American Economic Review 2009, 99:3, 909–936
http://www.aeaweb.org/articles.php?doi=10.1257/aer.99.3.909

Reference-Dependent Consumption Plans

By Botond Kőszegi and Matthew Rabin*

We develop a rational dynamic model in which people are loss averse over changes in beliefs about present and future consumption. Because changes in wealth are news about future consumption, preferences over money are reference-dependent. If news resonates more when about imminent consumption than when about future consumption, a decision maker might (to generate pleasant surprises) overconsume early relative to the optimal committed plan, increase immediate consumption following surprise wealth increases, and delay decreasing consumption following surprise losses. Since higher wealth mitigates the effect of bad news, people exhibit an unambiguous first-order precautionary-savings motive. (*JEL* D14, D81, D83, D91)

In this paper, we develop a general dynamic model of reference-dependent utility. Building on ideas and models in Astrid Matthey (2005), Christopher K. Hsee and Clair I. Tsai (2008), Miles Kimball and Robert Willis (2006), and Kőszegi and Rabin (2006, 2007), we assume that utility depends on recent changes in rational beliefs about present and future consumption, and bad news is more painful than good news is pleasant.[1] We derive implications of our model for preferences over receiving information about an exogenous future event, monetary risk preferences, and intertemporal consumption decisions. If news about more imminent consumption is felt more heavily than news about distant consumption, a person prefers to receive the same information sooner rather than later, increases consumption immediately in response to good surprises regarding wealth but delays cuts following bad surprises, and—since surprising herself with extra immediate consumption is pleasant—may overconsume early in life relative to the optimal committed plan. To reduce the impact of losses she may suffer, the decision maker prefers to receive bits of information together rather than apart, and—to lower the marginal utility associated with any surprise losses—prepares for future uncertainty by increasing savings.

We present the basic framework in Section I. In each period $t \in \{1, \ldots, T\}$, the decision maker consumes a K-dimensional consumption bundle, $\mathbf{c}_t$. Overall instantaneous utility in period t is the sum of reference-independent "consumption utility" that derives purely from $\mathbf{c}_t$, and gain-loss utility that derives from recent changes in beliefs about consumption in each dimension in each

* Kőszegi: Department of Economics, University of California, Berkeley, 508-1 Evans Hall #3880, Berkeley, CA 94720 (e-mail: botond@econ.berkeley.edu); Rabin: Department of Economics, University of California, Berkeley, 508-1 Evans Hall #3880, Berkeley, CA 94720 (e-mail: rabin@econ.berkeley.edu). This paper updates and replaces earlier versions titled "Dynamic Reference-Dependent Preferences," first presented and posted in June 2006. We are grateful to Avinash Dixit and Neel Mukherjee, and seminar audiences at the Conference in Honor of Eytan Sheshinski, University of California, Berkeley, CERGE-EI, University of Chicago, Cornell University, University of Edinburgh, Harvard University, London School of Economics, University of Michigan, Princeton University, University of California, San Diego, University of Southern California, and University of Texas, Austin, and especially three anonymous referees for useful comments. We thank the National Science Foundation (grant SES-0648659) for financial support. Kristóf Madarász provided research assistance.

[1] The model in this paper is closely related to contemporaneous work by Matthey (2006). Building on her earlier intuitions in Matthey (2005), she develops a formal model of preferences that is similar to our formulation. But her solution concepts embed different assumptions from ours about how people form their beliefs, and she develops a different set of results.

period starting with t. The person experiences "contemporaneous" gain-loss utility from any contrast between current consumption and her prior expectations of current consumption, and "prospective" gain-loss utility from changes in her beliefs about future consumption. In all these comparisons, she is loss averse: bad news about consumption is more unpleasant than good news is pleasant. Normalizing the weight $\gamma_{t,t}$ on contemporaneous gain-loss utility to one, in period t the decision maker puts weight $\gamma_{t,\tau} \leq 1$ on prospective gain-loss utility regarding outcomes in period $\tau > t$. Her goal in period t is to maximize the sum of instantaneous utilities starting in period t.

