A MODEL OF FOCUSING IN ECONOMIC CHOICE*

Botond Kőszegi
Adam Szeidl

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 decision maker 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. Applying our model to intertemporal choice, these results predict that a person exhibits present bias and time inconsistency when—such as in lifestyle choices and other widely invoked applications of hyperbolic discounting—the future effect of a current decision is distributed over many dates, and the effects of multiple decisions accumulate. But unlike in previous models, in our theory (1) present bias is lower when the costs of current misbehavior are less dispersed, helping explain why people respond more to monetary incentives than to health concerns in harmful consumption; and (2) time inconsistency is lower when a person commits to fewer decisions with accumulating effects in her ex ante choice. In addition, a person does not fully maximize welfare even when making decisions ex ante: (3) she commits to too much of an activity—for example, 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.

I. 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). 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

*Michele Muller-Itten and Takeshi Murooka provided excellent research assistance. We thank David Ahn, Stefano DellaVigna, Erik Eyster, Michael Grubb, Emir Kamenica, Joan de Martí Beltran, Andrea Prat, Matthew Rabin, Richard Thaler, and seminar participants for insightful discussions. Kőszegi thanks the Berkeley Center for Equitable Growth, the European Research Council (Starting Grant #313341), and the National Science Foundation (Award #0648659) and Szeidl thanks the Alfred P. Sloan Foundation and the National Science Foundation (Award #0752835) for financial support.

The Quarterly Journal of Economics (2013), 1–52. doi:10.1093/qje/qjs049.

theoretical implications of such phenomena for economic questions involving multiattribute choice are largely unexplored and unclear.

Building on evidence and related work we discuss in Section II, especially approaches by Tversky (1969), Loomes and Sugden (1982), and Bordalo, Gennaioli, and Shleifer (2012a, 2012b), in this article 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 helps explain facts that are puzzling from the perspective of existing theories and generates new economic insights. We also discuss other applications, including social preferences, and argue that our model provides a simple way to incorporate focusing into many more economic settings.

Section II presents our framework. 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* and welfare from a choice $c = (c_1, \ldots, c_K)$ is $U(c) = \sum_{k=1}^{K} u_k(c_k)$. But instead of consumption utility, the decision maker 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. Although this formulation allows a modeler to determine focus endogenously once C, the attributes, and $u_k(\cdot)$ are specified, we discuss some serious modeling issues—such as the possibilities that C is different from the agent's entire choice set, that attributes are affected by framing, and that the utility function is reference-dependent—one must confront when taking our framework to an economic setting.

In Section III, we identify the main general implications of our model. Because the decision maker focuses too much on a large advantage relative to multiple small disadvantages, she exhibits a "bias toward concentration": She is too prone to choose options with advantages concentrated in fewer attributes. Conversely, because the larger focus on a single advantage relative to a single smaller disadvantage just reinforces the advantage, the decision maker exhibits no bias in "balanced choices": She maximizes consumption utility when the advantages and disadvantages of relevant options are equally concentrated. As we illustrate in examples of social decisions and use repeatedly in our analysis of intertemporal choice, these results help identify ways in which specific economic choices are biased, and provide guidance as to which choices better reflect welfare.

In Section IV, we develop the main application in this article, 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. We first show that, unlike hyperbolic discounting, our theory does not predict present bias and time inconsistency in single balanced choices, such as whether to have 30 minutes of massage today or 60 minutes of massage next week. Nevertheless, our theory *does* predict present bias and time inconsistency in settings in which a sacrifice today results in small per period benefits for many periods in the future, and the benefits of repeated sacrifices accumulate. The bias toward concentration implies that when deciding whether to exercise on any given day, a person focuses too little on the many small future health benefits relative to the one big current cost, and hence—exhibiting present bias—she tends to exercise too little.[1] But when considering many possible sessions of exercise ex ante, the person focuses more on the large fitness gains she will enjoy every day if she exercises regularly, so—exhibiting time inconsistency—she prefers to exercise more. Because most of the field settings in which evidence consistent with hyperbolic discounting has been documented—such as exercise, harmful consumption, and consumption-savings—feature repeat “lifestyles choices” with distributed and accumulating benefits, our model provides an alternative explanation for this field evidence.

Furthermore, our theory makes novel comparative-statics predictions on the strength of present bias and time inconsistency. Most important, the bias toward concentration implies that an increase in the number of periods in which the future costs of current misbehavior are dispersed increases present bias. This prediction helps explain some empirically documented variation in harmful consumption that reasonable parameterizations of hyperbolic (and exponential) discounting models cannot

1. The logic for this result is 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.

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 even the hyperbolically discounted health consequences of consumption far outweigh the financial consequences, hyperbolic discounting implies that consumers should then be extremely responsive to proportionally small 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 consistent with our model if we make the reasonable assumption that monetary payments carry some immediate utility, so that the financial incentives are more concentrated than the health consequences.

A second new comparative static concerns variation in time inconsistency in repeated decisions with accumulating effects. The more decisions the consumer can commit 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 (e.g., quitting smoking forever) rather than only a trivial part of her future (e.g., quitting smoking for a week). Although there are other plausible explanations, we argue that this prediction is broadly consistent with existing evidence on the take-up of commitment devices in the field.

Beyond these new comparative statics, our model—based on a single well-defined utility function that reflects welfare—also helps sort out the welfare implications of time inconsistency. Supporting a commonly held presumption, for the foregoing types of lifestyle and consumption-savings choices our theory says that—it being a more balanced choice—ex ante behavior usually better reflects welfare than ex post behavior. But because focus is determined by the total rather than by the marginal benefit of effort, our model also implies that a person may "overcommit" to worthwhile tasks with decreasing marginal benefit. For 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 preparing for a marathon or working for a large bonus that the person evaluates as a separate attribute—a sequence of sacrifices leads to a single large and hence attention-grabbing goal, our theory predicts "future bias" in ex ante choices, and says 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 V with a recipe for translating a deterministic classical model into one with focus-dependent choice, with a brief discussion of other potential applications, and with some conceptual issues on how to extend the model to environments with uncertainty in a general way.

II. A Theory of Focusing

We formulate our basic model and review related research in Section II.A, and discuss issues with our simple formulation in Section II.B. Although we specify a way to apply the framework to some risky choices, our formal model and main emphasis are on riskless choices.

II.A. 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 decision maker'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 consumption utilities rather than by its vector of consumption levels. But instead of consumption utility, the decision maker maximizes *focus-weighted utility*

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

where g_k is the focus weight on attribute k. Our model's central assumption concerns how the g_k are determined:

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.

In addition, we make a key assumption regarding welfare:

ASSUMPTION 2. Welfare is given by consumption utility $U(c)$.

We now discuss our main assumptions in turn. Assumption 1 says that the decision maker focuses more on attributes in which her options generate a greater range of consumption utility. As we will argue, this assumption both seems empirically relevant and—because it derives focus from only the basic model ingredients introduced so far—provides a simple way to model focusing without appealing too much to determinants that are either difficult to observe or not relevant in most economic situations.

A number of theories explore versions of the hypothesis that larger differences attract more focus. To explain certain types of intransitivities, Tversky (1969) proposes a model of binary choice in which a decision maker does not notice small differences in an important attribute, but does notice and heavily weight larger differences. Bordalo, Gennaioli, and Shleifer (2012a, 2012b) assume that the salience of—and with it, the decision weight assigned to—a lottery's payoff in a state is increasing in the difference between this payoff and other available lotteries' payoff in the same state. Their framework, which also assumes diminishing sensitivity in salience, yields an explanation for risk-seeking behavior, the Allais paradox, and preference reversals. Bordalo, Gennaioli, and Shleifer (2012a) apply the same framework to choices among products, and explore its implications for context effects.[2] Modeling binary choices between two-outcome lotteries, Rubinstein (1988) assumes that the agent treats "similar" probabilities or outcomes as identical and uses this model to explain the Allais paradox. 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

2. In Bordalo, Gennaioli, and Shleifer's models, the weight on an option's payoff in an attribute or state is option-specific—it depends on how far the option is from others in the same attribute or state. In our model, in contrast, the weight is not option-specific—it depends only on the range of consumption utilities in the attribute. As we note shortly, the existing evidence does not allow us to pin down a particular specification, and we have chosen ours mainly because it simplifies the calculation of focus weights. For the questions we consider in this article, the two specifications seem to lead to similar intuitions.

individual's choice—because the scope for regret or rejoicing is greater in those states.[3] Going beyond this body of work, our article identifies some key general implications of the hypothesis that greater differences attract more focus, analyzes how choices relate to welfare, and applies the theory to choice over time.

