A Model of Focusing in Economic Choice[*]

Botond Kőszegi
UC-Berkeley

Adam Szeidl
Central European University,
UC-Berkeley, NBER, CEPR

March 8, 2012

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, the decisionmaker exhibits present bias and time inconsistency when—such as in lifestyle choices and other widely invoked applications of hyperbolic discounting—the future costs of current misbehavior are distributed over many dates, and the effects of multiple decisions accumulate. But unlike in previous models, (1) present bias is lower when the costs of current misbehavior are less dispersed, helping to explain why individuals respond more to monetary incentives than to health concerns in harmful consumption; and (2) time inconsistency is lower when the ex-ante choice integrates fewer decisions with accumulating effects. In addition, the agent does not fully maximize welfare even when making decisions ex ante: (3) she commits to too much of an activity—e.g., exercise or work—that is beneficial overall; and (4) makes "future-biased" commitments when—such as in preparing for a big event—the benefit of many periods' effort is concentrated in a single goal.

JEL Codes: D03, D40, D91

1   Introduction

People often focus disproportionately on, and hence overweight, certain attributes of their available options. For example, 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). And a shopper deciding whether to buy an unhealthy item may focus more on the item's price than on its health consequences, and hence be overly responsive to price relative to nutritional information (Abaluck 2011). The theoretical implications of such phenomena for economic questions involving multi-attribute choice are largely unexplored and unclear.

Building on evidence and related work we discuss in Section 2, especially approaches by Tversky (1969), Loomes and Sugden (1982), and Bordalo, Gennaioli and Shleifer (2012, forthcoming), in this paper 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 relevance of this determinant of focus in one economically important application, choice over time, showing that the model unifies seemingly conflicting evidence and intuitions, helps explain facts which are puzzling from the perspective of existing theories, and generates new economic insights. We also discuss other applications, and argue that because our model makes a prediction in any choice problem in which a classical model would, it provides a simple way to incorporate focusing into many more economic settings.

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 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 $U(c) = \sum_{k=1}^{K} u_k(c_k)$. But instead of consumption utility, the decisionmaker acts to maximize *focus-weighted utility* $\sum_{k=1}^{K} g_k u_k(c_k)$, 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 $\Delta_k(C) = \max_{c \in C} u_k(c_k) - \min_{c \in C} u_k(c_k)$ and $g(\cdot)$ is an increasing function. Because there 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.

In Section 3, we identify some general implications of our model that we use repeatedly in our analysis of choice over time, and that are likely to be important in other applications as well. Because the decisionmaker focuses too much on a large advantage relative to multiple small disadvantages, she exhibits a "bias towards concentration": she is too prone to choose options with a single advantage and multiple disadvantages relative to all alternatives, and more generally to choose options with advantages concentrated in fewer attributes. This property helps identify ways in which choices are biased in specific applications. But because the overly large focus on an advantage relative to a single smaller disadvantage just reinforces the advantage, the agent maximizes consumption utility in "balanced choices"—decisions in which the advantages and disadvantages of relevant options are equally concentrated. This observation allows one 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—and provides guidance in applications as to which choices better reflect welfare.

In Section 4, we explore one economically important application of our framework, intertemporal choice. We extend our basic setup to choice over time by assuming that utility outcomes on different dates correspond to different attributes, and 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. To isolate the implications of focusing alone, we assume no discounting in consumption utility, and for this reason our theory does not deliver a key prediction of hyperbolic-discounting models: that the agent exhibits present bias and time inconsistency in single balanced choices such as whether to perform a one-period task earlier or later. Nevertheless, we show that our theory makes the same predictions as hyperbolic discounting in exactly those field settings in which the latter model has been most invoked and documented; explains evidence about variation in present bias across situations which existing models cannot explain; and makes several plausible—although as yet untested—new predictions. Hence, while in itself not a complete theory as it does not match laboratory evidence on time inconsistency in single choices, we believe

that our model identifies a mechanism that is one potentially important ingredient of intertemporal choice in practice.

To be more specific, our theory predicts present bias and time inconsistency in intertemporal decisions where—such as in the case of exercise, harmful consumption, other lifestyle decisions, and consumption-savings—a large sacrifice today results in small per-period benefits for many periods in the future, and the benefits of repeated sacrifices accumulate. In a logic reminiscent of Akerlof's (1991) and Rick and Loewenstein's (2008) informal arguments that the benefits of present-oriented behavior often feel more tangible than the costs, our theory predicts that 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, and hence—exhibiting present bias—she tends to exercise too little. When considering the entire sequence of decisions ex ante, however, the person focuses on the large fitness gains she will enjoy every day if she exercises regularly, so—exhibiting time inconsistency—she prefers to exercise more.

Besides being consistent with the some of the key field evidence on present bias, our theory also generates new behavioral and welfare predictions. Most importantly, it predicts that—by reducing the per-period costs and hence the focus on these costs—an increase in the number of periods in which the future costs of current misbehavior are dispersed increases present bias. This comparative static provides a new testable prediction and may explain some empirically documented variation in present bias across situations that hyperbolic (and exponential) discounting models cannot explain. One central fact, documented for instance by Gruber and Kőszegi (2004), Volpp et al. (2008), and Abaluck (2011), is that harmful consumption is quite responsive to prices and other monetary incentives. Because for many harmful products the health consequences of consumption far outweigh the financial consequences, the logical implication of this fact in hyperbolic discounting and related models is that consumers are extremely responsive to changes in, or information regarding, the health consequences of consumption. We argue that this large responsiveness is empirically implausible and unobserved. But the combination of high responsiveness to monetary incentives and low responsiveness to health consequences is fully consistent with our model if we make the reasonable assumption that individuals think of monetary payments partly as separate

3

attributes, so that the financial incentives are concentrated but the health consequences are not.

In addition, our theory makes a novel prediction regarding the extent of time inconsistency in repeated decisions with accumulating effects. The more decisions the consumer is considering committing to, the larger—and hence the more salient—are the accumulated benefits in her ex-ante choice, and therefore the more time inconsistent she is. One 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 is more likely to have a noticeable impact on her life. While other plausible explanations are also possible, we argue that this prediction is broadly consistent with existing evidence on the take-up of commitment devices in the field.

Finally, because our model is based on a single well-defined utility function that reflects welfare, it helps sort out the welfare implications of time inconsistency. In particular, our theory often supports the commonly used presumption in the literature that ex-ante preferences are appropriate for welfare analysis, but also qualifies this perspective in specific ways. For the above types of consumption-savings and lifestyle choices trading off a concentrated cost in the present with many small benefits in the future, our theory says that—it being a more balanced choice—ex-ante behavior usually better reflects welfare. But because focus is determined by the total rather than the marginal benefit of effort, our model also implies that a person may "overcommit" to worthwhile tasks with decreasing marginal benefit. In one economically important example, a worker who faces attractive job prospects may focus more on the consumption benefit than on the effort cost of work, and hence agree to working too much. And in ex-ante choices where—such as when planning a wedding or preparing for a marathon—a sequence of sacrifices leads to a single large and hence attention-grabbing goal, our theory predicts "future bias" even for binary ex-ante choices or tasks with constant marginal benefits. For example, if an employee evaluates monetary payments as separate attributes, she may be excessively motivated by bonuses and similar concentrated payments she must work for over an extended period of time. In this case, while predicting the same kind of time inconsistency as existing models of present bias, our theory deviates from these models in saying that the ex-*post* choice better reflects welfare: since a single day's work cannot affect the bonus by

much, in ex-post choice the agent does not focus on the bonus too much.

We conclude the paper in Section 5 with discussing some potential uses of and issues with our model. Although in this paper we develop only one application, a recipe for translating a deterministic classical model into one with focus-dependent choice—and hence for taking our theory to many other economic domains—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.[1] Indeed, we briefly mention some potential applications of our theory, including social preferences and product design and pricing, that span a range of economic settings. 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 potential effect of unchosen options on behavior, the recipe might require adjustment—or the development of a theory of consideration-set determination—when the choice set is not an appropriate approximation of the consideration set.

2 A Theory of Focusing

We formulate our basic model of focusing in Section 2.1, review related theories and evidence in Section 2.2, and discuss issues with our simple formulation in Section 2.3. In this paper, we model only riskless choice; we are currently developing an extension to risky choice that generalizes previous approaches, and we discuss modeling issues that arise in the context of such an 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* if she chooses option $c = (c_1, \ldots, c_K) \in C$—which can be thought of as corresponding to classical outcome-based utility— is $U(c) = \sum_{k=1}^{K} u_k(c_k)$, and we will often represent an option by its vector $(u_1(c_1), \ldots, u_K(c_K))$ of

---

[1] 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).

consumption utilities rather than by its vector of consumption levels. 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 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:[2]

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.

Beyond the above assumptions relating to behavior, a key assumption of our theory is that—although she maximizes focus-weighted utility—the agent's welfare is given by consumption utility. This assumption captures the view (consistent with the psychology literature below) 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. For instance, at the moment of deciding whether to live in California or the Midwest a person may focus too much on the differences in weather—and hence she may make the wrong choice—but once in California, she rarely cares or even thinks about the weather in the Midwest, so her experienced utility is little affected by the comparison.

---

[2] 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.