Because preferences depend on the sequence of expectations, our model is complete only when combined with a theory of how expectations are formed. We follow much previous work on beliefs-based preferences and assume that beliefs must be rationally based on credible plans for state-contingent behavior.[2] In particular, an implemented plan must be a "preferred personal equilibrium": in each period it maximizes expected reference-dependent utility given the expectations generated by the plan, with the constraint that continuation plans must be consistent with a similar maximization in the future. In the Web Appendix (available at http://www.aeaweb.org/articles.php?doi=10.1257/aer.99.3.353) we propose a framework for thinking about the dynamic formation of rational beliefs and plans without imposing that *initial* beliefs about the future are necessarily correct, and explore some circumstances when this more basic rationality assumption justifies the use of preferred personal equilibrium.

In Section II, we explore some implications that our assumption of loss aversion over changes in beliefs has for informational preferences. On the one hand, when prospective gain-loss utility does not loom as large as contemporaneous gain-loss utility—so that possible bad news is less painful when outcomes are not imminent—the decision maker likes receiving the same information sooner rather than later. On the other hand, consistent with experimental evidence in Uri Gneezy and Jan Potters (1997) and others showing that people value the same asset less if they are to receive more interim information about its payout, the decision maker dislikes interim *piecemeal* information because it exposes her to possibly unnecessary bad news due to fluctuations in beliefs.[3]

In Section III, we show how our model might provide foundations for gain-loss utility over wealth, which is the central assumption of numerous models starting from Daniel Kahneman and Amos Tversky's (1979) original model of prospect theory. Because gains and losses in wealth generate utility-inducing news about the distribution of future consumption, the decision maker attends to these changes even if they are minuscule relative to the overall wealth risk she faces—and even if she holds wealth exclusively for future consumption. This insight also provides a new perspective on discussions of narrow bracketing in the literature: while treating a risk to wealth in isolation from future decisions and risks is typically considered a mistake, our model says that in some situations it is instead a manifestation of a preference over changes in beliefs.

Beyond providing foundations for reference-dependent monetary preferences, our theory helps unify much of the growing body of research on reference-dependent utility more generally. In contrast to prospect theory's focus on changes in wealth, a long-standing literature on "habit formation" assumes that people evaluate consumption in part by comparing it to consumption from

[2] For previous examples of this rational-expectations approach, see for instance Kőszegi and Rabin (2006, 2007) on reference-dependent preferences and Andrew Caplin and John Leahy (2001) and Kőszegi (2006b) on anticipatory utility. Another strand of models of beliefs-based preferences, such as George Akerlof and William T. Dickens (1982) and Markus Brunnermeier and Jonathan Parker (2005), assumes that agents can *choose* beliefs—even beliefs that are not consistent with rational expectations. Because agents are assumed to maximize their utility given their beliefs once the beliefs are chosen, they face a trade-off between forming beliefs that make them feel better and ones that help them make good choices.

[3] Our model shows a mechanism whereby loss aversion can generate "information-loving" and "information-averse" informational preferences by a decision maker, and as such complements models such as those of Caplin and Leahy (2001) and Kőszegi (2006a), which assume a taste or distaste for information as a primitive.

the recent past.[4] And a large literature building from Jack L. Knetsch (1989) and Kahneman, Knetsch, and Richard H. Thaler (1990) posits that willingness to pay for durable goods depends on recent ownership status. Since past consumption is in many circumstances a major determinant of expected future consumption, and changes in ownership status of a durable good are presumably pleasant or unpleasant mostly because of beliefs about the future use of that good, our theory is consistent with these approaches as well.

In Section IV, we present our main application, a two-period consumption-savings problem. A consumer with strictly concave consumption utility must split a possibly stochastic amount of lifetime wealth between periods 1 and 2, where we assume for simplicity that the interest rate is zero. We derive a number of behavioral and welfare results from the assumption that $\gamma \equiv \gamma_{1,2} < 1$, so that a deviation from expected period-1 consumption has a greater effect on utility than a similar change of period-1 plans regarding period-2 consumption. The most basic implication of this assumption is an asymmetry in the consumer's response to "surprises"—low-probability changes in her wealth. In the limit case, suppose the consumer had made plans expecting a given deterministic lifetime wealth, but then finds at the beginning of period 1 that her wealth will be a different deterministic level. Because the gain from increasing period-1 consumption is more pleasant than the gain from increasing plans for period-2 consumption, the decision maker fully consumes small increases in wealth in period 1. And because the loss from decreasing period-1 consumption is more painful than the loss from decreasing plans for period-2 consumption, she fully absorbs small decreases in wealth in period 2.