Although we are not aware of direct evidence pertaining to the details—such as the functional-form specification—of Assumption 1, a number of facts are consistent with the basic idea that larger differences attract more focus. 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 between the two locations (e.g., climate) and underweight important other determinants of life satisfaction. While these pieces of evidence support our central assumption, we view our model's as well as related models' ability to organize evidence and intuitions from a variety of domains as making the most compelling case for it.[5]

3. In less closely related work, 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; Schwartzstein (2012) assumes that an agent learning about the environment ignores factors that she expects not to be predictive. Finally, Cunningham (2011) develops a related model in which an agent who has encountered—either in her current decision or in a past decision—an option with a higher value in a given attribute becomes less sensitive to that attribute, and uses this model to understand context effects. Whether and how his findings can be reconciled with ours is an agenda for future research.

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.

5. Unless the focus-weight function $g(\cdot)$ is linear, our model is sensitive to affine transformations of the consumption-utility function $U(\cdot)$. This means that to keep the implications of our model unchanged when $U(\cdot)$ is rescaled, $g(\cdot)$ must be rescaled as well. Note that the elicitation procedure in Appendix B selects one of the equivalent scalings of $g(\cdot)$ and $U(\cdot)$.

Consistent with the psychology literature, we take the view that although focusing distorts how a person perceives her preferences near the moment of choice, it does not 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 might 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 not affected by the comparison. Hence, as captured in Assumption 2, we posit that the focus weights do not affect welfare.

II.B. Discussion

Our setup incorporates some implicit assumptions that play an important role in determining the effect of focus on choice. We now explicitly state these embedded assumptions and highlight some additional issues that are important for applying the model.

REMARK 1. Attributes are exogenous.

When applying our theory to a given economic setting, the definition of the relevant attributes must come from outside the theory. One way to minimize the danger that attributes are engineered to get specific results is to commit to restrictions on the set of attributes for a given domain of economic behavior. Restrictions that seem plausible to us include (1) that utilities realized on different dates are different attributes in intertemporal choices (as in Section IV); (2) that different individuals' utilities are different attributes in social choices (as in some examples in Section III); or (3) that quality and price are different attributes in models of industrial organization (as in Bordalo, Gennaioli, and Shleifer 2012a). While these restrictions add discipline, they sometimes fail to capture intuitively appealing ways that individuals might think of real-life decisions. For instance, an assistant professor may think of tenure in part as a single achievement rather than as a long sequence of benefits from it, so that specifying attributes as per period utilities does not fully capture her choice process. In addition, the approach of taking the attributes as fixed ignores the possibility that they may be influenced by framing, an issue which we discuss in Section III.

REMARK 2. The set C is exogenous and possibly different from the choice set.

Because not all comparisons necessarily affect a person's focus, we allow C to be different from the agent's entire choice set, thinking of it as a "consideration set" of her reasonable options. For example, it seems likely that in most situations a person quickly dismisses or does not even think about options that are extremely bad on all attributes, and hence they do not affect her focus—even though they would dominate the determination of focus if they were included in C. But while our general framework leaves C flexible, in applications we restrict ourselves and equate C with the choice set of a corresponding classical model—which as typically specified in practice does not include clearly inferior options.[6]

REMARK 3. The utility functions $u_k(c_k)$ depend only on the consumption levels c_k.

Following classical economic models, in our basic formulation consumption utility depends only on the agent's consumption level. If the agent's experienced utility depends strongly on other factors—such as her reference point or beliefs about the future—it may be necessary to start from a richer notion of consumption utility. With the modified consumption-utility function, one can apply our model in the same way as above, defining g_k as a function of the differences in (the now richer) consumption utility. As an important example, in Section III.D we discuss how prospect theory's (Kahneman and Tversky 1979) diminishing sensitivity modifies our predictions.[7]

6. As with the choice of attributes, our disciplined approach leaves out some potentially important issues. In particular, the set determining focus can (similarly to Bordalo, Gennaioli, and Shleifer's [2012a] "evoked sets") include options outside the consideration set and can be influenced by framing. For example, when a consumer is deciding whether to buy an item on sale, her focus may be influenced by the item's regular price, especially if the retailer conspicuously points this price out to her.

7. Our formulation also assumes additive separability in utility across attributes. One way to extend our framework to a setting where consumption utility is given by $U(c_1, \ldots, c_K)$ is the following. We posit that there is a "yardstick" option $c^0 \in C$, let $\Delta_k(C) = \max_{c\in C} U(c_k, c^0_{-k}) - \min_{c\in C} U(c_k, c^0_{-k})$, and define the focus-weighted utility of c with respect to c^0 as $wU(c) + (1-w)\sum_{k=1}^{K} g(\Delta_k(C))\left[U(c_k, c^0_{-k}) - U(c^0)\right]$. The first term allows complementarities in consumption utility to influence the agent's behavior, and the second term captures our hypothesis that greater differences on an attribute

REMARK 4. In some applications, we allow for immediate utility from monetary transactions ("pain of paying").

Although for most applications in this article, we posit the same sources and timing of utility as would a classical model, in some applications we also 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 experience monetary transactions as immediate utility is both intuitively compelling and supported in the literature: 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 money, such as narrow bracketing (Read, Loewenstein, and Rabin 1999, for example) and laboratory evidence on hyperbolic discounting (Thaler 1981, for example), 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 all of our conclusions survive in a model in which they generate *both* direct and indirect utility consequences.

REMARK 5. In some applications involving risky outcomes, we apply the model by defining expected utility as the relevant outcome determining focus.

While all of our formal examples and results are about deterministic choices, in the real-life settings we consider outcomes are almost always stochastic. To map stochastic outcomes to our deterministic setting, we simply think of expected utility in an attribute as the relevant outcome determining focus. Because in this specification large payoff differences in a low-probability state of the world do not affect focus very much, our model fails to capture many patterns in how individuals think about risk, including the patterns explained by the model of Bordalo, Gennaioli, and Shleifer (2012a, 2012b). We return to the issue of how to reconcile the expected-utilities-as-attributes and the state-by-state-payoffs-as-attributes models of salience in stochastic choice in Section V.

lead the agent to overweight that attribute. Our basic model obtains as a special case if U is additively separable.

III. Focus-Induced Patterns of Choice

In this section, we identify and illustrate some key general properties of our model that we use in our analysis of intertemporal choice in Section IV and that are likely to be important in other applications as well. We also discuss ways in which framing can affect attributes and hence the implications of our results, and comment on the relationship between the implications of our model and those of prospect theory's diminishing sensitivity.

III.A. Bias toward Concentration

For a simple example illustrating our model's most important property, suppose that a consumer purchasing a laptop computer is choosing whether to make one immediate payment of \$899 or 24 future monthly payments of \$39. Based on Remark 4, we posit that the consumption (dis)utility of payments is experienced immediately, and as in Section IV, we assume that consumption utilities realized in different periods correspond to different attributes. We also assume linear disutility from monetary payments and no discounting. 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 up-front 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. Because this prediction does not rely on liquidity constraints, it also explains why consumers sometimes resort to expensive financing even though they have liquid funds available (Bertaut, Haliassos, and Reiter 2009; Stango and Zinman 2009).

We now identify general conditions under which the bias toward concentration occurs. We define the *advantages* of an option c relative to an alternative c' as the set of values $u_k(c_k) - u_k(c'_k)$ for attributes k in which c is superior ($u_k(c_k) > u_k(c'_k)$). 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 with $F > f > 0$, and suppose that for some $c \in C$ (i) the advantages of c relative to any alternative in C are all greater than F; and (ii) any disadvantage any option in C has in the other attributes is lower than f. Then, the agent does not choose any $c' \in C \setminus \{c\}$ for which $U(c') - U(c) < F\left[\frac{g(F)}{g(f)} - 1\right]$.

Because $F > f$, conditions (i) and (ii) mean that the advantages of c are uniformly greater than any disadvantage any option has in the other attributes. Proposition 1 says that if this holds, the agent is biased toward c: A necessary condition for her to choose an alternative c' is that the consumption utility of c' exceeds that of c by the strictly positive margin $F\left[\frac{g(F)}{g(f)} - 1\right]$. Intuitively, the agent is biased toward c because she focuses more on its large advantages than on any of its possible small disadvantages; and the lower bound on the distortion in her focus, $\frac{g(F)}{g(f)}$, gives a lower bound on this bias. In the financing example, (i) and (ii) are satisfied with $F = 899$ and $f = 39$, confirming our earlier analysis that for $\frac{g(899)}{g(39)}$ sufficiently large, the agent chooses financing.