We follow two main principles when applying our framework. First, 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. If C included options that are extremely bad on all attributes—which are often available in economic decisions—these options could dominate the determination of relative focus. But in most situations a person quickly dismisses or does not even think about such "self-destructive" and other clearly inferior options, so that they do not influence her focus. While allowing some flexibility in the consideration set is necessary for our theory—as well as for related theories of context-dependent choice—to make correct predictions in general, we do not view C as a free parameter, and discipline should be used in its specification. Most importantly, because classical models as typically specified in practice do not include clearly inferior options in the choice set, in many settings it seems reasonable to simply equate the consideration set with the choice set of the corresponding classical model. This is the approach we take in Section 4 below. Our theory of choice from consideration sets can also be combined with any model of consideration-set determination.[3]

Second, for most applications in this paper we stick as close as possible to neoclassical economics in that—although we allow focus to distort decisions—we posit the same sources and timing of utility as would a classical model. But in some important applications, we do posit one non-classical source of utility: we assume that monetary transactions induce *direct* utility consequences, so that for instance an agent making a payment experiences an immediate utility loss. The idea that people treat monetary transactions as immediate utility is both intuitively compelling and widely supported in the literature on behavioral economics: it is central to Thaler's (1985, 1999) theory of mental accounting and Prelec and Loewenstein's (1998) "pain of paying," and some evidence on individuals' attitudes toward monetary gambles, such as narrow bracketing (Tversky and Kahneman 1981, Read, Loewenstein and Rabin 1999, Barberis, Huang and Thaler 2006), is difficult to explain without it. While in other applications we somewhat inconsistently assume that monetary transfers affect utility only indirectly through consumption, we show in Appendix A that

---

[3] 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.

all of our conclusions survive in a model in which they generate *both* direct and indirect utility consequences.

2.2 Related Theories and Evidence

To place our central assumption that larger differences attract more focus into context, we discuss a number of theories exploring versions of this hypothesis. While there are also differences in the precise formulation, the main new contribution of this paper is to identify some key general behavioral implications of this hypothesis, analyze how choices relate to welfare, and apply the theory to choice over time.

Tversky (1969) proposes a model of binary choice in which a decisionmaker does not notice small differences in an important attribute, but does notice and heavily weight larger differences. In a series of hypothetical-choice experiments, Tversky finds that subjects exhibit exactly the pattern of intransitivities predicted by his theory. Tversky also extends his model to a general theory of binary choice based on component-wise comparisons. In their model of choice between lotteries, Bordalo, Gennaioli and Shleifer (2012, forthcoming) assume that the salience of a state of the world—and with it, the decision weight assigned to that state—is increasing in the difference between lotteries' payoffs in that state. Their framework yields an explanation for the Allais paradox, frequent risk-seeking behavior, and preference reversals, and generates new predictions for risky choice that the authors confirm in experiments. Bordalo, Gennaioli and Shleifer (2011) apply the same framework to choices among products, and explore its implications for context effects. Loomes and Sugden's (1982) regret theory (although in interpretation unrelated to focusing) also yields the reduced-form assumption that states in which the difference between lotteries is greater carry a greater weight in an individual's choice—because the scope for regret or rejoicing is greater in those states. Finally, Gabaix (2011) posits 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.

As further motivation for our theory, we also mention some direct evidence for Assumption 1. 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.[4] 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. While these pieces of evidence support our central assumption, we view our model's (and the above related models') ability to organize evidence and intuitions from a variety of domains as making the most compelling case for it.

2.3   Comments on the Model

We conclude the discussion of our basic theory by highlighting several issues raised by our simple approach. First, although the evidence in Section 2.2 is not precise as to how focus depends on the number and dispersion of options, for simplicity and tractability Assumption 1 imposes a specific functional form for this relationship. The logic of our results in this 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 exploring the determinants of focus more precisely, which will help improve on our formulation.

Second, following classical economic models, in our formulation consumption utility depends only on the agent's consumption level. Because in some situations experienced utility depends on more aspects of the agent's environment—such as her reference point or the outcomes of others— for our theory to make correct predictions it is sometimes necessary to start from a richer notion of consumption utility. As an important example of how this can change implications, adding prospect theory's (Kahneman and Tversky 1979) diminishing sensitivity can weaken or reverse our reference-point-free model's prediction that individuals are more responsive to concentrated

---

[4] This study and the next pertain to predicted happiness rather than choice—but presumably individuals would make choices in these situations corresponding to their predicted happiness.

financial incentives than to dispersed financial incentives. At the same time, all predictions in this paper are fundamentally about consumption *utilities* rather than consumption levels, and when viewed in these terms they hold irrespective of the source of the consumption-utility vectors $(u_k(c_k))$ associated with options c.[5] Going further, we argue in Section 3 that the evidence for diminishing sensitivity is weakest in exactly those situations in which the bias toward concentration works against it, suggesting that our theory identifies a real force on behavior even for predictions regarding consumption levels.

Third, 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.

Fourth, 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 $\Delta_k(C) = \max_{c \in C} U(c_k, c_{-k}^0) - \min_{c \in C} U(c_k, c_{-k}^0)$, and define the focus-weighted utility of c with respect to c^0 as $wU(c) + (1-w) \sum_{k=1}^{K} g(\Delta_k(C)) \left[U(c_k, c_{-k}^0) - U(c^0) \right]$. 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

---

[5] For instance, even if consumption utility is reference-dependent, our theory says that a person is prone to a bias toward concentration in consumption utility (Section 3); and that, for a given reference point, a person exhibits both present bias and time inconsistency in exercise decisions as evaluated with respect to her consumption utility over the pain and health benefits of exercise (Section 4).

agent to overweight that attribute. Our basic model obtains as a special case when U is additively separable.[6]

3   Bias Toward Concentration

This section identifies general properties of our model that underlie many of its predictions in applications. We show that the decisionmaker is biased toward options whose advantages relative to alternatives are more concentrated than its disadvantages, and maximizes consumption utility when the advantages and disadvantages of alternatives are equally concentrated. We present these properties in a way that facilitates their application to intertemporal choice in Section 4 below.

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. Based on our discussion in Section 2 that a person may treat monetary transactions as immediate utility, we posit that the disutility of payments is immediate, and assume no discounting. Furthermore, as in Section 4 below, we assume that consumption utilities in different periods correspond to different attributes. Then, the consumer represents the decision of whether to finance as choosing between the streams $(-899, 0, \ldots, 0)$ and $(0, -39, \ldots, -39)$. Hence, although paying up front maximizes consumption utility, 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 the up-front payment.

Indeed, beyond the need to overcome liquidity constraints, our model's simple implication that "\$39 a month" seems less expensive than a price of \$899 may be one reason for the popularity of financing for purchases. Furthermore, because our prediction does not rely on liquidity constraints, it also explains why consumers sometimes resort to expensive financing even though they have

---

[6] Finally, 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."

liquid funds available (Bertaut, Haliassos and Reiter 2009, Stango and Zinman 2009).[7]

More generally than financing, the above logic says that consumers find paying less aversive if the price is split into multiple components they perceive as separate attributes. Consistent with this prediction, Gourville (1998) notes that retailers such as magazines or charities often frame aggregate expenses as a series of small "pennies-a-day" (per-issue or daily) payments, and documents that such a strategy increases demand. Similarly, the multitude of fees some firms impose for add-on services may (in addition to efficiency reasons) be in part motivated by an attempt to split prices. For instance, credit-card issuers, banks, and mobile-phone companies 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.[8]

We now identify general conditions under which the bias toward concentration occurs. For any $\epsilon > 0$ we say that the agent is ϵ-biased toward c in C if she does not choose an option $c' \in C \backslash \{c\}$ with $U(c') < U(c) + \epsilon$—that is, she does not choose an option whose consumption utility does not beat c's by at least an ϵ margin. We also say that the *advantages* of an option c relative to an alternative c' are the set of values $u_k(c) - u_k(c')$ for attributes k in which c is superior ($u_k(c) > u_k(c')$). We define disadvantages analogously. In the financing example, paying immediately has 24 dispersed advantages of 39 each, and financing has a single concentrated advantage of 899.

Proposition 1 (Bias Toward Concentration)**.** *Fix any $F, f > 0$. Then, there exists an $\epsilon > 0$ such that for any C and $c \in C$ satisfying the following conditions the agent is ϵ-biased toward c in C: (i) the advantages of c relative to any alternative in C are all greater than $F + f$; and (ii) any disadvantage any option in C has in the other attributes is lower than F.*

Proposition 1 says that if the advantages of c are uniformly greater than any disadvantage any option has in the other attributes, then the agent is biased toward c. Intuitively, the agent focuses too much on c's large advantages relative to any of c's possible disadvantages, biasing her choice

---

[7] 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.

[8] Agarwal et al. (2007, 2008) and Stango and Zinman (2009) document that consumers pay 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).

toward c. While somewhat special, this condition holds in several economic choices of interest. For instance, if a product—such as a mobile provider with a much faster network than all competitors—stands out in an attribute in a market with otherwise close competitors, consumers will be biased toward that product, and hence may buy it even if this means giving up too much in total on the other attributes.

By a similar logic, the agent is also biased toward c if it has a single advantage and multiple disadvantages relative to all alternatives:

Proposition 2 (Bias Toward Concentration 2). *Fix any $F, f > 0$. Then, there is an $\epsilon > 0$ such that for any C and $c \in C$ satisfying the following conditions the agent is ϵ-biased toward c in C: (i) there is an attribute k such that for any $c' \in C \setminus \{c\}$ we have $c_k > c'_k$ and $c_l \leq c'_l$ for all $l \neq k$; (ii) $u_l(c'_l) - u_l(c_l) \geq f$ for at least two attributes l; and (iii) $\Delta_l(C) \leq F$ for all attributes l.*

The condition of Proposition 2 holds in the financing example, where financing has a single advantage and multiple disadvantages relative to paying immediately. We will use this condition repeatedly in Section 4 for intertemporal decisions in which the agent can choose an option with a single one-time benefit and costs distributed over multiple periods. For example, overeating today generates immediate pleasure today and delayed costs for many periods into the future.

