on the product's packaging. Using only model ingredients introduced so far, our central assumption is that the decisionmaker focuses more on attributes in which her options generate a greater range of consumption utility:[4]

Assumption 1. The weights g_k are given by $g_k = g(\Delta_k(C))$, where $\Delta_k(C) = \max_{c' \in C} u_k(c'_k) - \min_{c' \in C} u_k(c'_k)$, and the function $g(\Delta)$ is strictly increasing in Δ.

Some evidence directly supports Assumption 1.[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. And in the context of choice bracketing, Read, Loewenstein and Rabin (1999) discuss evidence for "adding-up effects," whereby individuals pay more attention to small repeated costs (e.g., daily coffee expenditures) if they think of these costs as integrated—and hence larger. But rather than the direct evidence, we view our model's ability to organize evidence and intuitions from a variety of domains (including intertemporal choice and product design below) as making the most compelling case for it.

While the above evidence supports the general hypothesis that large differences attract focus, the precise role of the number and dispersion of options in mediating this effect is (as far as we know) unexplored empirically. Without evidence to go on, Assumption 1 imposes, primarily for simplicity and tractability, a specific functional form for $\Delta_k(C)$. The logic of our results in this

---

[4] Literally interpreted, Assumption 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. Since this adjustment would merely multiply $\tilde{U}(c, C)$ by a constant, it would not affect the choice or welfare implications of the model.

[5] The evidence by Dunn et al. (2003) and Schkade and Kahneman (1998) we discuss pertains to predicted happiness rather than choice—but presumably individuals would make choices in these situations corresponding to their predicted happiness.

6

paper does not seem to depend on the particular formulation, but some predictions of our model—especially regarding choice from sets with more than two relevant options—might be sensitive to it. We hope that our theory will encourage experimental and empirical work testing and improving this aspect of our formulation.

Our assumption that the weights g_k affect choice but not welfare captures the view (consistent with the psychology literature above) that our comparison-based focusing effect distorts how a person perceives her own preferences near the moment of choice, but does not significantly alter the experienced utility emanating from the choice. Suppose, for example, that when choosing a cellphone a consumer's focus is drawn too much to touch screens due to her comparison of phones with and without this feature. While this focusing effect may lead her to choose a phone with a touch screen, when using her new phone she will not usually care or even think about phones without a touch screen, so her experienced utility will not be much affected.

Because not all comparisons necessarily affect a person's focus, we allow C to be different from the agent's entire choice set, thinking of it as a "consideration set" of her reasonable options. For instance, in most situations a person quickly dismisses or does not even think about "self-destructive" and other clearly inferior options, so that they do not influence her focus. While most economic models of individual decisionmaking start from an exogenously given set from which the agent chooses, correct specification of this set is—due to the effect of unchosen options on behavior—more crucial in our setting than in most classical models. But this issue is not an impediment to applying our model: as we do in Sections 4 and 5 below, in many settings it seems reasonable to simply equate the consideration set with the choice set of the corresponding classical model as typically specified in practice. Our theory of choice from consideration sets can also be combined with any model of consideration-set determination.[6]

Our central hypothesis that large differences attract focus is closely related to key assumptions in several existing theories. Tversky (1977) proposes and provides evidence for a theory of intransitivities in binary choice in which a decisionmaker does not notice small differences in an important

---

[6] While we do not have a complete model in mind, we mention here two potentially desirable properties of a theory of consideration-set determination: (i) C should not contain dominated options; (ii) C should not have a proper subset C' such that any member of C' beats any member of $C \backslash C'$ in pairwise comparison.

attribute, but does notice and heavily weight larger differences. Our model assumes that the same idea applies to other attributes and more choices, and aims not just to explain certain intransitivities, but to apply the idea broadly to economic questions. In their model of choice between lotteries, Bordalo et al. (2010) and Bordalo (2011) make a number of assumptions about how relative payoffs affect salience, one of which is that salience of a state of the world is increasing in the difference between 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) also applies the framework to choices among products, and explores its implications for context effects. We study implications for different settings than do these authors, and our model uses *only* the assumption that larger differences attract more focus. 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. Our model assumes that it is the comparison between alternatives, rather than environmental uncertainty, that determines focus, and hence both yields different predictions and is most compellingly applied to different settings than Gabaix's theory.

Several comments about interpreting and applying our basic theory are in order. First, although our assumptions were not derived formally from patterns in choice behavior—as would be done in axiomatic theories—we show in Appendix B that if we know the relevant attributes, all ingredients of our model (the utility functions $u_k(\cdot)$ and focus-weight function $g(\cdot)$) can be identified from choices using a simple algorithm. This means that our theory's full set of predictions regarding choice *and* welfare can be identified from behavior in a limited number of settings. In this sense, our model satisfies the spirit of the revealed-preference criterion for economic theories. We also illustrate in Appendix B that if we know the set of *potential* attributes in a situation, under plausible conditions we can identify which of these the decisionmaker treats as identical versus separate attributes.

Second, to simplify the definition of the range of consumption utility in an attribute (which in turn determines focus), our formulation assumes additive separability in utility across attributes. 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 $U(c_1, ..., c_K)$. 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 $\Delta_k(C) = \max_{c \in C} U(c_k, c^0_{-k}) - \min_{c \in C} U(c_k, c^0_{-k})$, and define the focus-weighted utility of c with respect to c^0 as $wU(c) + (1-w) \sum_{k=1}^{K} g(\Delta_k(C)) \left[U(c_k, c^0_{-k}) - U(c^0) \right]$. 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.

Third, given the psychological evidence 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 often does not put appropriate weights on them. This phenomenon is slightly different from what is usually referred to in the economics literature as "(in)attention," that a person may not be aware of or understand some pieces of information. Hence, while we will use both terms in the paper, we refer to the phenomenon we explore primarily as "focusing."

2.2 Focusing in Intertemporal Choices

We now extend our model to intertemporal decisions, which requires us to specify how a person conceptualizes her dynamic choice problem as she makes decisions in each period. 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, we would assume that the agent 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 we assume that 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 focus is determined by the perceived consequences of only the decision at hand, rather than the entire sequence of current and future decisions.

Formally, there are T periods, $t = 1, \ldots, T$. In period t, the agent makes a choice x_t from the deterministic finite consideration set $X_t(h_{t-1})$, where $h_{t-1} = (x_1, \ldots, x_{t-1})$ is the history of choices up to period $t-1$. Continuing with the consumption-savings example, x_t can represent the bundle of goods to be consumed in period t, and $X_t(h_{t-1})$ the budget set given the path of past consumption. The decisionmaker's consumption utility in period t is $\sum_{s=t}^{T} \delta^{s-t} u_s(h_{s-1}, x_s)$, where u_s is the possibly history-dependent instantaneous utility function in period s. We assume that utilities realized on different dates are evaluated as separate attributes, and at each date also allow for multiple attributes with additively separable utilities. For any consumption profile $(x_1, \ldots, x_T)$ and date t, let $V_t(x_1, \ldots, x_T)$ be the induced vector of consumption utilities for current and future periods, $(\delta^{s-t} u_s(h_{s-1}, x_s))_{s=t}^{T}$.