The conditions of Proposition 1 hold in several other economic choices of interest. We briefly consider here how focusing affects choices over social allocations, assuming that other individuals' utilities enter the agent's utility function and correspond to different attributes. Suppose that a taxpayer facing no risk of auditing is contemplating whether to evade \$80 in taxes, which has costs—in the form of lower-quality government services—of \$0.10 each on 1,000 other people. The taxpayer weights the utilities of others by $\gamma \leq 1$ relative to her own well-being. Assuming a constant marginal utility of money of 1 for all individuals, evading maximizes the consumption utility of the taxpayer if and only if $\gamma \leq 0.8$. In our model, the taxpayer's decision can be represented as $(0, \ldots, 0)$ if she does not evade versus $(80, -0.1\gamma, \ldots, -0.1\gamma)$ if she does evade, where the first term measures her material utility and the other 1,000 terms measure the utility she derives from others' well-being. Applying Proposition 1 with $F = 80$ and $f = 0.1$, she evades if $\gamma < 0.8\left[\frac{g(80)}{g(0.1)}\right]$—that is, even for some $\gamma > 0.8$. Intuitively, the taxpayer is too likely to evade because she does not focus on the dispersed costs she inflicts on society. More broadly, Proposition 1 predicts that given their own social preferences, people are too likely to engage in selfish activities—such as environmental destruction—that have dispersed costs and are too likely to object to policies—such as privatizing inefficient utility companies—that result in small gains to many at the expense of larger costs to a few people. We also use Proposition 1 repeatedly in Section IV for intertemporal decisions in which—like in the financing example—the agent can choose an option with a large one-time benefit in a given period and small costs distributed over multiple periods.

To further develop intuition for the bias toward concentration, our next result identifies a key comparative static: that increasing the concentration of an option's advantages makes it more likely to be chosen. We say that an option c' combines the advantages of c^1 relative to c^2 in attributes k and l if c' can be obtained from c^1 by merging c^1's advantages relative to c^2 in k and l into a single advantage in k: $u_k(c_k^1) > u_k(c_k^2)$, $u_l(c_l^1) > u_l(c_l^2)$, $u_k(c'_k) - u_k(c_k^2) = \left[u_k(c_k^1) - u_k(c_k^2)\right] + \left[u_l(c_l^1) - u_l(c_l^2)\right]$, $u_l(c'_l) = u_l(c_l^2)$, and $u_j(c'_j) = u_j(c_j^1)$ for all $j \neq k, l$. We also say that two options $c^1, c^2 \in C$ span the range of consumption utilities in C in attribute k if for all $c \in C$, we have $\min\{u_k(c_k^1), u_k(c_k^2)\} \leq u_k(c_k) \leq \max\{u_k(c_k^1), u_k(c_k^2)\}$. Then:

Proposition 2 (Increasing Concentration). Suppose that the options c^1 and c^2 span the range of consumption utilities in C in attributes k and l, and c' combines the advantages of c^1 relative to c^2 in k and l. If the agent is willing to choose c^1 from C, then she strictly prefers to choose c' from $\left[C \setminus \{c^1\}\right] \cup \{c'\}$.

In the special case in which there are only two alternatives ($C = \{c^1, c^2\}$), Proposition 2 simply says that concentrating two—or, by induction, several—advantages of an option makes it more preferable. Conversely, concentrating some disadvantages of an option makes it less preferable. Hence, Proposition 2 predicts that a person is more likely to evade taxes in the foregoing example (i.e., to choose $(80, -0.1\gamma, \ldots, -0.1\gamma)$ over $(0, \ldots, 0)$) than to steal \$80 from a single individual who is thereby hurt by \$100 (i.e., to choose $(80, -100\gamma, 0, \ldots, 0)$ over $(0, 0, 0, \ldots, 0)$).

III.B. Balanced Choices

We now identify some classes of situations in which the agent always makes consumption-utility-maximizing decisions. As a motivating example, suppose that the laptop buyer 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 ordering.

To generalize the example, we say that the choices c^1, c^2 have *balanced trade-offs* if for some $K', p, m > 0$, the number of k such that $u_k(c_k^1) - u_k(c_k^2) = p$ is K', the number of k such that $u_k(c_k^2) - u_k(c_k^1) = m$ is K', and the number of k such that $u_k(c_k^1) = u_k(c_k^2)$ is $K - 2K'$. That is, relative to c^2, c^1 is better by p on K' attributes (its "pluses") and worse by m on K' attributes (its "minuses"). Then:

PROPOSITION 3 (RATIONALITY IN BALANCED TRADE-OFFS). Suppose that $c^1, c^2 \in C$ generate a balanced trade-off and span the range of possible outcomes in all attributes in C. If either c^1 or c^2 is a consumption-utility-maximizing choice in C, then the agent makes a consumption-utility-maximizing choice.

To understand Proposition 3, we consider first balanced *binary* decisions ($C = \{c^1, c^2\}$). Because c^1 or c^2 must then be a consumption-utility-maximizing option, Proposition 3 implies that the agent always makes a consumption-utility-maximizing decision. Letting $K' = 1$, the agent therefore maximizes consumption utility in the stealing example ($(80, -100\gamma)$ versus $(0, 0)$), in choices trading off utility at one earlier date with utility at one later date and also in decisions of whether to purchase a one-attribute product with a price she experiences as a single attribute.

Furthermore, as we show in Appendix B, the fact that the agent maximizes consumption utility when trading off two attributes helps identify our model's ingredients from behavior: We can use this property to elicit consumption utility, and then the bias in unbalanced choices to elicit $g(\cdot)$. As a result, although our assumptions were not derived formally from patterns in choice behavior—as would be done in axiomatic theories—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.[8]

For an example illustrating Proposition 3 in a decision with more than two options, we modify the tax-evasion example. Suppose that a politician with $\gamma = 1$ is choosing between having no tax-evasion policy (c^1), a policy that prevents 1,000 people from evading (c^2), and a policy that prevents 100 people from evading (c^3), where (as before) each instance of evasion benefits the evader by \$80 and hurts 1,000 others by \$0.10 each. This decision can be represented as choosing between the following per capita outcomes:

Choice	First 100 taxpayers	Next 900 taxpayers	1000 beneficiaries
c^1	80	80	0
c^2	0	0	100
c^3	0	80	10

It is easy to check that c^1 and c^2 span the range of outcomes, and—because of constant returns to a tougher policy—one of them must be utility-maximizing. Since c^1 and c^2 also constitute a balanced trade-off, Proposition 3 says that the politician makes the consumption-utility-maximizing choice of c^2.

While balanced decisions are likely to be rare, their logic also yields a natural comparative static: that the agent makes better decisions in "more balanced" choices in which the number of advantages and disadvantages is more similar. We formalize this idea for two-option decisions; our result can be extended to multiple similar decisions analogously to Proposition 3. Suppose that $C = \{c^1, c^2\}$, 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.

8. We also illustrate in Appendix B that if we know the set of *possible* attributes in a situation, under plausible conditions we can identify which of these the decision maker treats as identical versus separate 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 $\frac{P}{M} \neq 1$ and p. For any $R, R' > 0$ such that either $R' < R \leq 1$ or $R' > R \geq 1$ holds, if the agent maximizes consumption utility for $\frac{K_p}{K_m} = R'$, then she also maximizes consumption utility for $\frac{K_p}{K_m} = R$.

To illustrate the result, we return to the tax-evasion example, assume that $\gamma = 1$, and contrast a person's attitude toward a single instance of tax evasion and a broader pattern of tax evasion by 100 people in society, where again the personal benefit of evasion is \$80 and the cost is \$0.10 each on a 1,000 others. If the person were to choose whether a single other individual should evade taxes, she would make the 1,001-attribute choice between $(0, \ldots, 0)$ and $(80, -0.1, \ldots, -0.1)$; whereas if she were to choose whether 100 others should evade, she would make the 1,100-attribute choice between $(0, \ldots, 0)$ and $(80, \ldots, 80, -10, \ldots, -10)$. Then, Proposition 4 says that because $p = 80$ and $\frac{P}{M} = 0.8$ in both decisions but $\frac{K'_p}{K'_m} = 0.001$ for the first while $\frac{K_p}{K_m} = 0.1$ for the second, the person is more likely to make the rational choice and dislike tax evasion in the second decision. Intuitively, a person 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, but also dislike broad tax evasion because of her focus on its society-wide costs.