Balanced choices. We now identify some classes of situations in which the agent always makes consumption-utility-maximizing decisions. Unsurprisingly—given that the agent's behavior is distorted by focusing effects—such classes are quite narrow. Nevertheless, they allow us to provide revealed-preference foundations for our model, and act as a starting point for comparative statics on which types of situations are more conducive to good choices.

As a motivating example, 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 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

13

ordering.

To generalize the above example to other two-option choices, we say that the choices c^1, c^2 have *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). *If $C = \{c^1, c^2\}$ and c^1, c^2 have balanced tradeoffs, the decisionmaker makes a consumption-utility-maximizing choice.*

Proposition 3 is of immediate interest primarily in the special case $K' = 1$. This case applies, for instance, to a choice trading off utility at one earlier date with utility at one later date—such as when deciding whether to do a single indivisible task now or later—and the decision of whether to purchase a one-attribute product with a price the consumer experiences as a single attribute. In addition, because the agent maximizes consumption utility when trading off two attributes, her choices can be used to measure her relative preference for any two attributes. As we show in Appendix B, this property helps identify our model's ingredients from behavior: we can use rationality in balanced choices to elicit consumption utility, and then the bias in unbalanced choices to elicit $g(\cdot)$.

Beyond tradeoffs between two attributes, two-option balanced choices are relevant because some more complicated decisions ultimately boil down to such choices:

Corollary 1. *Suppose that $c^1, c^2 \in C$ generate a balanced tradeoff and span the range of possible outcomes in all attributes in C. If either c^1 or c^2 is the consumption-utility-maximizing choice in C, then the agent makes a consumption-utility-maximizing choice.*

Because c^1 and c^2 are balanced and one of them is utility-maximizing, the range of consumption utilities is greatest in exactly those attributes in which the utility-maximizing option beats all alternatives. Hence, again the agent's focus reinforces the advantages of the consumption-utility-maximizing choice. This result is most relevant when—such as in deciding whether to commit to an exercise regimen—a person makes multiple simultaneous binary decisions that are similar in

that they have the same (positive or negative) effect on consumption utility. Since the actions are similar, the utility-maximizing choice is either to take all of them or to take none of them. Corollary 1 says that if these two extremes constitute a balanced tradeoff, the agent makes the right decision.

While the situations guaranteeing consumption-utility-maximizing behavior are rare, we show that their logic also yields a natural comparative static: we establish that the agent makes better decisions in "more balanced" choices—choices in which the number of advantages and disadvantages is more similar. We consider only two-option decisions; our result can be extended to multiple similar decisions analogously to Corollary 1 above. Suppose $C = \{c^1, c^2\}$, with c^1 being better than c^2 by p utils on K_p attributes, and being worse than c^2 by m on K_m attributes. Let $P = K_p p$ and $M = K_m m$ be the total pluses and minuses of c^1 over c^2, respectively. Then:

Proposition 4 (More Rationality in More Balanced Choices). *Fix $P/M \neq 1$ and p. Then,*

[1] if $K_p'/K_m' \leq K_p/K_m \leq 1$ and the agent maximizes consumption utility for K_p', K_m', then she also maximizes consumption utility for K_p, K_m.

[2] if $K_p'/K_m' \geq K_p/K_m \geq 1$ and the agent maximizes consumption utility for K_p', K_m', then she also maximizes consumption utility for K_p, K_m.

Proposition 4 will be useful in our analysis of intertemporal choice for comparing ex-post and ex-ante decisions, as this change in perspective affects the relative concentrations of the costs and benefits of alternatives. As another example for how to use the proposition, suppose that a person is considering whether to go for a difficult run today ($K_m = 1$). Proposition 4 implies that, fixing P/M, the runner is more likely to make the consumption-utility-maximizing decision if she is training primarily for the Fall soccer season (K_p is small) than if—perhaps because of heart problems—she is training primarily to maintain good health in general (K_p is large).

Bias toward concentration and diminishing sensitivity. To conclude this section, we mention the relationship between our main prediction here and an important implication of prospect theory's (Kahneman and Tversky 1979) diminishing sensitivity. 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 inconclusiveness of the latter evidence.[11]

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, by predicting variation in present bias and time inconsistency across situations, our theory yields a unified explanation for a number of distinct and in part unexplained phenomena discussed in the literature, and also generates new predictions. Second, in contrast to existing work on present bias and time inconsistency, our model—and its guiding principle that decisions in more balanced choices are more likely to reflect a person's true well-being—can be used to provide unambiguous welfare conclusions without a priori assuming that ex-ante choice reflects

---

[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 consumer is more willing to make a related purchase than an unrelated purchase: the car-purchase decision tilts her focus toward the "car" attribute, which increases the perceived value of the car stereo as well.

welfare.[12]

In Section 4.1, we introduce our model of choice over time based on the focusing framework above. In the remaining subsections we develop our behavioral and welfare results through a stylized investment problem, and discuss how these results relate to previous models and evidence.

4.1 Focusing in Intertemporal Choices

We begin by extending our model of focusing 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 $\left\{\tilde{x}_\tau^t(h_{t-1}, x_t)\right\}_{\tau=t+1,\ldots,T}$, 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:

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

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 her future actions following each possible choice x_t are determined as the optimal behavior in the commitment problem starting 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.[13] The insights in this paper hold for both theories of how the agent forms beliefs about the future.

---

[13] 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.

4.2 Investing in the Future, One Step at a Time

We develop our central insights regarding choice over time by applying the framework above to a simple investment problem. 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 a period $t \in \{1, \ldots, T_i\}$ is $e_t \cdot B$, and any one investment generates a consumption-utility benefit of $A \left(\sum e_t \cdot B \right) / T_b$ in *each* of the last T_b periods. 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]

The key features of the above decision problem—that the effect of a period's investment is distributed over multiple periods, and that the effects of multiple investments accumulate—are present in many intertemporal decisions, including exercise, work, harmful consumption, and consumption-savings. Of course, to capture these basic forces in a technically convenient way, we have imposed a number of unrealistic assumptions. Most importantly, while the assumption that the investment and benefit periods are completely temporally separated is convenient in that it ensures an even distribution of costs as well as benefits, it is not satisfied in any of the relevant applications. In the case of exercise, for instance, a person can derive health benefits from past exercise while at the same time suffering disutility from current exercise. So long as the costs and benefits of effort are different attributes, such overlaps can (somewhat clumsily) be captured using our model by assuming that the decision problem is repeated multiple times, and the investment periods of a subsequent problem overlap with the benefit periods of the previous problem. In more general settings, the distribution of costs and benefits across periods may not be even, and—as in the case of consumption-savings—the costs and benefits in a period may accrue to the same attribute. This affects the focus weights and hence our precise predictions, but does not seem to eliminate the basic intuitions we identify. In particular, in Appendix C we consider a standard consumption-savings problem, and confirm our main insights below in that setting.[15]

---

[14] 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.

[15] Our assumption that choices are binary is also special. Once again, modifying this assumption can affect the

We begin by illustrating some of our formal results on ex-ante and ex-post choices in the above decision problem using two examples.

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, by the logic of Proposition 2's bias toward concentration, the consumer is too prone not to exercise.

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, $(-80, \ldots, -80, 100, \ldots, 100)$ and $(0, \ldots, 0)$. As a special case of Corollary 1, she commits to the consumption-utility-maximizing option of always exercising.

In Example 1, therefore, the consumer exhibits 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. In a logic reminiscent of Akerlof's (1991) and Rick and Loewenstein's (2008) arguments that the benefits of present-oriented behavior are often more "tangible" than the costs, 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, leading to more future-oriented choices. Our model also says that of these two perspectives, 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 (Planning a Wedding). In each of the first 100 periods, a happy parent can expend effort

---

focus weights and hence the precise predictions, but—as we show in the case of consumption-savings—does not seem to eliminate the basic intuitions we identify.

to improve her child's wedding, which occurs in period 101. A period's effort has a utility cost of 5 in that period and increases the utility from the wedding by 4. ($T_i = 100, T_b = 1, B = 5, A = 0.8$.)

Without commitment, the parent's decision in each period can be represented as a choice between $(-5, 4)$ and $(0, 0)$, where the 5 is the current effort cost of preparing and the 4 is the resulting improvement in the wedding. By Proposition 3, therefore, in each period the parent makes the consumption-utility-maximizing choice of not preparing for the wedding. From an ex-ante perspective, however, the parent's choice is effectively between $(-5, \ldots, -5, 400)$ and $(0, \ldots, 0)$, and these extreme options also determine her focus weights for the ex-ante choice. By Proposition 2, therefore, the parent is too prone to committing to a big wedding.

Like the consumer in Example 1, therefore, the parent exhibits time inconsistency in behavior in that her ex-post choice is more present-oriented than her ex-ante choice. In contrast to the exercise example, however, in this case our model says that the parent's ex-*post* choice maximizes her well-being, so that she is unbiased in ex-post choice and *future*-biased in ex-ante choice. Intuitively, because in her ex-ante choice the parent focuses on the large concentrated gain from having a great wedding and pays relatively less attention to the dispersed everyday costs, she tends to agree to overpreparing. But from the perspective of each period's work, she compares effort that day with just a marginally better wedding, making the effort less appealing.