We represent the decisionmaker's beliefs about how her choice in period t affects her future behavior by the functions $\{\tilde{x}_\tau^t(h_{t-1}, x_t)\}_{\tau=t+1,\ldots,T}$, which specify future choices as a function of h_{t-1} and x_t. For any history, these beliefs induce a set of lifetime consumption-utility profiles:

$$C_t(h_{t-1}) = \left\{ V_t(h_{t-1}, x_t, \tilde{x}_{t+1}^t(h_{t-1}, x_t), \ldots, \tilde{x}_T^t(h_{t-1}, x_t)) \middle| x_t \in X_t(h_{t-1}) \right\}. \tag{2}$$

We assume that $C_t(h_{t-1})$ is the consideration set that determines the decisionmaker's focus in period t, so that she applies the model of Section 2.1 to $C_t(h_{t-1})$.

A key question we explore in our analysis of intertemporal choice is whether and how the agent's behavior is different from 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-utility profiles. Applying (2), in this case the individual's focus in period 1 is determined by the range of consumption utilities generated by all possible consumption paths.

Our framework above of choice given beliefs about future behavior can be combined with any theory of how these beliefs are determined. Following standard economic methodology, in this paper we assume that the agent has rational (correct) beliefs. This implies that we can derive the agent's behavior in any decision problem using backward induction. In one alternative, "naive," theory, the decisionmaker believes that she will act in the future as she would now commit to do. In this formulation, the decisionmaker's beliefs at time t about about her future actions following each possible choice x_t are determined as the optimal behavior in the commitment problem starting

10

in period $t + 1$, given x_t. Intuitively, a naive agent formulates her beliefs about future behavior with a general global view of her decision problem, but makes each specific decision based on a local view.[7] The insights in this paper hold for both theories of how the agent forms beliefs about the future.

3   Bias Toward Concentration

This section identifies general properties of our model that underlie many of its predictions in specific applications. Most importantly, we show that the decisionmaker is biased toward options whose advantages relative to alternatives are more concentrated than its disadvantages—and conversely, she is biased against options whose disadvantages relative to alternatives are more concentrated than its advantages. At the same time, the agent is not biased when the advantages and disadvantages of alternatives are equally concentrated.

Bias toward concentration. For a simple example demonstrating our model's most important property, suppose that a consumer purchasing a laptop is choosing whether to make one immediate payment of \$899 or 24 future monthly payments of \$39. Positing that payments at different dates correspond to different attributes, and assuming no discounting, a zero interest rate, and linear consumption utility, the consumer represents the decision as choosing between the streams $(-899, 0, \ldots, 0)$ and $(0, -39, \ldots, -39)$.[8] Although paying up front is the financially superior option, in our model the consumer chooses financing if $g(899) \cdot 899 > g(39) \cdot 24 \cdot 39 = g(39) \cdot 936$—which may well be the case since $g(\cdot)$ is increasing. Intuitively, the single large upfront payment attracts more focus than the many smaller monthly payments, leading the agent to overweight it. Equivalently, the agent is biased toward financing because relative to the alternative, it has a large concentrated advantage: not having to pay a large sum immediately.

We now identify general conditions under which the bias toward concentration occurs. For any

---

[7] 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 in more specific terms.

[8] See Section 5 for a discussion of the assumption that the decisionmaker evaluates different payments as different attributes.

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 clearly holds for the financing example: relative to paying immediately, financing has exactly one advantage of \$899 (in the first payment), and disadvantages of only \$39 (in all the other payments). Proposition 1 uses this condition to formalize the bias toward concentration:

Proposition 1 (Bias Toward Concentration). *For 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 $\{c^1, c^2\}$.*

Proposition 1 says that the agent always chooses an option over an alternative with equal consumption utility and less concentrated advantages. By monotonicity and continuity, this also means that a person always chooses an option with more concentrated advantages if its consumption utility is higher, or even lower but sufficiently close. This bias toward concentration is the key behavioral implication of our model in that it drives the main predictions in intertemporal choice and product design below, including the predictions that we argue unify seemingly disparate evidence and intuitions.

We also formalize a 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 advantage in attribute k. Formally, this holds if $u_k(c') - u_k(c^2) = \left[u_k(c^1) - u_k(c^2)\right] + \left[u_l(c^1) - u_l(c^2)\right]$, $u_l(c') = u_l(c^2)$, and $u_m(c') = u_m(c^1)$ for all $m \neq k, l$. We define when c' combines two *dis*advantages of c^1 relative to c^2 analogously. Then:

Proposition 2 (Increasing Concentration). *If c' combines two 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 two 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 $\{c', c^2\}$.*

12

Proposition 2 says that concentrating an option's advantages makes it more preferable, while concentrating its disadvantages makes it less preferable. The reason is simple: because the larger differences created by the more concentrated advantages and disadvantages attract more of the agent's focus, they are weighted more heavily in her decision.

Our prediction of a bias toward concentration is sometimes in contradiction with an implication of prospect theory's (Kahneman and Tversky 1979) diminishing sensitivity, the property of reference-dependent preferences that an initial deviation from a reference point is felt more strongly than an increase in the deviation. As pointed out by Thaler and Johnson (1990) and others, diminishing sensitivity implies a preference for segregating gains and integrating losses relative to the reference point. But while there is compelling evidence for diminishing sensitivity based on people's attitudes toward outcomes closer to versus further from the reference point,[9] there is only limited evidence for a preference for separating gains and integrating losses.[10] Combined with diminishing sensitivity in consumption utility within each attribute, our model both accommodates the former evidence, and—because in the combined framework the latter prediction is weakened or reversed by the bias toward concentration—explains the lack of the latter evidence.[11]

Balanced choices. As an example of the other key property of our theory, suppose that the laptop buyer above is asked not whether she prefers financing over a lump sum payment, but whether she wants to make a lump-sum payment of $899 at the time of purchase, or a lump-sum payment of $936 a month later. This is a "balanced" choice because the advantages of the two alternatives are

---

[9] 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).

[10] 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.

[11] Furthermore, our theory seems useful for making sense of some examples commonly used to illustrate diminishing sensitivity in the literature. One popular example aimed at explaining add-on purchases argues that due to diminished sensitivity to additional expenditures following a big payment, consumers are more willing to buy a car stereo after buying the car itself than at other times (Thaler 1985, Chiu and Wu 2009, for example). Without judiciously chosen assumptions, 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: the consumer is more likely to purchase the car stereo when buying the car because the purchase decision tilts her focus toward the "car" attribute, which increases the perceived value of the add-on as well.

equally concentrated—on a single payment each. Then, because $g(899) \cdot 899 < g(936) \cdot 936$ for any increasing $g(\cdot)$, the consumer maximizes consumption utility and chooses the first option. While applying the focus weights changes her perceived utility of the two 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 $K', a, b > 0$, the number of k such that $u_k(c^1) - u_k(c^2) = a$ is K', the number of k such that $u_k(c^2) - u_k(c^1) = b$ is K', and the number of k such that $u_k(c^1) = u_k(c^2)$ is $K - 2K'$. That is, relative to c^2, c^1 is better by a on K' attributes and worse by b on K' attributes. Then:

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

We use this property of the model repeatedly in Section 4 for guidance about the kinds of situations in which choices are most likely to reflect welfare. And as we show in Appendix B, this same property helps identify our model's ingredients from behavior: we can use the rationality in balanced choices to elicit consumption utility, and then the bias in unbalanced choices to elicit $g(\cdot)$.[12]

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 $u = (u_1, \ldots, u_K)$ and $\delta = (\delta_1, \ldots, \delta_K)$, there is an $\bar{\epsilon} > 0$ such that if $0 < \epsilon < \bar{\epsilon}$, the decisionmaker makes a consumption-utility-maximizing choice from the set $\{u, u + \epsilon \cdot \delta\}$.*

For an intuition, notice that for small stakes, the focus weight on each attribute is near the minimal level $g(0)$. If $g(0) > 0$, therefore, the agent puts approximately the same relative weight on all attributes, yielding near-optimal behavior. Both because small-stakes decisions are all else

---

[12] Extending the logic of Proposition 3, the agent maximizes consumption utility 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 proportional pattern across attributes as its disadvantages.