III.C. Framing

Because our results on the bias toward concentration and balanced choices are driven by how payoffs are distributed across attributes, their implications can in practice depend on how the attributes are framed, even though such framing does not affect the economic outcomes associated with any alternative. In the financing example, for instance, the seller may explain that \$39 a month amounts to only about \$1.28 a day, or (conversely) a friend may advise the consumer to consider the total payment of \$936. It seems plausible that the first type of framing can make the consumer think about the daily costs as attributes, so that she finds the financing option even more attractive than in our original example, whereas the second type of framing can make her think about the total cost as an attribute, so that she finds the financing option less attractive than in our original example.[9]

Because our framework is based on exogenously given attributes, it leaves out such framing effects, and we hope that future research will help establish how framing works and what its limits are. Nevertheless, we believe that despite the possibility of framing, the qualitative predictions we develop in this article continue to be important—simply because per period utilities seem to be very natural attributes. For instance, we expect that framing monthly payments as a one-time payment usually does not make them feel as expensive as an actual one-time payment. Moreover, many types of framing—such as presenting $39 a month as 0.09 cents a minute—may just not work.

A more formal language for thinking about framing in our model is to assume that there is a set of potential attribute representations that a given utility vector can be decomposed into, and framing determines which one the agent invokes. In the financing example, the potential representations could be the ones with time intervals of at least one day. This framework accommodates both the seller's frame of $1.28 a day and the friend's frame of $936 in total, allows for some frames to be rejected by the agent (if, such as the 0.09-cents-per-minute frame, it is not among her potential representations), and also yields the natural prediction that a monopolist seller would frame prices by splitting them according to the most favorable potential representation. This language also raises important questions that are beyond the scope of this article, including what determines the set of potential representations and how a person thinks about attributes when competitors (e.g., the seller and the friend) are presenting different frames.[10]

9. Consistent with the notion that framing can influence attributes, 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 into separate attributes. 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.

10. Although their specific model of behavior is different, Mullainathan, Schwartzstein, and Shleifer (2008) also think of framing as influencing the way an agent represents a situation in her mind. Spiegler (2012) proposes a general

III.D. Relationship with Diminishing Sensitivity

To conclude this section, we comment on the relationship between the implications of our model and those of prospect theory's (Kahneman and Tversky 1979) diminishing sensitivity, the notion that sensitivity to changes in an outcome is lower if the outcome is further from the reference point. Because our theory predicts that people overweight a large advantage whereas diminishing sensitivity implies that larger deviations carry proportionally less weight, the two are seemingly in contradiction. We show that (depending on the environment) the two forces can in fact be orthogonal to, counteract, or strengthen each other, and we argue that focusing yields new insights on existing evidence and thought experiments about diminishing sensitivity.

Following Remark 3, we can add diminishing sensitivity to our model by incorporating it into the consumption-utility function, and letting focus be defined by differences in this modified utility function. As an example of a situation in which focusing and diminishing sensitivity are orthogonal, we consider Tversky and Kahneman's (1981) famous calculator example, in which a person is more likely to drive 10 minutes to save \$5 on a \$15 calculator than to drive 10 minutes to save \$5 on a \$125 jacket. Denoting the utility cost of driving 10 minutes by e and the utility of money by $v(\cdot)$, the former decision can be represented as choosing between $(0, v(-15))$ and $(-e, v(-10))$, while the latter decision can be represented as choosing between $(0, v(-125))$ and $(-e, v(-120))$. Because both of these are balanced choices, Proposition 3 implies that the person makes the consumption-utility-maximizing choice for any $v(\cdot)$. In particular, if $v(\cdot)$ is linear, our model predicts the same choice in the two circumstances, failing to capture Tversky and Kahneman's evidence. But if $v(\cdot)$ satisfies diminishing sensitivity relative to a reference point of zero—so that $v(-120) - v(-125) < v(-10) - v(-15)$—our model does predict that a person is more likely to drive in the calculator case, for exactly the same reason as in prospect theory. By a similar logic, our model also does not alter diminishing sensitivity's prediction that people are

model of the foregoing type of "competitive framing," and explores some natural hypotheses regarding how multiple frames interact. In one possible specification, he assumes that if firms choose the same frame, consumers adopt that frame; whereas if firms choose different frames, consumers become enlightened and look at the decision objectively.

more willing to pay an additional $10 for a better wine if—like in a restaurant—the base price is already higher (Bordalo, Gennaioli, and Shleifer 2012a).

As an illustration of how focusing can counteract diminishing sensitivity, we consider an example that is often used to illustrate diminishing sensitivity in the literature. Researchers have argued that due to diminished sensitivity to additional expenditures following a big payment, consumers are more willing to buy an add-on such as a car radio in the context of buying the car itself than at other times (Thaler 1985; Chiu and Wu 2009, for example). This explanation is incomplete: Applying diminishing sensitivity consistently to the "car" attribute as well, the consumer's sensitivity to the car radio should also be diminished when she is buying a car.[11] To see how this complicates the analysis, suppose the car costs $20,000 and its money-equivalent consumption value is $40,000, while both the price and the value of the radio is $500. Denoting the utility function in both the money and the car attribute by $v(\cdot)$, the consumer's choice is between not buying, $(v(0), v(0))$; buying the car only, $(v(40,000), v(-20,000))$; and buying the car as well as the radio, $(v(40,500), v(-20,500))$, where we assume for simplicity that $v(\cdot)$ is symmetric around 0. Then, because diminishing sensitivity implies that $v(40,500) - v(40,000) < v(-20,000) - v(-20,500)$, without focusing it predicts that the consumer does *not* purchase the car radio. Intuitively, because the value of the car exceeds its price—presumably, that is why the consumer is buying a car in the first place—the effect of diminishing sensitivity is greater on the car attribute than on the money attribute. Our theory, however, predicts a counterveiling force and thereby helps complete the logic of the example: because the focus weight on the car attribute is $g(v(40,500) - v(0))$ whereas on money it is only $g(v(0) - v(-20,500))$, the consumer attaches a greater weight to the car radio than to the money paid for it. As a result, she may well buy the radio. In addition, as an immediate implication of Proposition 2, the consumer is more likely to buy the car radio than to make an unrelated purchase of equal value and price.

11. One could also—in a somewhat ad hoc manner—assume that the consumer treats the radio as a separate attribute, but still integrates the money paid for it with that paid for the car. In this case, the explanation faces another problem: diminishing sensitivity predicts that the consumer should be equally willing to make any other purchase when she buys the car.

Intuitively, the car purchase decision tilts her focus toward the car attribute, which increases the perceived value of the car radio as well.[12]

Finally, we point out a simple example in which the bias toward concentration and diminishing sensitivity reinforce each other. Suppose that the agent has a reference point of zero and is choosing between $(v(0), v(-900), v(-900))$ and $(v(100), v(-950), v(-950))$. For example, a consumer might be deciding whether to add further to recently acquired debt for immediate consumption. Then, both because diminishing sensitivity makes the benefit of consumption feel larger than the per period monetary cost, and because the benefit is more concentrated, the consumer prefers to consume and add to her debt.

IV. Intertemporal Choice

This section explores some implications of our model for choice over time. In Section IV.A, we introduce our general model of choice over time based on the focusing framework, and in Section IV.B we use it to develop a stylized investment problem. We present our main behavioral and welfare results in Sections IV.C and IV.D, and discuss how these results relate to previous models and evidence. Because our model predicts patterns of present bias and time inconsistency, the main alternative theory to which we compare it is hyperbolic discounting (Laibson 1997; O'Donoghue and Rabin 1999b). In Section IV.E, we summarize the predictions of the two models and what our theory adds to the literature.

12. The bias toward concentration also counteracts diminishing sensitivity's prediction (pointed out for instance by Thaler and Johnson 1990) that people should have a preference for segregating gains and integrating losses relative to the reference point. While there is compelling evidence for diminishing sensitivity based on people's attitudes toward outcomes closer to versus further from the reference point, there is only limited evidence for a preference for separating gains and integrating losses. Thaler and Johnson (1990) and Linville and Fischer (1991) find that subjects do prefer to separate gains, but 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. The bias toward concentration is one potential explanation for the inconclusiveness of this evidence.

IV.A. 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 consideration set in a period is the set of lifetime consumption profiles associated with current options, given the agent's 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 agent's focus. Our formulation reflects the idea, broadly consistent with evidence on narrow bracketing (e.g., Tversky and Kahneman 1981; Rabin and Weizsäcker 2009), that focus is determined by the perceived consequences of only the decision at hand rather than of 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 decision maker'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 decision maker'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:

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

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

Our framework of choice given beliefs about future behavior can be combined with any theory of how these beliefs are formed. Following standard economic methodology, in this article 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 decision maker believes that she will act in the future as she would now commit to do.[13] The insights in this article hold for both theories of how the agent forms beliefs about the future.

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 with an additional period $t = 0$, such that the agent makes all decisions at $t = 0$, choosing from the set C^*_{ante} of all lifetime consumption-utility profiles.