We now turn to stating a general result about 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 is *future-biased* if $A^*_i < 1$. The agent exhibits *time inconsistency*—i.e. she is more present-oriented ex post than ex ante—if $A^*_{ante} < A^*_{post}$.

Proposition 5 (Biases in Investment). *Suppose $g(\cdot)$ is continuous. Then there is a strictly increasing continuous function $h(\cdot)$, which only depends on $g(\cdot)$ and B, such that $h(1) = 1$ and*

1. $A^*_{post} = h(T_b)$,

2. $A^*_{ante} = h(T_b/T_i)$.

In the rest of this section, we draw out some economic implications of Proposition 5. Notice first that the proposition implies a generalization of Example 1 to any setting in which repeated decisions have accumulating effects dispersed over multiple future periods: if $T_b \geq T_i > 1$, then (by Part 1) the agent is present-biased in ex-post choice ($A^*_{post} > 1$) and (by Part 2) she exhibits time inconsistency ($A^*_{post} > A^*_{ante}$). This means that our model has the same basic implications as hyperbolic discounting in exactly those settings in which the latter theory has been most invoked— harmful consumption, exercise, other lifestyle choices, and consumption-savings and borrowing decisions—providing a focus-based explanation for much of the present bias and time inconsistency that has been observed.[16]

Next, we turn to developing new comparative-statics and welfare predictions that are different from those of hyperbolic discounting, and that—we argue—are both intuitively compelling and supported by some evidence. Because our theory makes the same predictions as hyperbolic discounting in exactly those settings in which the latter theory has been most invoked and documented, but makes different and realistic predictions in other settings, it likely identifies a mechanism that is important in intertemporal choice. Nevertheless, because our model unrealistically predicts no present bias or time inconsistency for choices trading off the present with a single future period ($T_i = T_b = 1$)—such as deciding whether to do a single indivisible task now or later—it is not a complete theory of choice over time, and a complete theory would have to incorporate hyperbolic discounting or a related model as well.

4.3 Comparative statics

We start by stating a key result about how the environment shapes present bias and time inconsistency.

---

[16] See Laibson (1997a), 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 (1999, 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. Indeed, unlike for most phenomena in behavioral economics, the evidence for present bias and time inconsistency is strongest in these field settings—where our model makes the same prediction—and much weaker in laboratory decisions trading off current benefits with concentrated future costs—where our model does not make that prediction.

Corollary 2 (Comparative Statics). *Suppose $g(\cdot)$ is continuous. Then:*

1. A^*_{post} *is increasing in T_b and is independent of T_i.*

2. A^*_{post}/A^*_{ante} *is increasing in T_i.*

Part 1 of Corollary 2 says that an increase in T_b leads the agent to be more present-oriented in behavior and more present-biased relative to the utility-maximizing outcome. Intuitively, an increase in the number of periods T_b in which the consequences of current misbehavior are dispersed dilutes the benefits but not the cost of exerting effort today, leading the agent to focus relatively less on the benefits. This comparative static provides a novel, simple, testable prediction of our theory that may potentially explain a lot of variation in present bias across situations.

As a potentially important example, the above prediction helps resolve a body of puzzling evidence on individuals' high responsiveness to monetary incentives with regard to harmful consumption. One manifestation of this sensitivity is that individuals are often surprisingly responsive to the price of harmful substances; for example, Gruber and Kőszegi (2004) find a price elasticity of smoking of -0.6. Another manifestation is that individuals are quite responsive to experimental interventions that provide monetary incentives for better behavior. In Volpp et al.'s (2008) study, for instance, participants in a 16-week weight-loss program could earn on average \$21 per week if they kept losing 1 pound per week, and this intervention led participants to lose 13.1 pounds on average. Weight loss of this magnitude has immediate effects on outcomes such as blood pressure and serum lipid levels, and if maintained has large future health benefits as well.[17] Finkelstein, Linnan, Tate and Birken (2007) also find significant effects of moderate incentives on weight loss, and Lussier, Heil, Mongeon, Badger and Higgins (2006), Sindelar, Elbel and Petry (2007), and Sindelar (2008) discuss similar effects of financial incentives for managing substance use.

From the perspective of hyperbolic discounting and other existing models, the above kind of sensitivity is puzzling for a simple reason: the primary cost of consuming many harmful products

---

[17] The control group in the study lost 3.9 pounds on average during the program. At a follow-up, treated participants still retained an average weight loss of 9.2 pounds, but (with the authors' small samples) this was no longer statistically significantly greater than that of the control group. Whether and in what situations monetary incentives lead to sustained weight loss is an open question.

is in the form of long-term health consequences rather than in the form of short-term financial consequences, so the observed elasticity to the financial consequences implies empirically implausible and unobserved extreme sensitivity to the health consequences. As an illustration, consider what Gruber and Kőszegi's price elasticity of -0.6 implies for smokers' responsiveness to the future health consequences of smoking in a model of hyperbolic discounting, assuming a relatively low $\beta = 0.6$. Viscusi and Hersch (2007) estimate that the present value of just the mortality cost of smoking amounts to \$94-\$222 per pack. Taking a per-pack smoking cost of \$100, a five-percent reduction in the mortality cost should have about the same effect on smoking as a per-pack price change of $\$100 \cdot 0.6 \cdot 0.05$ or \$3. Assuming a price per pack of \$6, this means that the 5 percent reduction in mortality cost should have the same effect as a price drop of 50 percent! Extrapolating for illustration using the—local—price elasticity estimate of -0.6, a 5 percent reduction in mortality cost should increase the demand for smoking by 30 percent. This seems implausibly high: we are not aware of any evidence documenting large changes in smoking behavior following, for instance, diagnosis of potential future heart or lung problems, changes in health-insurance status, advances in cancer drugs, or other events that could impact the harm from cigarettes by a proportionally small amount. By similar calculations, the incentive effect of even a minor piece of news regarding the health benefits of losing weight should dwarf \$21 per pound—especially since in Volpp et al.'s study the monetary benefit from good behavior today accrues in the future—yet it seems unlikely that people would respond to such news by losing significant weight. Consistent with this perspective, Abaluck (2011) documents that consumers overrespond to prices relative to information about the nutritional content of foods.

In contrast to standard theories, our model can easily accommodate a combination of high responsiveness to monetary incentives and low responsiveness to health incentives. In the limit, when the health costs of current misbehavior are dispersed over many periods, the focus weight on these costs is close to $g(0)$. If $g(0)$ is sufficiently close to zero, therefore, our model says that the agent puts low weight on future health consequences, while, assuming that she evaluates monetary transactions in part as immediate utility, she may put much higher weight on current or future concentrated monetary incentives. Our theory also helps draw welfare and policy conclusions from

the evidence above. As Abaluck (2011) notes, it is unclear whether consumers' responsiveness to health information or their responsiveness to prices is relevant for welfare calculations. Our theory says that consumers make better decisions in more balanced choices, so that—since the benefit of consuming unhealthy food is concentrated—it is consumers' responsiveness to (concentrated) prices that is more appropriate for welfare and policy analysis.

Part 2 of Corollary 2 implies a sense in which 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 the agent integrates ex ante, it determines how much more she 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 is more likely to have a noticeable impact on her life. We are not aware of systematic tests of this prediction, nor of evidence that could be used to conclusively confirm or reject it. But—although the evidence is subject to multiple plausible interpretations—our prediction is consistent with observations that despite apparent desire to change one's behavior, the take-up of even effective short-term commitment devices in the field has been quite low. For instance, Giné, Karlan and Zinman (2010) offered, as a smoking cessation commitment device, an interest-free bank account in which participants could deposit funds that they forfeited if they did not pass a urine test in six months. While 43 percent of individuals offered the account claimed that they wanted to stop smoking in a year, and 53 percent said that they needed help to quit, only 11 percent took up the account. Similarly, although this comparison must be interpreted with extreme caution due to the different accounts and populations, only 28 percent took up the short-term commitment savings account SEED offered by Ashraf, Karlan and Yin (2006), but 78 percent took up the long-term SMarT account offered by Thaler and Benartzi (2004).[18]

---

[18] As in previous models, in the decision problem we have considered above time inconsistency can only be of one type: where the consumer is more present-oriented in ex-post choice than in ex-ante choice. In other settings, however, our model generates the opposite time inconsistency. The simple logic determining the direction of time inconsistency in our model is the following. A person is more present-oriented in ex-post choice than in ex-ante choice 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 is more *future*-oriented in ex-post choice than in ex-ante choice if 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. Our working

4.4 Welfare

Because our theory is based on a single utility function that reflects the agent's welfare, unlike existing models of time inconsistency it can be used to draw unambiguous welfare conclusions without a priori assuming that the ex-ante view reflects welfare. We emphasize here two potentially important new welfare implications of our theory. First, while Proposition 5 implies that for lifestyle choices a person's more future-oriented ex-ante choice is better than her ex-post choice ($A^*_{post} > A^*_{ante} \geq 1$ for $T_b \geq T_i > 1$), our theory does not endorse the view that ex-ante choices fully reflect welfare. An immediate caveat arises for $T_b > T_i$: in that case, the ex-ante choice is not balanced, so the agent is present-biased also in ex-ante choice, although less so than in ex-post choice. Furthermore, if the investments have non-constant marginal benefits, then even for $T_b = T_i$ the agent often does not maximize welfare ex ante:

Example 3 (Exercise with Decreasing Marginal Benefits)**.** 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. In addition, exercising in one of the periods 1 through 99 generates health benefits of α in each of the periods 101 through 200, while exercising in period 100 generates health benefits of $\alpha' \leq \alpha$ in the same periods.