14

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 possibility that the agent makes better decisions for small stakes than for larger stakes 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 (2011), in which the agent makes rational choices taking into account the cognitive or effort cost associated with attention or calculation. While in these models an increase in stakes increases cognitive effort and hence the quality of the agent's decision, in our model the same change can just exacerbate the agent's bias toward concentration.[13]

4 Intertemporal Choice

This section explores some implications of our model for choice over time. A fundamental insight— and a direct consequence of the bias toward concentration—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 often predicts a form of time inconsistency in behavior. Elaborating on these insights yields two main contributions. First, our theory endogenizes time preference and time inconsistency based on the concentration of the costs and benefits of available actions, providing a common explanation for several distinct phenomena discussed in the literature, as well as making new predictions. Second, in contrast to existing work on time inconsistency, our theory provides an unambiguous welfare measure without a priori assuming that ex-ante choice reflects 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.

---

[13] If the function $g(\Delta)$ has an upper bound, our model also implies near-optimal behavior if stakes are sufficiently large in *all* attributes. Even then, the agent exhibits a non-trivial bias toward concentration when differences are smaller in at least some attributes.

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, $t = 1, \ldots, T_i + T_b$. 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 last $T_b \geq 1$ 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 $u_t = A \left(\sum e_t \cdot B \right) / T_b$. The variable A measures the efficiency of the investment: investing maximizes consumption utility if and only if $A \geq 1$. Abstracting away from any time discounting in intrinsic preferences, we assume that $\delta = 1$.[14],[15]

We begin by illustrating some of our formal results on ex-ante and ex-post choices in the above decision problem using two examples. Recall from Section 2.2 that in the ex-ante or commitment problem the agent makes all decisions at $t = 1$, while in the ex-post or no-commitment problem she makes her decisions period by period given her beliefs about future behavior.

Example 1 (Exercise). A consumer decides 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 of the periods 101 through 200. ($T_i = T_b = 100, B = 80, A = 1.25$.)

Because $A > 1$, the consumption-utility-maximizing choice in any given period is to exercise. Without commitment, the consumer's decision in each period can be represented as a choice between $(-80, 1, ..., 1)$ and $(0, ..., 0)$, where the 80 is the current pain and the 1's are the future benefits from exercising in that period. 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.

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, $(-80, \ldots, -80,$

---

[14] There are also other considerations (e.g. possible overlaps between the investment and benefit periods and non-binary choices) that our simple setting ignores. 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.

[15] 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.

$100, \ldots, 100)$ and $(0, \ldots, 0)$, and these alternatives also generate the focus weights on all attributes. Because each alternative is better than the other on exactly 100 attributes $(T_i = T_b)$, this is a balanced choice, so that the consumer commits to the consumption-utility-maximizing option of always exercising.

In Example 1, therefore, the consumer exhibits a time inconsistency in behavior similar to that in models of present bias: she is more present-oriented in ex-post choice than in ex-ante choice. Intuitively, any one day's workout has an attention-grabbing concentrated current cost and easy-to-neglect dispersed future benefits, leading the consumer to focus little on the benefits. But because the incremental effects of daily exercise accumulate into large daily gains, from the perspective of a lifetime the consumer focuses more on these benefits. Furthermore, our model says that it is the consumer's balanced ex-ante choice that reflects her true well-being, so that she is present-biased in ex-post choice and unbiased in ex-ante choice.

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. $(T_i = 100, T_b = 1, B = 50, A = 0.8.)$

Without commitment, the scholar's decision in each period can be represented as a choice between $(-50, 40)$ and $(0, 0)$, where the 50 is the the current pain of writing and the 40 is the resulting incremental improvement in the book. Since this is a balanced choice, in each period the scholar makes the consumption-utility-maximizing choice of not working on the book. From an ex-ante perspective, however, the scholar's choice is effectively between $(-50, \ldots, -50, 4000)$ and $(0, \ldots, 0)$, so that if $g(4000)/g(50) > 1.25$, she commits to working on the book every period.

Like the consumer in Example 1, the scholar exhibits time inconsistency in behavior in that her ex-post choice is more present-oriented than her ex-ante choice. The reason is also the same: because the benefits of multiple periods of effort accumulate into a large gain, they attract more focus in the ex-ante than in the ex-post choice. In contrast to the exercise example, however, in this case our model says that the scholar's ex-*post* choice maximizes her well-being, so that she is unbiased in ex-post choice and *future*-biased in ex-ante choice. Because in her ex-ante choice the

scholar focuses on the large concentrated gain from having a quality book and pays relatively less attention to the dispersed everyday costs, the project is overly appealing to her.

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 $i = ante, post$, we say that the decisionmaker is *present-biased* 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 her present bias is greater if A^*_i is greater. Analogously, the decisionmaker is *future-biased* if $A^*_i < 1$, and her future bias is greater if A^*_i is lower. The agent is unbiased if $A^*_i = 1$. We say that the agent exhibits present-oriented time inconsistency—i.e. she is more present-oriented ex post than ex ante—if $A^*_{ante} < A^*_{post}$, and her time inconsistency is greater if A^*_{post}/A^*_{ante} is greater.

Proposition 5 (Biases in Investment). *Suppose $g(\cdot)$ is continuous. Then:*

1. *The agent is present-biased in ex-post choice for any $T_b > 1$, unbiased for $T_b = 1$, and her present bias is increasing in T_b. Her present bias in ex-post choice does not depend on T_i.*

2. *The agent is present-biased in ex-ante choice if $T_b/T_i > 1$, future-biased if $T_b/T_i < 1$, and unbiased if $T_b = T_i$. Her future bias in ex-ante choice is decreasing in T_b/T_i.*

3. *If $T_i = 1$, the agent is time consistent, and if $T_i > 1$, she exhibits present-oriented time inconsistency. The degree of time inconsistency is increasing in T_i.*

Part 1 of Proposition 5 generalizes the present bias in Example 1; Part 2 identifies when a future bias such as that in Example 2 occurs; and Part 3 generalizes the time inconsistency in both examples. We now turn to discussing some economic implications of the more detailed results in Proposition 5.

Lifestyle choices versus achievements. When $T_i \approx T_b \gg 0$, Parts 1 and 3 of Proposition 5 predict that in her ex-post choice, the agent exhibits underinvestment into the future that she disapproves

18

of ex ante. When $T_b \ll T_i$, in contrast, Parts 2 and 3 predict that in her ex-ante commitment choice, she exhibits overinvestment into the future that she disapproves of ex post. The former prediction means that our model has the same basic implications as hyperbolic discounting in consumption-savings and borrowing choices, harmful consumption, exercise, and other "lifestyle" choices—day-to-day decisions that each have potential benefits over many future periods. These decisions include many of the prime examples invoked in the literature on hyperbolic discounting.[16] The latter prediction says that people are too prone to commit themselves to extended work for "achievements" they evaluate in terms of one or few attributes. This is consistent with the view of several researchers that many people commit themselves to overly ambitious careers relative to what would make them happiest.[17] Hence, in some well-defined situations our model provides a focus-based explanation for behavior resembling hyperbolic discounting, while at the same time identifying situations in which individuals act in an overly *future*-oriented way. We next develop new welfare and comparative-statics predictions emanating from these focus-based foundations for different types of time preference.