IV.B. *Investing in the Future, One Step at a Time*

We develop our central insights regarding choice over time by applying the framework to a simple investment problem. There are $T_i + T_b$ periods, $t = 1, \ldots, T_i + T_b$. The agent makes an effort decision $e_t \in \{0, 1\}$ in each of the first $T_i \geq 1$ periods, and her investments provide benefits in the last $T_b \geq 1$ periods. Investment in a period $t \leq T_i$ generates a consumption-utility cost of $e_t \cdot B$ in period t and consumption-utility benefits of $\frac{A(e_t \cdot B)}{T_b}$ in *each* of the last T_b periods.[14] The variable A measures the efficiency of investment: Investing maximizes consumption utility

13. In this formulation, the decision maker'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. Although different in the specific theory of behavior, this perspective of the decision maker'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.

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.

if and only if $A \geq 1$. Abstracting away from any time discounting in consumption utility, we assume that $\delta = 1$.

The key features of the 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.[15] Of course, to capture these basic forces in a technically convenient way, we have imposed a number of unrealistic assumptions. Most important, while the assumption that the investment and benefit periods are temporally separated conveniently ensures that potential per period costs are equal across the investment periods and potential per period benefits are equal across the benefit periods, 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. In this case, the per period costs and benefits may be different across periods, and (depending on the application) 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.[16]

We begin by illustrating some of our formal results on ex ante and ex post choices in the decision problem using two numerical examples. Our first example is the simplest possible intertemporal decision problem.

Example 1 (Massage). A consumer must decide whether to have 30 minutes of massage in period 1 or 60 minutes of massage in period 2. She derives instantaneous utility of 80 from 30 minutes of massage and instantaneous utility of 100 from 60 minutes of massage. ($T_i = T_b = 1,\ B = 80,\ A = 1.25$)

15. Although we specified our dynamic model as deterministic, when we consider real-life examples, following Remark 5 we map stochastic outcomes to our deterministic setting using expected utility in a period as the relevant outcome.

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

In both the ex ante and ex post problems in Example 1, the consumer's choice is between $(80, 0)$ and $(0, 100)$—a balanced choice in which, by Proposition 3, she maximizes consumption utility. This example illustrates that, in contrast to hyperbolic discounting and the laboratory evidence supporting it, our model predicts no present bias or time inconsistency in single choices trading off utility at two different dates. In other types of decision problems, however, the situation is different:

EXAMPLE 2 (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 $(-80, 1, \ldots, 1)$ and $(0, \ldots, 0)$, where the 80 is the current pain and the 1's are the future benefits from exercising in that period. Proposition 1's bias toward concentration implies that if $1.25 < \frac{g(80)}{g(1)}$, the consumer does *not* exercise.

Now consider what the consumer would commit to in Example 2. 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 $(0, \ldots, 0)$, and these extreme options span the range of possible payoffs. Hence, as a special case of Proposition 3, the consumer commits to the consumption-utility-maximizing option of always exercising.

In Example 2, therefore, the consumer exhibits time inconsistency in behavior similar to that in models of hyperbolic discounting: 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 regular exercise accumulate into large health gains in the benefit periods, 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.

A key insight emerging from the comparison of Examples 1 and 2 is that—differently from hyperbolic discounting—our model predicts variation in the extent of present bias and time inconsistency across situations. In the rest of this section we develop formal results about how the environment shapes present bias and time inconsistency, explore the empirical content of these new comparative statics, and identify some welfare implications of our model.

IV.C. Patterns in Present Bias and Time Inconsistency

We first state our main formal result on the agent's behavior and biases. Notice that because of linearity and symmetry, in both the ex ante and ex post problems the agent prefers to exert effort either in all periods or in no period, and is indifferent only in a knife-edge case. Let A^*_{ante} and A^*_{post} be the cutoff levels of A above which the agent chooses to exert effort in the ex ante and ex post problems, respectively. For each $i = ante, post$, we say that the decision maker 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*—that is, she is more present-oriented ex post than ex ante—if $A^*_{ante} < A^*_{post}$. Note that time inconsistency in our model is not necessarily associated with present bias in ex post choice.

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 $h(1) = 1$, $A^*_{post} = h(T_b)$, and $A^*_{ante} = h(\frac{T_b}{T_i})$.

Proposition 5 implies a generalization of Example 2 to any setting in which repeated decisions have accumulating effects dispersed over multiple future periods ($T_b, T_i > 1$). In these situations, the Proposition implies that the agent is present-biased in ex post choice ($A^*_{post} > 1$), and 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.[17]

Furthermore, our theory identifies novel, simple, and testable comparative statics on present bias and time inconsistency:

COROLLARY 1. Suppose $g(\cdot)$ is continuous. Then:

(1) (Comparative Statics on Present-Biased Behavior) A^*_{post} is increasing in T_b and is independent of T_i.
(2) (Comparative Statics on Time-Inconsistent Behavior) Fixing T_b, $\frac{A^*_{post}}{A^*_{ante}}$ is increasing in T_i.

1. Comparative Statics on Present-Biased Behavior. Part 1 of Corollary 1 says that an increase in the number of periods in which the consequences of current misbehavior are dispersed (T_b) increases present bias. Intuitively, an increase in T_b dilutes the benefits but not the cost of exerting effort today, leading the agent to focus relatively less on the benefits. In hyperbolic discounting, in contrast, the agent's evaluation of an option depends only on its *total* (discounted) future cost, not on how that cost is distributed across periods. For example, because both in Example 1 and in Example 2 the total future cost of current misbehavior is 100, hyperbolic discounting predicts the same behavior in the two examples.

The comparative static may explain a lot of variation in present bias across situations. As a potentially important example, it 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

17. See Laibson (1997); Shui and Ausubel (2004); Laibson, Repetto and Tobacman (2007); Skiba and Tobacman (2008); and Meier and Sprenger (2010) on consumption-savings and borrowing. O'Donoghue and Rabin (1999a, 2006) and Gruber and Kőszegi (2001, 2004) on the consumption of harmful products, and DellaVigna and Malmendier (2004, 2006) and Acland and Levy (2010) on exercise. Indeed, unlike for most phenomena in behavioral economics, the evidence for present bias and time inconsistency is strongest in these field settings—where it is consistent with our model—and much weaker in laboratory decisions trading off current benefits with concentrated future costs—where it is inconsistent with our model.

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 one pound per week, and this intervention led participants to lose 13.1 pounds on average.[18] Finkelstein et al. (2007) also find significant effects of moderate incentives on weight loss, and Lussier et al. (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 foregoing kind of sensitivity is puzzling for a simple reason: The primary cost of consuming many harmful products 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 a 5% reduction in the mortality cost of smoking in a model of hyperbolic discounting, assuming an initial mortality cost of $100 per pack and a short-run discount factor of $\beta = .6$.[19] The reduction in mortality cost should have the same effect on smoking as a price decrease of $0.6 \cdot 0.05 \cdot \$100 = \$3$ per pack, which, assuming a price of $6 per pack, is a price drop of 50%![20] Extrapolating the local price elasticity of -0.6, the reduction in mortality cost should increase smoking by 30%. This seems implausibly high: We are not aware of

18. 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 this was no longer statistically significantly greater than the weight lost by the control group. Whether and how monetary incentives can be used to induce sustained weight loss is an open question.

19. Viscusi and Hersch (2008) estimate that the present value of the mortality cost of smoking amounts to $94–$222 per pack. Existing estimates of β typically fall in the 0.5–0.9 range, with a concentration of observations around 0.7–0.8. For example, Thaler's (1981) laboratory evidence implies a one-month discount factor of 0.75; Shui and Ausubel (2004) estimate a β of around 0.8 based on a structural model of credit card take-up and borrowing, and Laibson, Repetto, and Tobacman (2007) find $\beta = 0.7$ using a structural model of life-cycle consumption.

20. This calculation assumes—as we do below for our model—that the consumer experiences monetary payments as immediate disutility, so that the financial cost of smoking is immediate whereas the health cost is delayed.

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, and Ippolito and Ippolito (1984) find a very low value of life when estimating it based on the responsiveness of smoking to health information.

In contrast to standard theories, our model can accommodate a combination of high responsiveness to monetary incentives and low responsiveness to health incentives. As we have emphasized, in the examples the per period cost of current misbehavior (smoking or overeating today) is tiny but accrues in many future periods. Hence, the focus weight on these costs is close to $g(0)$, so that if $g(0)$ is low, our model says that the agent puts low weight on the future health consequences of misbehavior. At the same time, assuming (based on Remark 4) that she evaluates monetary transactions in part as immediate utility, she may put much higher weight on the current or future concentrated monetary consequences of misbehavior. We illustrate this argument using a simple formal example in Appendix A.

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.