Example 3 reduces to Example 1 if $\alpha = \alpha' = 1$. But suppose instead that $\alpha = 1$ and $\alpha' = 0.75$, so that here period 100's exercise session has future per-period benefits of only 0.75. In this example, the extreme choices determining the consumer's focus weights in ex-ante choice are $(0, \ldots, 0)$ and $(-80, \ldots, -80, 99.75, \ldots, 99.75)$. Hence, because the consumer focuses more on the benefits of exercise than on its costs, she may agree to exercise every day—even on the day in which this is not worth it. In fact, our theory predicts a novel comparative static: for any given α', the higher is α, the more likely the consumer is to commit to exercising in period 100. Intuitively, the more beneficial is exercise overall, the more the consumer's focus is distorted toward the benefits of exercise, and hence the more suboptimal the exercise sessions she is also willing to agree to.

Example 3 has an interesting reinterpretation in the context of work and career decisions. Thinking of the investments as the effort put into work and the benefits as the future consumption

paper formalizes these arguments (Kőszegi and Szeidl 2011, Section 4.2).

made possible by work, the example says that if work is beneficial overall, individuals may agree to work too much. Furthermore, the comparative static above says that individuals who have "better" jobs (for whom α is higher) are especially prone to overworking.

The second potentially important welfare prediction of our model arises even in binary choices, and follows from generalizing Example 2. If $1 \approx T_b \ll T_i$, the agent is *future*-biased in ex-ante choice, tending to overcommit to making investments over time for a temporally concentrated benefit. While decisions trading off dispersed costs with a temporally concentrated future consumption benefit seem less common and of more limited economic importance than decisions trading off a current benefit with dispersed future costs, it is worth noting that this prediction has further implications if a person places a sufficiently large weight on bonuses or other concentrated monetary rewards. In this case, the person will commit to work too much for such rewards, and by implication firms may strategically use them to motivate employees. Because this implication depends on the strength of utility from money relative to the utility from consumption, however, at this point it is only a tentative implication of our model.[19]

5   Other Potential Applications and 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. Combined with the natural assumption that other individuals' utilities correspond to different attributes, our theory yields a novel focus-based model of social preferences, with the forces identified in this paper immediately generating some interesting implications. Our theory predicts that individuals will be more selfish when—such as with tax evasion—the cost is hurting many others by a little bit than when—such as with stealing—the cost is hurting one person by a lot. Indeed, many social problems, including environmental destruction,

---

[19] Similarly, if people view "achievements" such as getting a promotion as a single attribute, they may—consistent with the view of several researchers, including Scitovsky (1976), Loewenstein, O'Donoghue and Rabin (2003), Kahneman, Krueger, Schkade, Schwarz and Stone (2006), and Hamermesh and Slemrod (2008)—commit themselves to overly ambitious careers relative to what would make them happiest. Because we have no way of determining whether and when people think of achievements in such terms, this suggestion is even more tentative than that for monetary rewards.

underprovision of public resources, as well as tax evasion, have (from an individual's perspective) this benefit-to-one-costs-to-many structure. Relatedly, our model predicts contradictory attitudes to such socially suboptimal behavior: while an individual may forgive or even approve tax evasion by a typical single other individual because she does not focus on the dispersed costs to other citizens, she may strongly support a policy of cracking down on tax evasion because of her focus on its society-wide costs.

Our theory also has potential use for understanding 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.[20] As the flip side of the incentive for price splitting we have discussed in Section 3, our theory's bias toward concentration implies that 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." These predictions raise several natural further questions, including which value attributes firms concentrate on in different situations, how firms design, bundle, and price products to take advantage of consumers' distorted focus when making purchase decisions, and how consumer focus affects the competitiveness of an industry. For example, as an analogue of Example 3, our theory implies that individuals may be too prone to buy add-ons that increase the value of a base product, and are the more so the higher is the base product's value relative to price.

We briefly mention several other likely worthwhile applications. Research indicates that retirees take too much of their retirement wealth in one lump sum rather than as an annuity,[21] and our model says that this may simply happen because a lump-sum payment looks very large relative to an annuity's monthly payments. Similarly, our theory predicts that an employee may be motivated by some features of her employment contract—such as a bonus, major promotion, or other large

---

[20] See Spiegler (2011) for a review of a small literature that analyzes traditionally marketing questions using economics methods.

[21] 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).

goal—not only because they can generate higher consumption utility, but also because of her disproportionate focus on these features. 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.

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. (2012, forthcoming), 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 in progress), we combine these possibilities.

References

Abaluck, Jason, "What Would We Eat if We Knew More: The Implications of a Large-Scale Change in Nutrition Labeling," 2011. Working Paper, MIT.

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.

Akerlof, George A., "Procrastination and Obedience," *American Economic Review*, 1991, *81* (2), 1–19.

Asheim, Geir B., "Procrastination, Partial Naivete, and Behavioral Welfare Analysis," 2008. Working Paper, University of Oslo.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *The Quarterly Journal of Economics*, 2006, *121* (2), 635–672.

Barberis, Nick, Ming Huang, and Richard H. Thaler, "Individual Preferences, Monetary Gambles and Stock Market Participation: A Case of Narrow Framing," *American Economic Review*, 2006, *96* (4), 1069–1090.

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, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," 2011. Working Paper, Harvard University.

___ , ___ , **and** ___ , "Salience Theory of Choice Under Risk," *Quarterly Journal of Economics*, 2012, forthcoming.

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.

Carroll, Christopher D., "Theoretical Foundations of Buffer Stock Savings," 2011. Working paper, Johns Hopkins University.

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.

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`.

Finkelstein, Eric A., Laura A. Linnan, Deborah F. Tate, and Ben E. Birken, "A Pilot Study Testing the Effect of Different Levels of Financial Incentives on Weight Loss Among Overweight Employees," *Journal of Occupational and Environmental Medicine*, 2007, *49* (9), 981–989.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," 2011. Working paper, NYU Stern. URL: http://pages.stern.nyu.edu/ xgabaix/papers/sparsebr.pdf.

Genesove, David and Chris Mayer, "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 2001, *116* (4), 1233–1260.

Giné, Xavier, Dean Karlan, and Jonathan Zinman, "Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation," *American Economic Journal: Applied Economics*, 2010, *2* (4), 213–35.

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).

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, "A Model of Focusing in Economic Choice," 2011. Working Paper, UC Berkeley, available at `http://elsa.berkeley.edu/~botond/focusingold.pdf`.

⸺ **and** ⸺ , "Focusing in Risky Choices," in progress.

Laibson, David I., "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

___ , "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.

Loomes, Graham and Robert Sugden, "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty," *Economic Journal*, 1982, *92* (368), 805–824.

Lussier, Jennifer Plebani, Sarah H. Heil, Joan A. Mongeon, Gary J. Badger, and Stephen T. Higgins, "A Meta-Analysis of Voucher-Based Reinforcement Therapy for Substance Use Disorders," *Addiction*, 2006, *101* (2), 192–203.

Meier, Stephan and Charles Sprenger, "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2010, *2* (1), 193–210.

Odean, Terrance, "Are Investors Reluctant to Realize Their Losses?," *Journal of Finance*, 1998, *53* (5), 1775–1798.

O'Donoghue, Ted and Matthew Rabin, "Addiction and Self-Control," in Jon Elster, ed., *Addiction: Entries and Exits*, Russel Sage Foundation, 1999.

___ **and** ___ , "Optimal Sin Taxes," *Journal of Public Economics*, 2006, *90* (10-11), 1825–1849.

Prelec, Drazen and George Loewenstein, "The Red and the Black: Mental Accounting of Savings and Debt," *MARKETING SCIENCE*, 1998, *17* (1), 4–28.

Rabin, Matthew and Georg Weizsäcker, "Narrow Bracketing and Dominated Choices," *American Economic Review*, 2009, *99* (4), 1508–1543.

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

Rick, Scott and George Loewenstein, "Intangibility in Intertemporal Choice," *Philosophical Transactions of the Royal Society of London*, 2008, *363* (1511), 3813–3824.

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.

Sindelar, Jody, Brian Elbel, and Nancy M. Petry, "What Do We Get for Our Money? Cost-Effectiveness of Adding Contingency Management," *Addiction*, 2007, *102* (2), 309–316.

Sindelar, Jody L., "Paying for Performance: The Power of Incentives over Habits," *Health Economics*, 2008, *17* (4), 449–451.

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.

___ **and Shlomo Benartzi**, "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, 2004, *112* (1), S164–S187.

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

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

Viscusi, W. Kip and Joni Hersch, "The Mortality Cost to Smokers," 2007. NBER Working Paper #13599.

Volpp, Kevin G., Leslie K. John, Andrea B. Troxel, Laurie Norton, Jennifer Fassbender, and George Loewenstein, "Financial Incentive-Based Approaches for Weight Loss," *Journal of the American Medical Association*, 2008, *300* (22), 2631–2637.

Yaari, Menahem E., "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies*, 1965, *32* (5), 137–150.

A   Utility from Monetary Transactions

In this paper, we model how the agent thinks about the consequences of spending or receiving money in two different ways. As discussed in Section 4 and formalized in greater generality in Appendix C below, when modeling consumption-savings decisions we directly apply our theory of focusing to the corresponding classical decision problem, which implies that the cost of consuming more today is in the form of reduced consumption in future periods. But in some applications, including the decision of whether to finance a purchase and the effect of monetary incentives on health behavior, we assume that the agent treats monetary receipts and payments as current (concentrated) utility, which implies that the cost of spending money is in the form of an immediate "pain of paying."