Welfare. The existing literature on present bias has not converged on an entirely satisfactory and well-founded view of the welfare ranking of different choices. One approach evaluates welfare using long-run preferences, imposing—exogenously—that a self's high relative weight on the present is always mistaken or irrelevant. Another approach applies the Pareto criterion to the intertemporal incarnations of the individual, often making no prediction about the welfare-optimal choice.[18] Going beyond both of these approaches, our model yields an unambiguous welfare ranking in all situations without a priori assuming that the ex-ante perspective measures welfare. 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. For instance, in Example 1 the agent's welfare is aligned with her

---

[16] See for instance Laibson (1997), Laibson, Repetto and Tobacman (2007), Shui and Ausubel (2004), Skiba and Tobacman (2008), and Meier and Sprenger (2010) on consumption-savings and borrowing, O'Donoghue and Rabin (1999a, 2006) and Gruber and Kőszegi (2001, 2004) on the consumption of harmful products, and DellaVigna and Malmendier (2004, 2006) and Acland and Levy (2010) on exercise.

[17] See Scitovsky (1976), Loewenstein, O'Donoghue and Rabin (2003), Kahneman et al. (2006), and Hamermesh and Slemrod (2008).

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

ex-ante preference to exercise, while in Example 2 her welfare is aligned with her ex-post preference not to write the book.[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 about 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. 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.[20]

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.

Finally, Part 3 of Proposition 5 implies that the degree of time inconsistency is increasing in the number of investment periods T_i. Intuitively, because T_i determines how many benefits considered separately ex post are integrated ex ante, it determines how much more the agent focuses on the benefits ex ante. A potentially important implication of this insight is that a person is more likely to commit to future-oriented behavior if the commitment applies to a substantial part of her future rather than only a trivial part of her future—because with the former type of commitment she

---

[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] Hyperbolic discounting says that this might not make much difference. In fact, if the long-run discount factor is less than 1, hyperbolic discounting predicts that the employee is more likely to do the five-hour job in the latter case, when she cannot put off the work for as long.

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 conclude this section by showing through an example that our model also predicts intuitively plausible circumstances in which the agent exhibits *future-oriented time inconsistency*—whereby she is *less* present-oriented ex post than ex ante:

Example 3 (Adding Achievements)**.** Suppose the scholar from Example 2 can not only write a book, but also an editorial for the newspaper. The editorial has an effort cost of 5 in periods 91 through 100, and a benefit of 45 in period 102.

First consider what the scholar does when making the book and editorial-writing decisions together ex ante. Then, the range of possible effort levels in periods $t = 91, ..., 100$ is 0 to 55, so the scholar's focus on effort ($g(55)$) is greater than her focus on the benefit of the editorial ($g(45)$). Hence, while she is (as in Example 2) likely to commit to writing the book, she prefers not to take on the editorial assignment.

Now consider whether the scholar would commit to the editorial in period 91, having already committed to the book. At this point, whether to take on the new assignment is equivalent to choosing between $(-50, -50, \ldots, -50, 0)$ and $(-55, \ldots, -55, 45)$. Hence, if $g(45)/g(5) > 50/45$, the scholar commits to the editorial as well. Intuitively, because the scholar's focus on effort is determined by the two projects' *total* costs in the ex-ante choice but only by the editorial's *incremental* costs in the ex-post choice, effort looms less important in the ex-post choice, so that she is more prone to agree to the editorial. Consistent with this prediction, many people—among them academics—appear to gradually take on more and more tasks that they themselves think are in sum making them worse off.

---

[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 the extreme prediction of no time inconsistency in single choices, we view the above comparative static as the more important prediction of our model.

4.2 General Conditions for Time-Inconsistent Behavior

In this section, we provide general conditions under which present-oriented time inconsistency and future-oriented time inconsistency arise in our model. Our results formalize a simple intuition. A person exhibits present-oriented time inconsistency if (as in Examples 1 and 2) the range of *future benefits* is narrower in ex-post than in ex-ante choice, because in that case she focuses relatively less on these future benefits in her ex-post decision. Conversely, a person exhibits future-oriented time inconsistency if (as in Example 3) the range of *investment costs* is narrower in ex-post than in ex-ante choice, because in that case she focuses relatively less on these costs in her ex-post decision. These results show that the basic intuitions about time-inconsistent behavior identified in Section 4.1 hold more generally as well.

To state our results, we first introduce some notation and concepts. Suppose there are κ attributes at time t, and let the corresponding utility functions be $u_{t1}(\cdot)$ through $u_{t\kappa}(\cdot)$, so that $u_t = \sum_{k=1}^{\kappa} u_{tk}$. We allow $u_{tk}(x_t, h_{t-1})$ to depend on both current and past consumption choices, but require that the consideration set X_t does not depend on previous choices. We continue to abstract away from time discounting and set $\delta = 1$.

To explore the question of time inconsistency, we identify the types of choices the agent prefers in single-stage deviations—where she is allowed to deviate from her ex-ante committed plan in period t, assuming that in all other periods she is still committed to the plan. Formally, for an ex-ante committed plan $(x_1^*, \ldots, x_T^*)$, we ask how, given history $x_1^*, \ldots, x_{t-1}^*$, the decisionmaker's choice in the continuation problem $X_t \times \{x_{t+1}^*\}, \ldots, \times \{x_T^*\}$ compares to x_t^*.

Our first result identifies circumstances under which the agent prefers to behave in a more present-oriented way in a single-stage deviation than she would commit to ex ante. We say that the current utility of a current action is not controlled by previous choices if, whenever u_{tk} depends on x_t, it does not depend on the history h_{t-1}. We say that period-t choice involves a clean tradeoff between present and future if whenever an action at t increases utility in period t, it reduces utility in all future periods. Formally, for any $x_t, x_t' \in X_t$, either $u_{tk}(x_t', h_{t-1}) \geq u_{tk}(x_t, h_{t-1})$ for all k and $u_{\tau k}(h_{\tau, -t}, x_t') \leq u_{\tau k}(h_{\tau, -t}, x_t)$ for all $\tau > t$, k, and feasible x_{t+1} through x_T, or vice versa. Both of these conditions are satisfied by the investment problem of Section 4.1. Proposition 6 generalizes

the time inconsistency predicted by our model in that problem:

Proposition 6 (Present-Oriented Time Inconsistency). *Take any $t \in \{2, \ldots, T\}$, and suppose that (i) the current utility of current actions in period t is not controlled by previous choices; and (ii) period-t choice involves a clean tradeoff between present and future. Then, the decisionmaker would prefer a single-stage deviation that 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 her period-t choice equally ex ante and 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 in period t than ex ante. As a result, she pays (weakly) less relative attention to the future in period t than ex ante, so she acts in a (weakly) more present-oriented manner.