2. Comparative Statics on Time-Inconsistent Behavior. Part 2 of Corollary 1 implies a sense in which the degree of time inconsistency is increasing in the number of investment periods T_i. To illustrate, recall Example 2 on exercise. As we have shown, if the agent decides whether to commit to exercising every period, she chooses to commit. But if she decided whether to commit to exercising on a *single* future date, her commitment decision would be identical to an ex post decision, in which—as we have also shown—she might decide not to exercise. Intuitively, the combined impact of multiple accumulating investments is greater and hence attracts greater attention if the number of investment periods (T_i) is higher, inducing more future-oriented behavior. Hyperbolic discounting, in contrast, predicts that if a person is unwilling to commit to any one of a sequence of actions with additive accumulating effects, she is also unwilling to commit to the entire sequence of actions.

A testable prediction and potentially important implication of our insight is that a person is more likely to commit to a longer sequence of future-oriented decisions—such as quitting smoking for a lifetime, rather than for a week—because with the longer commitment she has a more salient impact on her life. Although the evidence is subject to multiple plausible interpretations, this 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% of individuals offered the account claimed that they wanted to stop smoking in a year, and 53% said that they needed help to quit, only 11% took up the account. Similarly, although this comparison must be interpreted with caution due to the different accounts and populations, only 28% took up the short-term commitment savings account SEED offered by Ashraf, Karlan, and Yin (2006), but 78% took up the long-term SMarT account offered by Thaler and Benartzi (2004).[21]

21. 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 is the following. A person is more present-oriented in ex post choice than in ex ante choice if (as in Example 2) the range of *future benefits* is narrower in ex post than in ex ante choice, because in that

IV.D. 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.[22] We emphasize here two potentially important new welfare implications of our theory. First, 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 nonconstant 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 any 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 2 if $\alpha = \alpha' = 1$. But suppose instead that $\alpha = 1$ and $\alpha' = 0.75$, so that period 100's exercise session has future per period benefits of only 0.75. In this example, the extreme choices determining the consumer's

---

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

22. Existing research either imposes exogenously that ex ante preferences reflect welfare—thereby dismissing a self's high relative weight on the present as always welfare-irrelevant—or maintains that a welfare judgment cannot be made when ex post and ex ante choices differ. For examples of the former view, see DellaVigna and Malmendier (2004), Gruber and Kőszegi (2004), and O'Donoghue and Rabin (2006); for examples of the latter view, see Laibson (1997), Bernheim and Rangel (2008), and Asheim (2008).

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 exercising every day—even on the day in which this is not worth it. This occurs for the same reason that she is time inconsistent in Example 2: Because she is considering multiple decisions with accumulating benefits, she puts higher weight on the benefits of each individual exercise session than she would if she was considering the session in isolation. 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 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 choices with constant marginal benefits and binary choices, and can be illustrated using the following example:

Example 4 (Training for a Marathon). In each of the first 100 periods, a runner can train to improve her performance at a marathon, which occurs in period 101. A period's training has a utility cost of 5 in that period and increases the utility from the marathon by 4. The runner does not care about the health effects of training. ($T_i = 100, T_b = 1$, $B = 5$, $A = 0.8$.)

Without commitment, the runner'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 training and the 4 is the resulting improvement in her experience at the marathon. By Proposition 3, therefore, in each period the runner makes the consumption-utility-maximizing choice of not preparing for the marathon. From an ex ante perspective, however, the runner'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 1, therefore, the runner is too prone to commit to preparing for the marathon.

Like the consumer in Example 2, therefore, the runner exhibits time inconsistency. In contrast to that exercise-for-health example, however, in this case our model says that the runner'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 runner focuses on the large concentrated gain from having a great marathon and pays relatively less attention to the dispersed everyday training costs, she tends to agree to overpreparing. But when considering whether to train on an individual day ex post, the runner compares effort that day with just a marginally better marathon, making the effort less appealing.

More generally, if $1 \approx T_b < T_i$, Proposition 5 implies that 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 relies on the utility of money carrying sufficiently large weight relative to the utility from consumption, however, at this point it is only a tentative implication of our model.[23]

23. 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 et al. (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.

TABLE I
A SUMMARY OF THE MAIN PREDICTIONS OF FOCUSING AND HYPERBOLIC DISCOUNTING

	Focusing	Hyperbolic disc.
Behavior		
Time inconsistency in single choice	no	yes
Time inconsistency in lifestyle choices	yes	yes
Higher responsiveness to concentrated than to dispersed incentives	yes	no
More likely to make long-term than short-term commitments	yes	no
Welfare		
Present bias in single choice	no	yes
Present bias in lifestyle choices	yes	yes
Ex ante perspective generally better for welfare	yes	yes
Overcommitment to concentrated goals	yes	no
Overcommitment to overall beneficial lifestyle	yes	no

IV.E. Focusing versus Hyperbolic Discounting

Table I summarizes the key predictions of our focusing model for intertemporal choice and contrasts them with those of hyperbolic discounting. Because our theory makes the same predictions as hyperbolic discounting in exactly those lifestyle choices in which the latter theory has been most invoked and documented, but also makes different and realistic comparative-statics predictions, it likely identifies a mechanism that is important in intertemporal choice. Yet because our theory does not match evidence in single choices, 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.

V. OTHER POTENTIAL APPLICATIONS AND CONCLUSION

By virtue of defining focus-dependent utility based on consumption utility and the decision maker's consideration set, our theory opens the way for analyzing the role of focus in many economic settings using one generally applicable model. A recipe for translating a deterministic classical model into one with focus-dependent choice—and hence for taking our theory to other economic domains—is to (1) specify the relevant attributes in the given setting; (2) take the utility function from the classical model; and (3) equate the consideration set with the choice set of the classical model.[24]

In addition to choice over time and social preferences, two applications we have considered in this article, our framework also has potential use for understanding profit-maximizing firms' product design and pricing behavior, helping incorporate into economics a key marketing question: how firms design and position products to manipulate consumer attention.[25] As a simple example, the bias toward concentration implies that a firm has an incentive to concentrate product value on a single attribute, a prediction that seems consistent with marketing analyses of successful brand positioning and consumer "value propositions." Further natural questions include which value attributes firms concentrate on in different situations, and how firms design, bundle, and price products to take advantage of consumers' distorted focus when making purchase decisions. For example, as an analog 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 two other likely worthwhile applications. Research indicates that retirees take too much of their retirement wealth in one lump sum rather than as an annuity,[26] and our model says that this may simply happen because a lump-sum payment looks very large relative to an annuity's monthly payments. And our theory predicts that an employee may be

24. A similar recipe works for applying our model to a new setting for which no appropriate classical model exists: specify (1) the relevant attributes; (2) the utility function; and (3) the consideration set. Note that in such an application, a classical model would also have to make assumptions corresponding to (2) and (3).

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

26. 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 the tendency to under-annuitize, 59% of respondents in the Health and Retirement Survey report that they would accept $500 less in Social Security benefits in exchange for one 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).

motivated by some features of her employment contract—such as a bonus, major promotion, or other large goal—not only because they can generate higher consumption utility, but also because of her disproportionate focus on these features.

Being defined for riskless choices only, our model is (without additional assumptions) not applicable to situations in which uncertainty is a central part. Uncertainty raises a new conceptual issue: how to define attributes with uncertain payoffs. To illustrate two extreme specifications, suppose that smoking a cigarette increases the probability of developing lung cancer at later dates by a tiny amount. If state-by-state payoffs in a period are the attributes, as in Bordalo, Gennaioli, and Shleifer (2012a, 2012b), 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 expected utility in a period is the attribute, the trivial change in probability leads the consumer to underweight the possibility of lung cancer, making smoking much more attractive. Whether state-by-state payoffs or expected utilities are the "right" attributes seems to depend on the nature of the particular decision. We believe that states are the natural attributes in decisions—such as Kahneman-Tversky-style choices between lotteries—in which the state-space representation of uncertainty is explicit, while expected utilities are more natural in decisions—such as smoking—in which there is no immediate state-space representation. Combining these possibilities in a single model is an agenda for future research.

Appendix A. Utility from Monetary Transactions

In this article, we model how the agent thinks about the consequences of spending or receiving money in two different ways. As discussed in Section IV and formalized in greater generality in Appendix C, 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 $0 < \lambda < 1$.[27] With this extension to a richer notion of consumption utility, we simply apply our model of dynamic decision making from Section IV.A. 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.

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

27. At the cost of some conceptual (and notational) complexity, our framework can be extended to nonlinear consumption utility. The main conceptual complication is to determine the appropriate assumption regarding the 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.

1. 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 IV.B, the agent is deciding whether to consume 80 now or 1 in each of 100 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,..., 0)
	Money	(−80, 0,..., 0)
1's later	Consumption	(0, 1,..., 1)
	Money	(0, −1,..., −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 IV.B—holds if $\frac{g(80)}{g(1)}$ is sufficiently large. More generally, unless $g(\cdot)$ is too concave, the agent is always present-biased in an ex post consumption-savings decision.