To show that our results are not driven by conflicting sets of assumptions, in this section we propose a natural way to combine these two ways of modeling the utility consequences of monetary transactions. This combined framework assumes that the agent derives utility *both* from monetary receipts and payments directly, and from the future consequences of spending. We illustrate that in this combined model all of our results survive, resolving the tension between the two approaches to money.

Formally, suppose that in each period t, monetary transfers m_t affect the agent's instantaneous consumption utility directly, and constitute one separate attribute in addition to those derived from neoclassical utility. For simplicity, we assume that the utility of both consumption and transfers is linear, with marginal utilities 1 and λ, respectively. To capture the hypothesis that paying leads to immediate disutility but this disutility is not so large as to make consumption undesirable, we assume that $0 < \lambda < 1$.[22] With this extension to a richer notion of consumption utility,

---

[22] At the cost of some conceptual (and notational) complexity, our framework can be extended to non-linear consumption utility. The main conceptual complication is to determine the appropriate assumption regarding the

we simply apply our model of dynamic decisionmaking from Section 4.1. In particular, in each period t, the agent chooses c_t, m_t from a consideration set that depends on previous choices. Given her beliefs about how her current choice affects future consumption and monetary transfers, the current consideration set induces a consideration set of lifetime consumption-payment paths, and the agent's focus weights derive from this set.

While our combined formulation raises some additional questions regarding the psychology of money that are beyond the scope of this paper, we illustrate briefly that it does accommodate in one framework all the predictions we have emphasized in this paper. It is worth noting that in each of the settings below, the introduction of monetary utility would not in itself affect predictions—it is only in combination with our model of focusing that it does.

Consumption-Savings. To illustrate how the combined framework affects predictions in a consumption-savings context, consider the stylized decision in which, analogously to our examples in Section 4.2, the agent is deciding whether to consume 80 now or 1 in each of one-hundred future periods. We assume that consumption must be paid for at the time it occurs. The agent then chooses between the following two consumption-transfer profiles:

decision	attribute	profile
80 now	consumption	$(80,0 , \ldots,0)$
	money	$(-80, 0, \ldots, 0)$
1's later	consumption	$(0,1,\ldots, 1)$
	money	$(0,-1,\ldots, -1)$

Hence, the agent chooses immediate consumption if

$$g(80)80 - g(\lambda \cdot 80)\lambda \cdot 80 - 100g(1) + 100g(\lambda)\lambda > 0,$$

or

$$\frac{g(80)}{g(1)} + \lambda\frac{g(\lambda)}{g(1)} \cdot \frac{5}{4} > \lambda\frac{g(\lambda \cdot 80)}{g(1)} + \frac{5}{4},$$

which—as in our analysis in Section 4.2—holds if $g(80)/g(1)$ is sufficiently large. More generally, unless $g(\cdot)$ is too concave, the agent is always present-biased in an ex-post consumption-savings decision.

---

shape of the monetary-utility function, which determines the marginal utility of transfers. Because monetary utility likely stands in for the future usefulness of money, it seems natural to assume that the monetary-utility function is some $\lambda < 1$ times the indirect utility of money that derives from future consumption.

Intuitively, monetary utility introduces a "pain of paying" for consumption that counteracts the agent's bias toward current consumption we have identified in Section 4.2: because consuming now requires a concentrated payment now rather than dispersed smaller payments in the future, in terms of monetary utility the agent prefers future consumption. But since the pain of paying is a weaker force than the preference for consumption, this preference for paying later tends to be outweighed by the agent's present bias in consumption.

Financing or Paying Immediately. Beyond what consumption profile a person chooses, our combined model also has implications for how she prefers to pay for a given purchase. Recall the example of a laptop purchase in Section 3, in which a consumer decides whether to make one lump-sum payment now or 24 payments in the future. Slightly modifying our example to hold the total payment constant, suppose that the choice is between an immediate payment of \$936 and 24 monthly payments of \$39. We assume that the method of payment does not affect future consumption. With these assumptions, the model reduces trivially to our formulation in the text: the agent is indifferent between the two payment methods in terms of consumption, but prefers financing due to monetary utility.

Of course, if financing comes at the cost of higher total payments—as in our example of a one-time payment of \$899 versus 24 payments of \$39—choosing financing has the additional effect of lowering consumption in future periods. But if the consumer cares sufficiently little about this effect—which is often the case as the effect is dispersed—she still prefers financing.

Monetary Incentives. Finally, we illustrate that in this combined framework the agent cares more about monetary incentives than about incentives deriving from dispersed future health consequences, so that she is more responsive to the former incentives than to the latter incentives. Consider a person who chooses whether to smoke in the current period, where smoking yields a one-time pleasure she values at $B > 0$. For simplicity, we assume that cigarettes have a a price of zero to start with, and that the person initially believes that smoking has no future health cost.[23] With these parameters, the consumer clearly decides to smoke. To illustrate the contrast between monetary incentives and incentives deriving from health consequences, we ask whether the con-

---

[23] While it would affect the precise predictions, assuming instead that the price and perceived health cost of cigarettes are positive to start with does not affect the basic intuitions we identify below.

sumer will still smoke (a) if she must pay \$100 for it; and (b) if she finds out that smoking has health costs she values at \$1 in each of 100 periods. For simplicity, we suppose that if she pays the \$100, the consumer must reduce consumption by \$1 in each of the same 100 periods (considering a full consumption-savings problem will not affect the qualitative insights). There are then four pertinent attributes, the pleasure of smoking, health outcomes, consumption, and money. Dropping the (for this decision irrelevant) health attribute, the consumer's choice in case (a) is:

smoke?	attribute	profile
	smoking	(0,0 , …,0)
no	consumption	(0,0 , …,0)
	money	(0,0 , …,0)
	smoking	(B,0 , …,0)
yes	consumption	(0,-1 , …,-1)
	money	(-100,1 , …,1)

Analogously, dropping the (for this decision irrelevant) consumption and money attributes, the consumer's choice in case (b) is:

smoke?	attribute	profile
no	smoking	(0,0 , …,0)
	health	(0,0 , …,0)
yes	smoking	(B,0 , …,0)
	health	(0,-1 , …, -1)

It is easy to check that for any $\lambda > 0$ the monetary incentive is stronger: whenever the agent abstains in choice (b), she strictly prefers to abstain in choice (a). Intuitively, the loss in future consumption from paying in choice (a) and the health consequences in choice (b) have the same effect on the consumer's focus-weighted utility; but in choice (a) the concentrated monetary payment smoking requires makes it less attractive.

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 that (i) we know how products map into attributes and there are at least three attributes; (ii) we can manipulate individual attributes in the decisionmaker's choices; (iii) $g(\cdot)$ is strictly increasing; and (iv) the utility functions $u_k(\cdot)$ 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

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

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.[24] 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.

For an example, suppose that a consumer's mobile-phone provider offers both a data plan and a mobile TV plan. These two plans are two potential attributes, and we would like to identify whether the consumer treats them as distinct attributes or as parts of the same "service quality" attribute. Assuming that there is a third attribute that we know is distinct from these two, the bias towards concentration allows us to elicit how the consumer treats the two attributes. Intuitively, if the two attributes are parts of the same attribute, then offering them jointly rather than separately will make them more attractive to the consumer.

Suppose the third attribute that we know is distinct is monetary utility on some date, and consider the following elicitation procedure. (i) We find the amount of money, m, such that the consumer is indifferent whether to pay m for the data plan alone. (ii) Given m, we find the *additional* payment, m', such that the consumer is indifferent whether to pay m' to get the mobile TV plan *in addition to* the data plan. (iii) We ask the consumer whether she is willing to pay $m + m'$ for both plans. If the two plans are the same attribute, the consumer will be indifferent, while if the plans are separate attributes, then she will strictly prefer not to buy them at this price.

To see the logic formally, begin with the case when the two plans are different attributes. Let the two plans be attributes 1 and 2 and the payments be attribute 3, and denote the consumption

---

[24] We include a third attribute and set $x_0 > 0$ because the first option would otherwise be clearly dominated by the other alternatives, so that it might not make the agent's consideration set. If the first option is sufficiently superior in the third attribute, however, it will presumably be in the consideration set, even if it is eventually not chosen.

utilities of the two plans by u_1 and u_2. In addition, we normalize money when no payment is made to zero. Since steps (i) and (ii) above generate balanced choices, we must have $u_1 = u_3(0) - u_3(-m)$ and $u_2 = u_3(-m) - u_3(-m - m')$. Hence, the choice in step (iii) is between $(u_3(0) - u_3(-m), u_3(-m) - u_3(-m - m'), u_3(-m - m') - u_3(0))$ and $(0, 0, 0)$, so that by Proposition 1's bias toward concentration, the consumer strictly prefers the second option. Now consider the case when the two plans are on the same attribute, with the plans being on attribute 1 and money being on attribute 2. Steps (i) and (ii) still yield $u_1 = u_3(0) - u_3(-m)$ and $u_2 = u_3(-m) - u_3(-m - m')$. The consumer's choice in step 3 is then between $(u_1 + u_2, u_3(-m - m') - u_3(0)) = (u_3(0) - u_3(-m - m'), u_3(-m - m') - u_3(0))$ and $(0, 0)$, and hence by Proposition 3 the consumer 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 consumer treats mobile TV as a feature of mobile-phone services that is separate from other attributes, we cannot tell whether in her mind there are one or two such mobile TV 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 mobile TV attribute.