We next identify conditions under which the agent behaves in a *less* present-oriented way in a single-stage deviation than she would commit to ex ante. We say that a future attribute τ^*, k^* is fully controlled by period-t choice if $u_{\tau^* k^*}$ depends only on x_t. This attribute involves a clean tradeoff with all other attributes if, for any x_t and x_t' such that $u_{\tau^* k^*}(x_t') > u_{\tau^* k^*}(x_t)$, for any other τ, k with $\tau \geq t$ it must be that $u_{\tau,k}(x_\tau, x_{\tau-1}, \ldots, x_{t+1}, x_t', h_{t-1}) \leq u_{\tau,k}(x_\tau, x_{\tau-1}, \ldots, x_{t+1}, x_t, h_{t-1})$. When these conditions are satisfied, the attribute (τ^*, k^*) can be thought of as a concentrated achievement that may require multiple periods of sacrifice the agent can commit to in period t, as in Example 3. Proposition 7 generalizes the time inconsistency in that example:

Proposition 7 (Future-Oriented Time Inconsistency). *Consider any $t \in \{1, \ldots, T-1\}$, and suppose that (i) there is a future attribute (τ^*, k^*) fully controlled by period-t choice, and (ii) this attribute involves a clean tradeoff with all other attributes. Then, the decisionmaker would prefer a single-stage deviation that yields at least as high utility in attribute (τ^*, k^*) as her ex-ante plan.*

Because of (i), the range of consumption utility from the achievement is the same ex ante and in period t. But because the agent has fewer choices in period t, the range of costs is (weakly) smaller in period t than ex ante. Hence, the agent pays (weakly) more relative attention to the future benefits in period t than ex ante, leading to more future-oriented behavior.

5 Product Design and Pricing

In this section, we explore some implications of our model for profit-maximizing firms' product design and pricing behavior, helping to incorporate into economics a key marketing question: how firms design and position products to manipulate consumer attention. 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, 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. By fully specifying how product features determine consumer focus, we can make specific predictions on the features firms will use to influence consumers' focus.

Throughout this section, we assume that there is a representative consumer in the market with known preferences. A product is a vector $(v_1, ..., v_{K_1}, -p_1, ..., -p_{K_2})$, where the first K_1 dimensions are value attributes and the last K_2 dimensions are price attributes, and the consumer's consumption utility is the sum of her consumption levels in the $K_1 + K_2$ attributes. In particular, our formulation means that if $K_2 > 1$, firms can separate the price into multiple components that the consumer treats as distinct attributes. This assumption reflects the idea, most clearly elaborated in Thaler's (1985, 1999) theory of mental accounting, that individuals often do not integrate different monetary transactions in their minds. Thaler and Johnson (1990) argue, for instance, that temporal separation facilitates the separation of monetary transactions in a person's mind, but payments can be framed as separate by or for the consumer in other ways as well.[22]

A firm selling the above product makes a profit of

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

Hence, a 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

---

[22] The assumption that the consumer evaluates different payments as different attributes is entirely consistent with our consumption-utility-based model, as well as with the assumption of the fungibility of money *in consumption utility*. Specifically, if for pre-existing wealth W we define consumption utility in money as $U_m(W - p_1 - \cdots - p_{K_2})$, and assume a linear $U_m(\cdot)$, we get the reduced-form specification above. Assuming a strictly concave $U_m(\cdot)$ yields a model in which utility is not additively separable in attributes.

differentiable and satisfies $\phi'(0) = 0, \phi''(c) > 0$ for all $c \geq 0$, as well as $\lim_{c \to \infty} \phi'(c) = \infty$.

Our assumptions that there is a single representative consumer and that value attributes are perfectly substitutable in production and consumption make our results starker and easier to state. Because the basic forces we identify do not seem to depend on these assumptions, we expect that our qualitative behavioral predictions extend to settings with heterogenous consumers and imperfectly substitutable attributes. In these settings, stronger welfare losses are likely to emerge due to the misallocation of goods and factors of production.

5.1 Product Design Under Monopoly

We first analyze how a monopolist designs and prices products to take advantage of consumers' bias toward concentration. Suppose that a monopolist can sell exactly one product, and that the outside option of the consumer is not purchasing, which is represented by the vector $(0, ..., 0)$. Then:

Proposition 8 (Price Splitting and Value Concentration). *Suppose that $g(\cdot)$ is differentiable. The optimal product has $v_k > 0$ for exactly one $k \in \{1, \ldots, K_1\}$, $v_j = 0$ for all $j \neq k$, and $p_k > 0$ for all $k = 1, \ldots, K_2$. If $K_2 = 1$, the level of value production is efficient ($\phi'(v_k) = 1$). If $K_2 > 1$ and $G(\Delta) \equiv g(\Delta)\Delta$ is strictly convex, there is overproduction of value ($\phi'(v_k) > 1$).*

Because the consumer focuses more on and hence is repelled by concentrated disadvantages, the monopolist splits the price into multiple components. And because the consumer focuses more on and hence is attracted by concentrated advantages, the monopolist concentrates the value on a single attribute. To understand the efficiency implications, first note that given the concentrated value, 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 overweights the single large value relative to the split prices, so the monopolist produces inefficiently high value.[23]

---

[23] To understand the role of the extra assumption on $G(\Delta)$ in this last result, first note that under Assumption 1, $G(\Delta)$ is strictly increasing and convex near zero. We make the natural assumption that this property extends to the relevant range of medium stakes. Intuitively, this 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. Then, the consumer responds more to adding to the concentrated value of the product than to adding to the dispersed price, so that the monopolist increases the value beyond the efficient level.

Proposition 8's key predictions are supported by some evidence from economics and marketing. Artificial price splitting, whereby the price is separated into multiple components over which the consumer has no discretion, is frequently used in practice. In one specific setting, Woodward and Hall (2010) argue that borrowers underestimate mortgage-broker compensation (an underestimation the authors document) partly because such compensation is artificially split into many fees. Similarly, Gourville (1998) notes retailers such as magazines or charities often frame aggregate expenses as a series of small "pennies-a-day" (per-issue or daily) payments. And although these charges may be less salient not only because of our model, but simply because they are hidden, the observation that retailers often do not include taxes or shipping costs in their posted prices is also consistent with price splitting. Experiments document that all these strategies work in increasing demand.[24]

In addition, although it serves other purposes as well, financing may be so common in part due to our model's prediction of price splitting: as in the lead-off example of Section 3, a laptop that costs "$39 a month" may seem less expensive than one with a price of $899. Because our model's prediction 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).[25] And although they have an efficiency rationale and can in principle be avoided by consumers, the multitude of fees imposed by firms for additional services may also in part be due to price splitting. For instance, credit-card issuers, banks, mobile-phone companies, and airlines make a large part of their profits from imposing many relatively small fees that may not seem like much to consumers when getting the product, but that can easily add up to significant amounts.[26]

Finally, although the reasons given are varied and sometimes not clearly related to focusing, Proposition 8's prediction of value concentration is reminiscent of views in the marketing literature

[24] See Morwitz, Greenleaf and Johnson (1998), Chakravarti, Krish, Paul and Srivastava (2002) and Xia and Monroe (2004) for artificial price splitting, Gourville (1998) for pennies-a-day strategies, Chetty, Looney and Kroft (2009) for posting pre-tax versus post-tax prices, and Hossain and Morgan (2006) for shipping costs.

[25] 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.

[26] Agarwal et al. (2007, 2008) and Stango and Zinman (2009) document that consumers pay very significant amounts in banking and credit-card fees. For mobile phones, the Federal Communication Commission is considering regulation to help consumers avoid the "bill shock" from unexpectedly large bills due to many minor charges (Federal Communication Commission 2010).

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" to consumers is most 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 product with many characteristics is counterproductive.

5.2 Competition and Relative Positioning

We next explore how competition affects firms' incentives to position products and improve product quality. Firms often seem to use the positioning strategy of building 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. 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: what determines which strategy is optimal?