Intuitively, monetary utility introduces a "pain of paying" for consumption that counteracts the agent's bias toward current consumption we identified in Section IV.B: 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.

2. 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 III, 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.

3. 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 price of zero to start with, and that the person initially believes that smoking has no future health cost.[28] 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 consumer 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

28. Although 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.

decision irrelevant) health attribute, the consumer's choice in case (a) is:

Smoke?	Attribute	Profile
No	Smoking	(0, 0, ..., 0)
	Consumption	(0, 0, ..., 0)
	Money	(0, 0, ..., 0)
Yes	Smoking	(B, 0, ..., 0)
	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.

Appendix 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 (1) we know how products map into attributes and there are at least three attributes; (2) we can manipulate individual attributes in the decision maker's choices; (3) $g(\cdot)$ is strictly increasing; and (4) 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 decision maker 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.[29] For some p satisfying $0 < p < x$, we find the $\delta_x(p)$ that makes the decision maker 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 toward 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. (1) We find the amount of money, m, such that the consumer is indifferent whether to pay m for the data plan alone. (2) 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. (3) 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

29. 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.

indifferent, whereas 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 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 (1) and (2) 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 (3) 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 (1) and (2) 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.

Appendix 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 IV extend to this setting. Consider a consumer who lives for a finite number $T > 1$ of periods, has consumption utility

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

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

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

for $t = 1, \ldots, T$ 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 $0 < \alpha \leq 1$.[30]

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.[31] By allowing for equality in the condition we accommodate the useful benchmark $\delta = 1$ and $R = 0$, in which consumption in each period is $C_t = \frac{W_1}{T}$.

30. If the coefficient of relative risk aversion is greater than 1, then $u(C)$ goes to minus infinity as C approaches 0, 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 decision maker 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.

31. 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.

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 $b_t^i = \frac{C_t^i}{W_t^i}$ 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 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 IV. 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:

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

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 (1997), 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 $t+2,\ldots, T$, and $h(1-b_{t+1})$ can be interpreted as the focus weight associated with these remaining periods. Finally, the term involving $\delta(1+R)^{\alpha}$ reflects the intuition 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 $t+2,\ldots, T$ 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, $u'(C_{t+1})$ must equal the marginal benefit of saving an extra dollar:

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

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

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

Now note that

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

and hence for all $\tau \geq t+1$, $\frac{g_t^\tau}{g_{t+1}^\tau} = g(\delta u((1+R)(1-b_{t+1})))$, 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

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

Given that $b_T^{post} = 1$, we can use this equation to recursively solve for $b_{T-1}^{post}, \ldots, b_1^{post}$. 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 $W_1(1+R)^{t-1}$, the focus weight over utility in that period, using our functional forms, is $W_1^{\alpha\theta}[\delta(1+R)^\alpha]^{\theta(t-1)}$. 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 equation (4) substituting $\theta = 0$ and the discount factor γ to obtain

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

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

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

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 $b_t^{post} < b_t^{ante}$.

APPENDIX D. PROOFS

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

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

where we use that $\Delta_k(C) > F$ 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 nonempty. Thus the agent does not choose c' over c. ■

Proof of Proposition 2. As in the previous proof, assume that the vectors $c \in C$ are measured in utility terms. The assumptions of the Proposition imply that $\Delta_k(C) = c_k^1 - c_k^2$ and $\Delta_l(C) = c_l^1 - c_l^2$. Denoting $C' = C \setminus \{c^1\} \cup \{c'\}$, for any $c^3 \in C' \setminus \{c'\} = C \setminus \{c^1\}$ we have

$$\begin{aligned}\tilde{U}(c', C') - \tilde{U}\big(c^3, C'\big) &= g(\Delta_k(C) + \Delta_l(C))(\{c_k^1 - c_k^3\} + \{c_l^1 - c_l^2\}) \\ &\quad + g(\Delta_l(C'))(c_l^2 - c_l^3) + \sum_{j \neq k,l} g(\Delta_j(C)) \cdot (c_j^1 - c_j^3) \\ &> g(\Delta_k(C) + \Delta_l(C))(c_k^1 - c_k^3) + g(\Delta_l(C))(c_l^1 - c_l^2) \\ &\quad + g(\Delta_l(C))(c_l^2 - c_l^3) + \sum_{j \neq k,l} g(\Delta_j(C)) \cdot (c_j^1 - c_j^3) \\ &\geq g(\Delta_k(C))(c_k^1 - c_k^3) + g(\Delta_l(C))(c_l^1 - c_l^3) + \\ &\quad \sum_{j \neq k,l} g(\Delta_j(C)) \cdot (c_j^1 - c_j^3) \\ &= \tilde{U}\big(c^1, C\big) - \tilde{U}\big(c^3, C\big) \\ &\geq 0,\end{aligned}$$

where the first equality uses the assumption that c' concentrates the advantages of c^1 over c^2 in attributes k, l; the following inequality uses that $g(\cdot)$ is strictly increasing, that $\Delta_k(C) > 0$, that $c_l^1 - c_l^2 > 0$, that $\Delta_l(C) \geq \Delta_l(C')$, and that $c_l^3 \geq c_l^2$; and the next inequality uses that $g(\cdot)$ is increasing and $c_l^3 \geq c_l^2$. It follows that the agent strictly prefers c' over any alternative in C'. ■

Proof of Proposition 3. Continue to assume that the vectors in C are measured in utility terms. Suppose c^1 is a utility-maximizing option, and assume $c_k^2 = 0$ for all k. Denote by A the K' attributes in which c^1 is superior to c^2, and by B the attributes in which c^1 is inferior to c^2. Since the utility of c^1 exceeds that of c^2, we have $p \geq m$.

Take any alternative $c' \in C$ with $U(c') < U(c^1)$. Then,

$$\sum_{i \in A} (p - c_i') > \sum_{i \in B} (c_i' + m)$$

where the spanning assumption implies that all terms in both sums are nonnegative. Multiplying the left-hand side by $\frac{g(p)}{g(m)} \geq 1$ implies

$$\sum_{i \in A} g(p)(p - c_i') > \sum_{i \in B} g(m)(c_i' + m)$$

which means that the focus-weighted utility of c^1 exceeds that of c'. Hence, the agent does not choose 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 $\frac{P}{M}$ are held fixed, and hence changes in K_p and K_m affect the inequality only through the $\frac{K_p}{K_m}$ term on the right-hand side.

To prove in the case $\frac{K_p}{K_m} \leq 1$, first assume that $P < M$, so that the right choice is c^2. Since $g(\cdot)$ is strictly increasing, if the agent is willing to choose c^2 (i.e., the above inequality is violated) for $\frac{K_p}{K_m} = R'$, then she strictly prefers to choose c^2 (the above inequality goes strictly the other way) for $\frac{K_p}{K_m} = R > R'$. 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. The case $\frac{K_p}{K_m} \geq 1$ can be shown analogously. ■

Proof of Proposition 5. 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 $\Delta_s = \frac{AB}{T_b}$. As a result, the consumer chooses to invest if $Bg(B) < ABg(\frac{AB}{T_b})$. Since the right-hand side is continuous and increasing in A, the value A^*_{post} is defined by $g(B) = A^* \cdot g(\frac{A^*B}{T_b})$. Let the function $h(\cdot)$ be implicitly defined by the equation $g(B) = h(z) \cdot g(\frac{Bh(z)}{z})$. 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, $A^*_{post} = h(T_b)$.