C   Consumption-Savings

We now apply our intertemporal model to a classical consumption-savings problem, and show how the intuitions about present bias and time inconsistency identified in the investment model of Section 4 extend to this setting. Consider a consumer who lives for a finite number $T > 1$ of periods, has consumption utility

$$\sum_{t=1}^{T} \delta^{t-1} u\left(C_t\right),$$

enters each period t with wealth W_t, and faces the intertemporal budget constraints

$$W_{t+1} = \left(W_t - C_t\right)\left(1 + R\right)$$

for $t = 1, ..., T$ and $W_{T+1} \geq 0$, where $R \geq 0$ is the constant interest rate. We assume that consumption utility exhibits constant relative risk aversion, and to guarantee that focus weights are well-defined even when $C = 0$ is in the consideration set, we impose that the coefficient of relative risk aversion is less than one: $u(C) = C^{\alpha}$, with $0 < \alpha \leq 1$.[25]

To ensure that the solution of the neoclassical consumption-saving model remains well-behaved when the horizon T becomes large, we assume that $\delta(1+R)^{\alpha} \leq 1$. This inequality—the weak version of which is labeled by Carroll (2011) the "return impatience condition"—means that over time, the highest achievable per-period utility cannot grow faster than the discount rate shrinks. When the inequality fails, the infinite-horizon neoclassical model does not admit a solution because there exist policies which generate arbitrarily large lifetime utility; moreover, in the finite-horizon case, the consumer is so patient that as T grows without bound, she spends an arbitrarily small amount on current consumption.[26] By allowing for equality in the condition we accommodate the useful benchmark $\delta = 1$ and $R = 0$, in which consumption in each period is $C_t = W_1/T$.

For tractability, we work with a power function for the focus weight as well: $g(\Delta) = \Delta^{\theta}$. Then, standard arguments imply that both in the ex-post and ex-ante decisions of the focusing model, as well as in the rational ($\theta = 0$) model, consumption in any period is a constant fraction of available wealth. We let $b_t^i = C_t^i/W_t^i$ denote the average and also marginal propensity to consume in period t, where $i \in \{ante, post, rat\}$ refers to one of the three model variants. It is easy to see that b_t^i depends only on $T - t$. We say that the consumer is present-biased in ex-i choice if $b_t^i > b_t^{rat}$, and that she exhibits time inconsistency in period t if $b_t^{post} > b_t^{ante}$.

Proposition 6. *Assume that $\delta(1+R)^{\alpha} \leq 1$ and $\theta > 0$. Then*

[1] For any $T > 1$, the consumer is present-biased in ex-ante choice if $\delta(1+R)^{\alpha} < 1$, and she

---

[25] If the coefficient of relative risk aversion is greater than one, then $u(C)$ goes to minus infinity as C approaches zero, so that if arbitrarily low consumption levels are in the consideration set, the focus weights are not defined. With such a utility function, however, it seems reasonable to assume that the decisionmaker would not consider very low (and hence extremely unattractive) consumption levels when making her decision. If we impose any constant lower bound on consumption (i.e. $c_t \geq c_{min}$ for some $c_{min} > 0$), then our model is again well-defined; and, although we cannot solve the model in closed form, the forces we identify below continue to be active in the resulting consumption-saving problem.

[26] The return impatience condition does permit parameters such that $\delta(1+R) > 1$, and hence does allow for a consumption path that increases over time. It only rules out parameters for which the consumption-to-wealth ratio grows over time.

is unbiased in ex-ante choice if $\delta (1 + R)^\alpha = 1$.

[2] The consumer is present-biased in ex-post choice for any $t < T - 1$.

[3] The consumer exhibits time inconsistency in periods $t < T - 1$.

Part 1 of Proposition 6 follows from the observation that in ex-ante choice, the focus weight on a future period is a function of the maximum possible discounted consumption utility in that period. That maximum utility is determined by the combination of discounting and the utility gains from investing, explaining the role of $\delta (1 + R)^\alpha$ in the result. Intuitively, in ex-ante choice the consumer is more sensitive to changes in R than in the standard model, because a higher interest rate makes saving more attractive both for neoclassical reasons and—through increasing potential future utility—by increasing the focus weight on the future. The intuition for Parts 2 and 3 of Proposition 6 parallels the analysis of Section 4. Because a narrower range of future outcomes is considered, the consumer focuses relatively less on the future in ex-post choice, which generates both time inconsistency and present bias. For example, in the benchmark case when $\delta = 1$ and $R = 0$, the consumer decides rationally ex ante, and is present-biased ex post.

To prove these results, we first derive an Euler equation for ex-post choice. Let $h (C) = g (u (C))$ measure the focus weight as a function of consumption when consuming zero is in the consideration set. Then, dropping the *post* superscript for convenience, the following Euler equation characterizes ex post choice:

$$u' (C_t) = u' (C_{t+1}) \delta (1 + R) \cdot [h (b_{t+1}) b_{t+1} + h (1 - b_{t+1}) (1 - b_{t+1})] \cdot [\delta (1 + R)^\alpha]^\theta . \qquad (3)$$

As usual, the left hand side is the marginal cost, and the right hand side is the marginal benefit of saving an additional dollar. The term $u' (C_{t+1}) \delta (1 + R)$ is what a standard model would imply. However, as in Laibson (1997b), time-inconsistent preferences imply that we cannot directly substitute out future marginal utility using C_{t+1} only. The next term on the right-hand side collects the effect of a marginal increase in savings on focus-weighted utility in both $t + 1$ and in the future. In that term, b_{t+1} is the share of the saved dollar consumed in $t + 1$, and $h (b_{t+1})$ is the focus-weight associated with that consumption. In turn, $1 - b_{t+1}$ is the share of the saved dollar preserved for periods $t + 2,...,T$, and $h (1 - b_{t+1})$ can be interpreted as the focus weight associated with these

remaining periods. Finally, the term involving $\delta\left(1+R\right)^{\alpha}$ reflects the intuition discussed above that the focus weight on the future is affected both by discounting and by the potential utility gains from investing. The key to this first-order condition is that—because they are determined by the rationally anticipated future marginal propensities to consume—the relative focus weights across utilities over dates $t+2,...,T$ are the same for both period t's and period $t+1$'s self.

For a formal proof of the Euler equation, let g_t^s denote relative focus weight on a future period s versus the present period t. In period t, $u'\left(C_t\right)$ must equal the marginal benefit of saving an extra dollar:

$$g_t^{t+1}\delta\left(1+R\right)\frac{\partial C_{t+1}}{\partial W_{t+1}}u'\left(C_{t+1}\right)+g_t^{t+2}\delta^2\left(1+R\right)^2\frac{\partial C_{t+2}}{\partial W_{t+2}}u'\left(C_{t+2}\right)+...+g_t^T\delta^{T-t}\left(1+R\right)^{T-t}\frac{\partial C_T}{\partial W_T}u'\left(C_T\right).$$

Similarly, at $t+1$, $u'\left(C_{t+1}\right)$ must equal the marginal benefit of saving an extra dollar:

$$g_{t+1}^{t+2}\delta\left(1+R\right)\frac{\partial C_{t+2}}{\partial W_{t+2}}u'\left(C_{t+2}\right)+g_{t+1}^{t+3}\delta^2\left(1+R\right)^2\frac{\partial C_{t+3}}{\partial W_{t+3}}u'\left(C_{t+3}\right)+...+g_{t+1}^T\delta^{T-t-1}\left(1+R\right)^{T-t-1}\frac{\partial C_T}{\partial W_T}u'\left(C_T\right).$$

Now note that

$$g_t^\tau = g\left(\delta^{\tau-\tau}u\left(\left(1+R\right)^{\tau-t}\cdot\left(1-b_{t+1}\right)\left(1-b_{t+2}\right)\cdot...\cdot\left(1-b_{\tau-1}\right)\cdot b_\tau\right)\right)$$

and hence for all $\tau\geq t+1$, $g_t^\tau/g_{t+1}^\tau = g\left(\delta u\left(\left(1+R\right)\left(1-b_{t+1}\right)\right)\right)$, where the terms in g come from discounting, return accumulation, and the fact the agent at date t subtracts consumption in period $t+1$. Because this ratio is constant for all $\tau\geq t+1$, combining the above equations for future marginal utilities yields the Euler equation.

Substituting the functional forms of u and h into the Euler equation implies, after manipulations, that

$$\frac{1}{b_t^{post}} = 1+\frac{1}{b_{t+1}^{post}}\left(\delta^{1+\theta}\left(1+R\right)^{\alpha+\alpha\theta}\right)^{1/(1-\alpha)}\left[\left(b_{t+1}^{post}\right)^{1+\alpha\theta}+\left(1-b_{t+1}^{post}\right)^{1+\alpha\theta}\right]^{1/(1-\alpha)}. \qquad (4)$$

Given that $b_T^{post} = 1$, we can use this equation to recursively solve for $b_{T-1}^{post},...,b_1^{post}$. We have thus characterized ex-post behavior in this model.

Now consider ex-ante choice. Because the highest consumption level in a period t is $W_1\left(1+R\right)^{t-1}$, the focus weight over utility in that period, using our functional forms, is $W_1^{\alpha\theta}\left[\delta\left(1+R\right)^\alpha\right]^{\theta(t-1)}$.