To address this 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 ($K_1 = 2$) 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. Firm 2 sells an exogenously given competing product with qualities $\bar{v}_1$, $\bar{v}_2 < \bar{v}_1$, and price $\bar{p}$.[27] We assume that the consumer's consideration set includes these two products as well as the option of buying neither product. Under a somewhat stronger assumption on $g(\cdot)$, qualitatively similar results hold if the consideration set did not include not buying.[28]

---

[27] Because we explore only firm 1's best response rather than a full game in which both firms act strategically, our analysis identifies only whether different kinds of firms would ideally *like* to differentiate their products, and how these incentives play out in equilibrium depends on the precise environment (e.g., timing and technologies) in which firms operate. In a simultaneous-move game, for instance, the differing incentives of strong and weak firms to match attributes would generate a mixed-strategy equilibrium.

[28] Specifically, in that case our basic finding—that a strong firm wants to be strongest on its competitor's weak attribute, while a weak firm prefers the opposite—holds if the function $G(\Delta)$ defined in Proposition 8 is strictly

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

Proposition 9 (Incentives to Differentiate). *Suppose $g(\cdot)$ is differentiable.*

1. If $v_i < \bar{v}_i$ for $i = 1, 2$, then $v_1 > 0$, $v_2 = 0$.

2. If $v_1 + v_2 > \bar{v}_1 + \bar{v}_2$ and (holding the other parameters constant) either $g(v_1 + v_2)$ is sufficiently large or $\bar{v}_2$ is sufficiently small, then $v_1 = 0$, $v_2 > 0$.

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 produces on attribute 1 only ($\bar{v}_2 = 0$), 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 v to attribute 1 is optimal: it maximizes focus on firm 1's product without affecting focus on the competitor's product. If $v > \bar{v}_1$, however, then allocating 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.[29]

The setting explored in Proposition 9 is one example where a richer model would generate additional welfare effects. While under our current assumptions it does not matter whether a firm produces a given value on the same attribute as the competitor or a different one, this would matter for instance if consumers had heterogeneous preferences for product attributes. Then, focus-driven product design may result in a suboptimal product mix for consumers.

Having analyzed a firm's incentive to differentiate its product, to conclude this section we explore how quality levels depend on whether firms produce value on the same attribute or different attributes. In "attentionally differentiated" competition, $K_1 = 2$ and firm $i = 1, 2$ can only produce attribute i. In "attentionally undifferentiated" competition, $K_1 = 1$ and both firms produce on

---

convex. For a discussion of this assumption, see Footnote 23.

[29] The intuition is somewhat more complicated if $\bar{v}_2 > 0$. In that case, firm 2 benefits if firm 1 allocates v and hence draws attention to attribute 2. The additional conditions in Part 2 of Proposition 9 guarantee that the effect discussed above dominates this effect: as $\bar{v}_2$ approaches zero, this effect vanishes, and as $g(v_1 + v_2)$ increases without limit, the above effect becomes arbitrarily strong.

this attribute. In both cases, we assume that there is only one payment dimension ($K_2 = 1$). We also assume that the consumer's consideration set includes both options as well as not buying, but a qualitatively similar result holds even if not buying is not in the consideration set. Firms simultaneously choose quality levels and price, with firm i aiming to maximize $p_i - \phi(v_i)$. Then:

Proposition 10 (Differentiated versus Undifferentiated Competition). *Suppose $g(\cdot)$ is differentiable. For any symmetric pure-strategy equilibrium, value is higher with attentionally differentiated than with attentionally undifferentiated competition, and in either case is 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 but not firm 2. With undifferentiated competition, however, any increase in the consumer's focus benefits both 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 interpreted, Proposition 10 says that in industries in which product differentiation is easier, production of value should be higher. 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. But following the introduction of mobile phones—which turned the phone into a personal item and opened the scope for greater differentiation—telephone technology improved rapidly.

6 Conclusion

By virtue of defining focus-dependent utility based on consumption utility and the decisionmaker's consideration set, our theory opens the way for analyzing the role of focus in many economic settings

using one generally applicable model. Besides intertemporal choice and product design, which we have started to explore in this paper, there seem to be many other 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 attracting focus to their candidates' strengths. Furthermore, our basic finding that individuals like concentrated payments to them but dislike concentrated payments to others may have many implications in settings beyond those in this paper. For example, research indicates that retirees take too much of their retirement wealth in one lump sum rather than as an annuity,[30] and our model says that this may simply happen because a lump-sum payment because it looks very large relative to an annuity's monthly payments.

Being defined for riskless choices only, our model is not applicable to situations in which uncertainty is a central part. Uncertainty raises a new conceptual issue for our model: whether to think of states or probabilities (or both) as the relevant attributes. To illustrate the difference, suppose that smoking a cigarette increases the probability of developing lung cancer at later dates by a tiny fraction. If different states are different attributes, as in Bordalo et al. (2010) and Bordalo (2011), our model says that the large decrease in consumption utility in certain states draws the consumer's focus to the possibility of lung cancer, making smoking aversive. But if the probabilities are the attributes, the trivial change in probability leads the consumer to underweight the possibility of lung cancer, making smoking much more attractive. In ongoing work (Kőszegi and Szeidl 2011, hopefully), we attempt to combine these possibilities.

---

[30] 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). 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).

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 Salience Theory," 2011. Working Paper, Harvard University.

___ , **Nicola Gennaioli, and Andrei Shleifer**, "Salience 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.

Chakravarti, Dipankar, Rajan Krish, Pallab Paul, and Joydeep Srivastava, "Partitioned Presentation of Multicomponent Bundle Prices: Evaluation, Choice and Underlying Processing Effects," *Journal of Consumer Psychology*, 2002, *12* (3), 215–229.

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," *Theoretcial Economics*, 2011, *6*, 127–155.

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 Kőszegi, "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).

Hossain, Tanjim and John Morgan, "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay," *B.E. Journal of Economic Analysis and Policy: Advances in Economic Analysis and Policy*, 2006, *6* (2), 1–27.

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.

Kőszegi, 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.

Morwitz, Vicki G., Eric A. Greenleaf, and Eric J. Johnson, "Divide and Prosper: Consumers' Reactions to Partitioned Prices," *Journal of Marketing Research*, 1998, *35* (4), 453–463.

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** ___ , "Doing It Now or Later," *American Economic Review*, 1999, *89* (1), 103–124.

___ **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.

Read, Daniel, George Loewenstein, and Matthew Rabin, "Choice Bracketing," *Journal of Risk and Uncertainty*, 1999, *19* (1-3), 171–197.

Reis, Ricardo, "Inattentive Producers," *Review of Economic Studies*, 2006, *73* (3), 793–821.

Salant, Yuval, "Procedural Analysis of Choice Rules with Applications to Bounded Rationality," *American Economic Review*, 2011, *101* (2), 724–748.

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.

—— , "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, "Intransitivity of Preferences," *Psychological Review*, 1969, *76* (1), 31–48.

—— , "Features of Similarity," *Psychological Review*, 1977, *84* (4), 327–352.

—— **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.

Xia, Lan and Kent B. Monroe, "Price Partitioning on the Internet," *Journal of Interactive Marketing*, 2004, *18* (4), 63–73.