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 $\Delta_s = AB \cdot \frac{T_i}{T_b}$ for $s = T_i + 1, \ldots, T_i + T_b$. Effort is strictly preferred if and only if $Bg(B) < AB(AB \cdot \frac{T_i}{T_b})$, and A^*_{ante} is defined by $g(B) = A^* \cdot g(A^*B \cdot \frac{T_i}{T_b})$. The definition of $h(\cdot)$ implies that $A^*_{ante} = h(\frac{T_b}{T_i})$. ■

Proof of Corollary 1. Immediate. ■

Central European University and
University of California at Berkeley
Central European University, NBER, and CEPR

References

Abaluck, Jason, "What Would We Eat if We Knew More: The Implications of a Large-Scale Change in Nutrition Labeling," Working Paper, MIT, 2011.
Acland, Dan, and Matthew R. Levy, "Habit Formation, Naivete, and Projection Bias in Gym Attendance," Working Paper, Harvard University, 2010.
Akerlof, George A., "Procrastination and Obedience," *American Economic Review*, 81, no. 2 (1991), 1–19.
Asheim, Geir B., "Procrastination, Partial Naivete, and Behavioral Welfare Analysis," Working Paper, University of Oslo, 2008.
Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 121 (2006), 635–672.
Bernheim, B. Douglas, and Antonio Rangel, "Beyond Revealed Preferences: Choice Theoretic Foundations for Behavioral Welfare Economics," NBER Working Paper No. 13737, January 2008.
Bertaut, Carol C., Michael Haliassos, and Michael Reiter, "Credit Card Debt Puzzles and Debt Revolvers for Self Control," *Review of Finance*, 13, no. 4 (2009), 657–692.
Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," Working Paper, Harvard University, 2012a.
———, "Salience Theory of Choice Under Risk," *Quarterly Journal of Economics*, 127 (2012b), 1243–1285.
Brown, Jeffrey R., "Understanding the Role of Annuities in Retirement Planning," in *Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs*, Annamaria Lusardi, ed., 178–208 (Chicago: University of Chicago Press, 2009).
Brown, Jeffrey R., Marcus D. Casey, and Olivia S. Mitchell, "Who Values the Social Security Annuity? New Evidence on the Annuity Puzzle," NBER Working Paper No. 13800, 2008.
Carroll, Christopher D., "Theoretical Foundations of Buffer Stock Savings," Working Paper, Johns Hopkins University, 2011.
Chiu, Andrew, and George Wu, "Prospect Theory," in *Wiley Encyclopedia of Operations Research and Management Science*, James J. Cochran, ed. (New York: Wiley, 2009). Cunningham, Tom, "Comparisons and Choice," Working Paper, 2011.
Cunningham, Tom, "Comparisons and Choice," Working Paper, 2011.
Davidoff, Thomas, Jeffrey R. Brown, and Peter A. Diamond, "Annuities and Individual Welfare," *American Economic Review*, 95, no. 5 (2005), 1573–1590.
DellaVigna, Stefano, and Ulrike Malmendier, "Contract Design and Self-Control: Theory and Evidence," *Quarterly Journal of Economics*, 119 (2004), 353–402.
———, "Paying Not to Go to the Gym," *American Economic Review,* 96, no. 3 (2006), 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*, 29 (2003), 1421–1432.
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*, 49, no. 9 (2007), 981–989.
Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality", Working paper, NYU Stern, 2011, http://pages.stern.nyu.edu/~xgabaix/papers/sparsebrconsumer.pdf.

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,* 2, no. 4 (2010), 213–235.
Gourville, John T., "Pennies-a-Day: The Effect of Temporal Reframing on Transaction Evaluation," *Journal of Consumer Research*, 24, no. 4 (1998), 395–408.
Gruber, Jonathan, and Botond Kőszegi, "Is Addiction 'Rational?' Theory and Evidence," *Quarterly Journal of Economics*, 116, no. 4 (2001), 1261–1305.
———, "Tax Incidence When Individuals Are Time-Inconsistent: The Case of Cigarette Excise Taxes," *Journal of Public Economics*, 88, nos. 9–10 (2004), 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*, 8, no. 1 (2008).
Ippolito, Pauline M., and Richard A. Ippolito, "Measuring the Value of Life Saving from Consumer Reactions to New Information," *Journal of Public Economics*, 25, nos. 1–2 (1984), 53–81.
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*, 312, no. 5782 (2006), 1908–1910.
Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, no. 2 (1979), 263–291.
Kőszegi, Botond, and Ádám Szeidl, "A Model of Focusing in Economic Choice," Working Paper, UC Berkeley, 2011, http://elsa.berkeley.edu/~botond/focusing.pdf.
Laibson, David I., "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112 (1997), 443–477.
Laibson, David I., Andrea Repetto, and Jeremy Tobacman, "Estimating Discount Functions with Consumption Choices over the Lifecycle," NBER Working Paper No. 13314, 2007.
Lehenkari, Mirjam, "The Hedonic Editing Hypothesis: Evidence from the Finnish Stock Market," *Journal of Behavioral Finance*, 10, no. 1 (2009), 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*, 75, no. 1 (1998), 5–18.
Linville, Patricia W., and Gregory W. Fischer, "Preferences for Separating or Combining Events," *Journal of Personality and Social Psychology*, 60, no. 1 (1991), 5–23.
Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 118 (2003), 81–123.
Loomes, Graham, and Robert Sugden, "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty," *Economic Journal*, 92, no. 368 (1982), 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*, 101, no. 2 (2006), 192–203.
Meier, Stephan, and Charles Sprenger, "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2, no. 1 (2010), 193–210.
Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer, "Coarse Thinking and Persuasion," *Quarterly Journal of Economics*, 123, no. 2 (2008), 577–619.
O'Donoghue, Ted, and Matthew Rabin, "Addiction and Self-Control," in Addiction: Entries and Exits, Jon Elster, ed. (New York: Russell Sage Foundation, 1999a).
———, "Doing It Now or Later," *American Economic Review*, 89, no. 1 (1999b), 103–124.
———, "Optimal Sin Taxes," *Journal of Public Economics*, 90, nos. 10–11 (2006), 1825–1849.

Prelec, Drazen, and George Loewenstein, “The Red and the Black: Mental Accounting of Savings and Debt,” *Marketing Science*, 17, no. 1 (1998), 4–28.
Rabin, Matthew, and Georg Weizsäcker, “Narrow Bracketing and Dominated Choices,” *American Economic Review*, 99, no. 4 (2009), 1508–1543.
Read, Daniel, George Loewenstein, and Matthew Rabin, “Choice Bracketing,” *Journal of Risk and Uncertainty*, 19, nos. 1–3 (1999), 171–197.
Rick, Scott, and George Loewenstein, “Intangibility in Intertemporal Choice,” *Philosophical Transactions of the Royal Society of London*, 363, no. 1511 (2008), 3813–3824.
Rubinstein, Ariel, “Similarity and Decision-Making Under Risk (Is There a Utility Theory Resolution to the Allais Paradox?),” *Journal of Economic Theory*, 46, no. 1 (1988), 145–153.
Schkade, David A., and Daniel Kahneman, “Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction,” *Psychological Science*, 9, no. 5 (1998), 340–346.
Schwartzstein, Joshua, “Selective Attention and Learning,” Working Paper, Dartmouth College, 2012.
Scitovsky, Tibor, *The Joyless Economy: The Psychology of Human Satisfaction* (New York: Oxford University Press, 1976).
Shui, Haiyan, and Lawrence M. Ausubel, “Time Inconsistency in the Credit Card Market,” 14th Annual Utah Winter Finance Conference, 2004. Available at http://ssrn.com/abstract=586622.
Sindelar, Jody L., “Paying for Performance: The Power of Incentives over Habits,” *Health Economics*, 17, no. 4 (2008), 449–451.
Sindelar, Jody, Brian Elbel, and Nancy M. Petry, “What Do We Get for Our Money? Cost-Effectiveness of Adding Contingency Management,” *Addiction*, 102, no. 2 (2007), 309–316.
Skiba, Paige Marta, and Jeremy Tobacman, “Payday Loans, Uncertainty, and Discounting: Explaining Patterns of Borrowing, Repayment, and Default,” Vanderbilt Law and Economics Research Paper No. 08-33, 2008.
Spiegler, Ran, *Bounded Rationality and Industrial Organization* (New York: Oxford University Press, 2011).
———, “Competitive Framing,” Mimeo, Tel Aviv University, 2012.
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*, 99, no. 2 (2009), 424–429.
Thaler, Richard, “Some Empirical Evidence on Dynamic Inconsistency,” *Economics Letters*, 8 (1981), 201–207.
———, “Mental Accounting and Consumer Choice,” *Marketing Science*, 4, no. 3 (1985), 199–214.
———, “Mental Accounting Matters,” *Journal of Behavioral Decision Making,* 12, no. 3 (1999), 183–206.
Thaler, Richard, and Shlomo Benartzi, “Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving,” *Journal of Political Economy*, 112, no. 1 (2004), S164–S187.
Thaler, Richard, and Eric J. Johnson, “Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice,” *Management Science*, 36, no. 6 (1990), 643–660.
Tversky, Amos, “Intransitivity of Preferences,” *Psychological Review*, 76, no. 1 (1969), 31–48.
Tversky, Amos, and Daniel Kahneman, “The Framing of Decisions and the Psychology of Choice,” *Science*, 211 (1981), 453–458.
Viscusi, W. Kip, and Joni Hersch, “The Mortality Cost to Smokers,” *Journal of Health Economics*, 27 (2008), 943–958.
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*, 300, no. 22 (2008), 2631–2637.
Yaari, Menahem E., “Uncertain Lifetime, Life Insurance, and the Theory of the Consumer,” *Review of Economic Studies*, 32, no. 5 (1965), 137–150.