Because these weights are exponential in t, the ex-ante decision is observationally equivalent to the outcome of a neoclassical consumption-savings problem with a different discount factor $\gamma = \delta^{1+\theta} \cdot (1+R)^{\alpha\theta}$, where the first δ comes from neoclassical discounting and the remaining terms come from the focus weights. Using this observation we can compute the optimal consumption path with (4) substituting $\theta = 0$ and the discount factor γ to obtain

$$\frac{1}{b_t^{ante}} = 1 + \frac{1}{b_{t+1}^{ante}} \left(\delta^{1+\theta} (1+R)^{\alpha+\alpha\theta} \right)^{1/(1-\alpha)}. \tag{5}$$

Finally, the choices of a rational agent can similarly be characterized using $\theta = 0$ in (5) with the recursion

$$\frac{1}{b_t^{rat}} = 1 + \frac{1}{b_{t+1}^{rat}} \left(\delta (1+R)^{\alpha} \right)^{1/(1-\alpha)}. \tag{6}$$

Comparing the recursions for the ex-ante and the rational choice, given that $b_T^{ante} = b_T^{rat} = 1$, it is easy to show inductively that, for all $t < T$, when $\delta (1+R)^{\alpha} < 1$ we have $b_t^{ante} > b_t^{rat}$, and when $\delta (1+R)^{\alpha} = 1$ we have $b_t^{ante} = b_t^{rat}$. And comparing the recursions for the ex-ante and ex-post choice shows that, because $b_{t+1}^{1+\alpha\theta} + (1-b_{t+1})^{1+\alpha\theta} < 1$ when $b_{t+1} < 1$, for all $t < T - 1$ we have $b_t^{post} < b_t^{ante}$.

D   Proofs

Proof of Proposition 1. For simplicity, assume that the vectors $c \in C$ are already measured in utility terms. Let $\epsilon = [g(F+f) - g(F)](F+f)/g(F) > 0$. Fix some alternative $c' \in C$ for which $U(c') < U(c) + \epsilon$ and let A denote the attributes k in which $c_k > c'_k$. The focus-weighted utility difference between c and c' is

$$
\begin{aligned}
\tilde{U}(c,C) - \tilde{U}(c',C) &= \sum_{k \in A} g(\Delta_k(C)) \cdot (c_k - c'_k) + \sum_{l \notin A} g(\Delta_l(C)) \cdot (c_l - c'_l) \\
&> g(F+f) \sum_{k \in A} (c_k - c'_k) + g(F) \sum_{l \notin A} \cdot (c_l - c'_l) \\
&> [g(F+f) - g(F)](F+f) + g(F)(U(c) - U(c')) \\
&> [g(F+f) - g(F)](F+f) - g(F)\epsilon \\
&= 0.
\end{aligned}
$$

44

where we use that $\Delta_k(C) > F + f$ and $\Delta_l(C) < F$ for all $k \in A$ and $l \notin A$; that $c_k - c'_k > 0$ if and only if $k \in A$; and that A is non-empty. Thus the agent does not choose c' over c. $\quad\square$

Proof of Proposition 2. We begin by defining ϵ. Because $g(\cdot)$ is continuous, it is also uniformly continuous on a compact interval. Thus for any $\beta > 0$ there exists $\gamma > 0$ such that $|g(x) - g(y)| < \beta$ when $|x - y| < \gamma$ for $x, y \in [0, F]$. Pick $\beta > 0$ and the corresponding $\gamma > 0$ small enough such that $\beta F + \gamma g(F) < g(f)f/2$ and also $\gamma < f/2$. Let $\epsilon_1 = (-\beta F - \gamma g(F) + g(f)f)/g(F)$ and $\epsilon_2 = f - \gamma$, both of which are positive. Denote the minimum over $x \in [\gamma, F]$ of $[g(x) - g(x - \gamma)]2f/g(x)$, a continuous function, by ϵ_3, which is also positive. We define $\epsilon = \min[\epsilon_1, \epsilon_2, \epsilon_3]/2$.

As in the previous proof, we assume that the vectors in C are measured in utility terms, and for notational simplicity denote $\Delta_i(C)$ by Δ_i. Additionally, for each attribute we normalize to 0 the lowest consumption utility in C in that attribute; and relabel attributes to ensure $k = 1$. With these normalizations, the assumption of the proposition implies that $c_1 = \Delta_1$ and $c_2 = ... = c_K = 0$.

Suppose, toward a contradiction, that $c^* \neq c$ is chosen from C even though $U(c^*) < U(c) + \epsilon$.

Case I. First consider the case when the maximum difference in attributes $2, ..., K$, denoted Δ_j, exceeds $\Delta_1 - \gamma$. Denote by c' the option in C for which $c'_j = \Delta_j$. The consumption utility of c' is at least $\Delta_1 - \gamma + f$ which is higher than $U(c) + \epsilon = \Delta_1 + \epsilon$ because $\epsilon < \epsilon_2 = f - \gamma$. Thus c' is different from c^*. We now show that the agent prefers c' over c^*, a contradiction.

We consider two subcases. First, suppose that $\Delta_1 \geq \Delta_j$. The focus-weighted utility of c' is at least

$$\tilde{U}(c', C) \geq g(\Delta_1 - \gamma)(\Delta_1 - \gamma) + g(f)f \geq (g(\Delta_1) - \beta)(\Delta_1 - \gamma) + g(f)f$$
$$> g(\Delta_1)\Delta_1 - \beta F - \gamma g(F) + g(f)f = g(\Delta_1)\Delta_1 + g(F)\epsilon_1.$$

In contrast, $\tilde{U}(c^*, C)$ is at most $g(\Delta_1)(\Delta_1 + \epsilon)$, which is smaller than the above expression since $\epsilon < \epsilon_1$ and $\Delta_1 \leq F$.

Now suppose $\Delta_1 < \Delta_j$. Then

$$\tilde{U}(c', C) - \tilde{U}(c^*, C) \geq g(\Delta_j)\Delta_j + g(f)f - g(\Delta_j)(\Delta_1 + \epsilon) > g(f)f - g(F)\epsilon > 0$$

because $\epsilon < \epsilon_1 < g(f)f/g(F)$.

Case II. Next consider the case in which $\Delta_j \leq \Delta_1 - \gamma$. Note, this implies in particular that $\Delta_1 \geq \gamma$. In this case

$$\tilde{U}(c^*, C) \leq g(\Delta_1)(\Delta_1 - 2f + \epsilon) + g(\Delta_1 - \gamma)2f$$

because c^* has a total advantage of at least $2f$ over c in attributes $2, ..., K$, and has total utility of at most $\Delta_1 + \epsilon$. But then

$$\tilde{U}(c, C) - \tilde{U}(c^*, C) \geq g(\Delta_1)(2f - \epsilon) - g(\Delta_1 - \gamma)2f > g(\Delta_1)(2f - \epsilon_3) - g(\Delta_1 - \gamma)2f$$

$$> g(\Delta_1)\left(2f - 2f\frac{g(\Delta_1) - g(\Delta_1 - \gamma)}{g(\Delta_1)}\right) - g(\Delta_1 - \gamma)2f = 0.$$

Thus c is chosen over c^*, a contradiction. □

Proof of Proposition 3. Immediate. □

Proof of Corollary 1. Continue to assume that the vectors in C are measured in utility terms. Suppose c^1 is the utility-maximizing option, and assume $c_i^2 = 0$ for all i. Denote by A the K' attributes in which c^1 is superior to c^2 and by a the corresponding utility difference. Analogously, denote by B the K' attributes in which c^1 is inferior to c^2, and by b the utility difference. Since the utility of c^1 exceeds that of c^2, we have $a > b$.

For any alternative $c' \in C$, the fact that c^1 is the utility-maximizing option implies

$$\sum_{i \in A}(a - c_i') > \sum_{i \in B}(-b - c_i')$$

where the spanning assumption implies that all terms in both sums are non-negative. Multiplying the left hand side by $g(a)/g(b) > 1$ implies

$$\sum_{i \in A}g(a)(a - c_i') > \sum_{i \in B}g(b)(-b - c_i')$$

which means that the focus-weighted utility of c^1 exceeds that of c'. □

Proof of Proposition 4. The focus-weighted utility of c^1 exceeds that of c^2 if and only if $Pg(p) > Mg(m)$ which is equivalent to

$$\frac{P}{M}g(p) > g\left(\frac{K_p}{K_m}\frac{M}{P}p\right).$$

By assumption both p and P/M are held fixed, and hence changes in K_p and K_m affect the inequality only through the K_p/K_m term on the right hand side.

To prove [1], first assume that $P < M$. Then the right choice is c^2, and the agent will make the right choice if the above inequality fails, which is more likely when K_p/K_m is higher. Next assume that $P > M$. Then the right choice is c^1. Since $K_p \leq K_m$, $P > M$ implies $p > m$, and hence $Pg(p) > Mg(m)$ always holds. Thus the agent always makes the right choice. Part [2] can be shown analogously. $\qquad\square$

Proof of Proposition 5.

1. In the ex-post choice of period t, the focus weight on attribute t is set by $\Delta_t = B$, while the focus weight on attributes $s = T_i + 1, \ldots, T_i + T_b$ is determined by $\Delta_s = AB/T_b$. As a result, the consumer chooses 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 $g(B) = A^* \cdot g(A^*B/T_b)$. Let the function $h(\cdot)$ be implicitly defined by the equation $g(B) = h(z) \cdot g(Bh(z)/z)$. Since $g(\cdot)$ is strictly increasing, $h(\cdot)$ is well-defined, strictly increasing, depends only on $g(\cdot)$ and B, and satisfies $h(1) = 1$. Moreover, by the above formula, $A^*_{post} = h(T_b)$.

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 definition of $h(\cdot)$ implies that $A^*_{ante} = h(T_b/T_i)$. $\quad\square$

Proof of Corollary 2. Immediate. $\qquad\square$