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

$$
\tilde{U}\left(c^1, C\right) - \tilde{U}\left(c^2, C\right) = \sum_{i \in A(c^1, c^2)} g(\Delta_i(C)) \cdot \Delta_i(C) - \sum_{j \in A(c^2, c^1)} g\left(\Delta_j(C)\right) \cdot \Delta_j(C)
$$

$$
\geq g\left(\min_{i \in A(c^1, c^2)} \Delta_i(C)\right) \cdot \sum_{i \in A(c^1, c^2)} \Delta_i(C) - g\left(\max_{j \in A(c^2, c^1)} \Delta_j\right) \cdot \sum_{j \in A(c^2, c^1)} \Delta_j(C)
$$

$$
> g\left(\min_{i \in A(c^1, c^2)} \Delta_i(C)\right) \cdot \sum_{i \in A(c^1, c^2)} \Delta_i(C) - g\left(\min_{i \in A(c^1, c^2)} \Delta_i(C)\right) \cdot \sum_{j \in A(c^2, c^1)} \Delta_j(C) = 0,
$$

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

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

Proof of Proposition 3. Immediate. □

Proof of Proposition 4. The focus-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 $\epsilon \to 0$ 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. □

Proof of Proposition 5.

1. Attention on attribute t is set by $\Delta_t = B$, while attention on attributes $s = T_i + 1, \ldots, T_i + T_b$ is determined by $\Delta_s = AB/T_b$. The consumer strictly prefers to invest if $Bg(B) < ABg(AB/T_b)$. Since the right-hand side is continuous and increasing in A, the value A^*_{post} is defined by $g(B) = A^* \cdot g(A^* B/T_b)$. 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 attributes is set by $\Delta_t = B$, $t = 1, \ldots, T_i$ while focus on the benefit attributes is defined by $\Delta_s = AB \cdot T_i/T_b$ for $s = T_i + 1, \ldots, T_i + T_b$. Effort is strictly preferred if and only if $Bg(B) < AB \cdot g(AB \cdot T_i/T_b)$, and A^*_{ante} is defined by $g(B) = A^* \cdot g(A^* B \cdot T_i/T_b)$. The claims now all follow from the strict monotonicity of $g(\cdot)$.

3. Immediate from the above expressions. □

Proof of Proposition 6. Consider $x'_t \in X_t$ such that $u_t(x'_t, h^*_{t-1}) \leq u_t(x^*_t, h^*_{t-1})$. We show that in the ex-post decision of period t, the decisionmaker prefers x^*_t to x'_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 x_t. Both ex ante and ex post, the choice of x_t only affects period-t attributes in $\tilde{K}$ and future attributes in periods

$\tau > t$. The fact that ex ante the agent chooses x_t^* rather than x_t' means that

$$\sum_{k \in \tilde{K}} g(\Delta_{tk}) \left[u_{tk}(x_t^*, h_{t-1}^*) - u_{tk}(x_t', h_{t-1}^*) \right] \geq \sum_{k, \tau > t} g(\Delta_{\tau k}) \left[u_{\tau k}(x_t', h_{\tau,-t}^*) - u_{\tau k}(x_t^*, h_{\tau,-t}^*) \right].$$

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. □

Proof of Proposition 7. Consider $x_t' \in X_t$ such that $u_{\tau^* k^*}(x_t') \leq u_{\tau^* k^*}(x_t^*)$. We show that in the t-only choice, the decisionmaker prefers x_t^* to x_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 x_t^* rather than x_t' means that

$$g(\Delta_{\tau^* k^*}) \left[u_{\tau^* k^*}(x_t^*) - u_{\tau^*, k^*}(x_t') \right] \geq \sum_{(\tau, k) \neq (\tau^*, k^*), \tau > t} g(\Delta_{\tau k}) \left[u_{\tau k}(x_t', h_{\tau,-t}^*) - u_{\tau k}(x_t^*, h_{\tau,-t}^*) \right].$$

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. □

Proof of Proposition 8. Recall that $G(\Delta) = \Delta g(\Delta)$. The monopolist maximizes

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

subject to

$$\sum_{i=1}^{K_2} G(p_i) \leq \sum_{i=1}^{K_1} G(v_i). \tag{3}$$

Note that because $g(\cdot)$ is strictly increasing, for any $x, y > 0$ we have $G(x + y) > G(x) + G(y)$.

We prove by contradiction that having $v_i, v_j > 0$ for two distinct i, j cannot be optimal. Suppose that this is the case. Instead choosing value $v_i + v_j$ on attribute i and value zero on attribute j

37

increases the right-hand side of the constraint 3 due to the inequality on $G(\cdot)$ above, allowing the monopolist to increase prices without increasing costs, a contradiction. Given that ϕ' is unbounded, it is now easy to see that choosing a very large value for any v_i is dominated by setting all $v_i = 0$. It follows that the above problem can be viewed as maximizing a continuous function on a compact set and hence has a well-defined optimum.

Because $\phi'(0) = 0$, the monopolist chooses $\sum_{i=1}^{K_1} v_i > 0$ and $\sum_{i=1}^{K_2} p_i > 0$. We now prove by contradiction that $p_k > 0$ for all price attributes k. Suppose that $p_i = 0$ and $p_j > 0$. Due to the inequality on $G(\cdot)$ above, instead choosing $p_j/2$ on both attributes lowers the left-hand side of constraint 3, allowing the monopolist to increase the total price without increasing costs, a contradiction.

Now we turn to the efficiency implications. Given that the monopolist produces value on exactly one attribute, when $K_2 = 1$ the consumer faces a balanced choice, and hence maximizes consumption utility. As a result, the monopolist chooses value as for a classical consumer, i.e., efficiently.

Now suppose $K_2 > 1$. 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 $g(\cdot)$ is strictly increasing, the constraint implies that $v > p$ whenever $K_2 > 1$. Since this problem has a solution and it is not at $v = 0$, it must be interior. Differentiability of the objective function and the constraint then implies that the optimum must satisfy the first order condition, which, using the Lagrangian, can be written as

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

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. $\qquad \square$

Proof of Proposition 9. (i) By assumption, the focus weights are set by $\bar{v}_1$ and $\bar{v}_2$, and hence the consumer focuses more on attribute 1. Suppose that $v_2 > 0$, 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 focus weights and hence the perceived value of firm 2's product. But this change does move more of firm 1's value on the attribute which enjoys more attention, increasing the perceived value of that product.

(ii) Suppose by way of contradiction that $v_1 > 0$. 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 results in 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 between the latter and former expressions 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.

Finally, we briefly show how our results change when the consideration set includes only the two products. We assume that $G(\Delta)$ is a convex function. First, we establish a slightly weaker analogue of (i): that if $v_i < \bar{v}_i$ for $i = 1, 2$, then $v_1 > v_2$. Consider the optimal behavior of firm 1 for a fixed value of $v_1 + v_2 = v$. We denote $\bar{v}_1 + \bar{v}_2 = \bar{v}$. Note that under the assumption in the proposition, the relative disadvantage of firm 1's product is $G(\Delta_1) + G(\Delta_2)$ where $\Delta_1 + \Delta_2 = \bar{v} - v$, and firm 1's objective is to minimize this subject to $\Delta_i \leq \bar{v}_i$. Due to convexity, in an interior optimum we must have $G'(\Delta_1) = G'(\Delta_2)$ or $\Delta_1 = \Delta_2$, so that $v_1 > v_2$. This interior optimum is not feasible if $\Delta_2 > \bar{v}_2$. In that case all feasible allocations have $\Delta_1 > [\bar{v} - v]/2$, and the minimum cost feasible allocation is when Δ_2 is as high as possible, or $v_2 = 0$. Thus, again $v_1 > v_2$. Next, we establish a stronger analogue of (ii): that if $v_1 + v_2 > \bar{v}_1 + \bar{v}_2$, then $v_1 = 0, v_2 > 0$. Consider an optimum and label attributes by i, j such that $\Delta_i \geq \Delta_j$. Then we must have $v_i > \bar{v}_i$. Consider the deviation of marginally increasing v_i and marginally reducing v_j. We claim that the marginal benefit to firm i from this deviation is $G'(\Delta_i) - G'(\Delta_j)$. This is immediate when $v_j > \bar{v}_j$ in which case the relative advantage of firm 1's product is $G(\Delta_i) + G([v - \bar{v}] - \Delta_i)$. When $v_j < \bar{v}_j$, note that $\Delta_j = \bar{v}_j - v_j = \bar{v} - v - (\bar{v}_i - v_i) = \Delta_i - [v - \bar{v}]$ and hence the relative advantage to firm i is $G(\Delta_i) - G(\Delta_j) = G(\Delta_i) - G(\Delta_i - [v - \bar{v}])$, so that the claim again follows. Since G is convex, by $\Delta_i \geq \Delta_j$ the gain is increasing in Δ_i, hence firm 1 optimally chooses a corner solution of allocating all value on a single attribute. The same argument also implies that allocating all value to the second attribute is better, because this is the case when Δ_i and hence $G(\Delta_i) - G(\Delta_j)$ is largest. $\qquad \square$

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

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

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

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

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

Finally, we sketch a proof establishing that the results extend when the consideration set of the buyer only includes the two products. First consider undifferentiated competition. Following

41

similar steps as given above yields the condition $\phi'(v) = 1$. Intuitively, the consumer now faces a balanced choice, chooses rationally, and hence the firm maximizes value. Next consider differentiated competition. Now we obtain $\phi'(v) = (g(v) + g'(v)v)/g(0)$. The denominator changes relative to (5) because now the price attribute receives a focus weight of $g(0)$. Comparing these expressions, we still find higher value under attentionally differentiated competition, and (weakly) higher value than the monopolist, who, like the firms in attentionally undifferentiated competition produces efficiently. □

B Eliciting Model Ingredients from Behavior

In this section, we outline an algorithm for eliciting the utility functions $u_k(\cdot)$ and focus-weight function $g(\cdot)$ from behavior by observing choices from a number of specifically chosen consideration sets. Once these ingredients are elicited, our model provides a prediction on both behavior and welfare for any finite consideration set. Note that to elicit the utility function in a classical model of individual choice based on axiomatic foundations, one would also be required to observe choices in a number of consideration sets (typically equated with choice sets, as in our elicitation). Hence, with the caveat that we also need to know the relevant attributes, our model is as falsifiable as a classical model.

Our elicitation works by first eliciting consumption utility using the fact that the agent makes consumption-utility-maximizing choices in balanced decisions, and then eliciting $g(\cdot)$ by measuring the agent's bias toward big differences. We assume (i) that we know how products map into attributes and there are at least three attributes; (ii) that we can manipulate individual attributes in the decisionmaker's choices; (iii) that $g(\cdot)$ is strictly increasing. For simplicity, we also assume that (iv) the utility functions u_k are differentiable. Without loss of generality, we normalize $u_k(0) = 0$ for each k and $u_1'(0) = 1$. Since $g(1) > 0$, we can also without loss of generality normalize $g(1) = 1$.

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 $p \to 0$ gives

$$u_1'(0) = u_k'(x)\delta'(0).$$

This procedure elicits $u_k'(x)$, and hence (using the normalization that $u_k(0) = 0$) the entire utility function $u_k(\cdot)$. We can then use the elicited utility function for some $k > 0$ to similarly elicit the entire utility function $u_1(\cdot)$.

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 $\{(0, 0, x_0), (1, x - p, 0), (1 - \delta, x, 0\}$. The component x_0 can take any value sufficiently low for the agent not to choose the first option.[31] For some p satisfying $0 < p < x$, we find the $\delta_x(p)$ that makes the decisionmaker indifferent between the last two options. Hence

$$g(1) \cdot 1 + g(x) \cdot (x - p) = g(1) \cdot (1 - \delta_x(p)) + g(x) \cdot x,$$

or

$$g(x) = \frac{\delta_x(p)}{p} \cdot g(1) = \frac{\delta_x(p)}{p}.$$

To conclude, we note that if an observer knows the set of potential attributes and is able to manipulate these potential attributes separately, then our model can be used to identify whether two potential attributes are distinct or form part of the same attribute. As a result, if the pool of potential attributes known to the observer contains all true attributes, then it is possible to identify the true attributes from observational data.

---

[31] The sole role of including a third attribute and setting $x_0 > 0$ is to allow for the possibility that if the first option is clearly dominated by the other alternatives, it might not make the agent's consideration set. If the first option is sufficiently superior in the third attribute, it will presumably be in the consideration set, even if it is eventually not chosen.

For an example that is relevant to some of our applications, suppose that an employee may receive an increase in her bonus in December and another increase in her bonus in June. These two bonuses are two potential attributes, and we would like to identify whether the employee treats them as distinct attributes or as parts of the same attribute, her total annual bonus. Assuming that there is some attribute that we know is distinct from these payments, the bias towards concentration allows us to elicit how the employee treats the two payments. Intuitively, if the two payments are parts of the same attribute, then offering them jointly will make them more attractive then offering first the December and then the June payment.

Consider the following elicitation procedure. (i) We find the hours worked on one day, e, such that the agent is indifferent regarding whether or not to work e for the December payment increase W_D, assuming no payment increase in June. (ii) Given hours e and December payment W_D, we find the additional number of hours on the same day, e', such that the agent is indifferent regarding whether or not to work e' to also get the June payment increase W_J. (iii) Given these data, we can identify if the two attributes are distinct by asking the employee whether she accepts working $e + e'$ hours today in exchange for *both* future payments. If the payments are the same attribute, the employee will be indifferent, while if the payments are separate attributes, then she will prefer not to take the job.

To see the logic formally, begin with the case when the two payments are different attributes. Assuming that the utility of no payment increase or zero work is normalized to zero, the first elicitation, being a balanced trade-off, yields $-u_1(e) = u_D(W_D)$. The second elicitation, also a balanced trade-off, yields $-(u_1(e + e') - u_1(e)) = u_J(W_J)$. Thus the comparison in (iii) is between $(-u_D(W_D) - u_J(W_J), u_D(W_D), u_J(W_J))$ and $(0, 0, 0)$ and hence the second option is strictly preferred by Proposition 1. Now consider the case when the two payments are on the same attribute. The first elicitation still yields $-u_1(e) = u(W_D)$, but the second elicitation now gives $-(u_1(e + e') - u_1(e)) = u(W_D + W_J) - u(W_D)$. The comparison in (iii) is now between $(-u(W_D + W_J), u(W_D + W_J))$ and $(0, 0)$ and hence by Proposition 3 the employee is indifferent.

Some assumptions in our model, however, cannot be elicited from choice behavior, and therefore must come from outside our theory. Clearly, our theory does not offer a way to for a modeler to

formulate a set of potential attributes for a particular situation. In addition, even if the set of potential attributes is known, identifying the true attributes from behavior clearly requires us to manipulate attributes separately. 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. In this case, it seems to us to be a natural assumption that a person would view essentially identical attributes as one, so that there is only a single book achievement attribute.