We also identify ways in which the consumer's behavior may be suboptimal among the strategies available to her. We replicate in our setting a result Rebecca Stone (2005) established for deterministic wealth, that expectations-based preferences can generate overconsumption for a completely different reason than present-biased preferences in the sense of David Laibson (1997) and Ted O'Donoghue and Rabin (1999), and temptation disutility in the sense of Faruk Gul and Wolfgang Pesendorfer (2001). To understand the intuition, suppose the consumer had made the ex ante optimal plan to consume an equal amount in the two periods. If γ is relatively small, pleasantly surprising herself with extra consumption in period 1 at the cost of lowering consumption plans for period 2 increases utility ex post, so that without commitment the optimal plan is not consistent.

Interestingly, our model differs from the self-control theories above in predicting a strong role of current uncertainty in exacerbating overconsumption: there are γ such that if the consumer faces sufficient uncertainty regarding her wealth, she overconsumes in period 1 for all wealth levels; yet for any wealth level on the support of her wealth distribution, if she knew in advance this would be her wealth level, she would not overconsume. Because the sense of gain from an upward revision of a deterministic period-1 consumption plan is often smaller than the sense of loss from the corresponding downward revision of a deterministic period-2 consumption plan, a deterministic plan for consumption acts as a bright-line commitment device. When there is uncertainty, however, the possibility of higher or lower consumption is already incorporated into expectations, so the implications of a revision in plans are evaluated much less asymmetrically.

Finally, we consider environments where the consumer's wealth is possibly stochastic, and uncertainty is resolved in period 2. Because an unusually low level of consumption utility is more painful than an unusually high level of consumption utility is pleasant, the consumer dislikes uncertainty in period-2 consumption utility. Saving more decreases the pain of uncertainty by lowering marginal utility. This novel type of precautionary-savings motive implies an

[4] See, for instance, James S. Duesenberry (1952), Harl E. Ryder Jr. and Geoffrey M. Heal (1973), and Tibor Scitovsky (1976), with David Bowman, Deborah Minehart, and Rabin (1999) more recently combining this approach with Kahneman and Tversky's prospect theory.

unambiguous first-order effect of increased uncertainty on savings, and seems more intuitive than expected-utility-of-wealth theories of precautionary savings.

We conclude the paper in Section V by noting further natural applications of our model, and especially pointing out some of its limitations.

I. The Model

We consider discrete-time models where there are $T + 1$ periods, 0 through T. Period 0 corresponds to the time at which the person first makes plans. Referring to decisions as the irreversible implementation of choices, we assume that the first relevant decision is made and the first relevant outcome occurs in period 1 or later. While in Web Appendix A we discuss situations and propose formal solution concepts for cases where the person had not been focusing on the decision problem until she is confronted with it, in the text we are by construction assuming planning begins before the implementation of decisions.

In each period $t \geq 1$, a K-dimensional consumption vector $\mathbf{c}_t = (c_t^1, \ldots, c_t^K)$ is realized. The timing within such a period is the following. The decision maker starts with beliefs $F_{t-1} = \{F_{t-1,\tau}\}\,|_{\tau=t}^{T}$ inherited from period $t-1$, where $F_{t-1,\tau} = (F_{t-1,\tau}^1, \ldots, F_{t-1,\tau}^K)$ are the beliefs regarding the K dimensions of consumption in period τ. Then some uncertainty may be resolved, and the decision maker takes an action. Further uncertainty may then be resolved, and the decision maker forms new beliefs $\{F_{t,\tau}\}_{\tau=t}^{T}$, where the beliefs $F_{t,t}$ assign probability 1 to $\mathbf{c}_t$.[5]

The decision maker's period-t instantaneous utility u_t depends on consumption in period t, and on the changes in period t to beliefs about contemporaneous and future consumption:

$$u_t = m(\mathbf{c}_t) + \sum_{\tau=t}^{T} \gamma_{t,\tau} N(F_{t,\tau}\,|\,F_{t-1,\tau}). \tag{1}$$

The term $m(\mathbf{c}_t)$ is "consumption utility," which can be thought of as corresponding to classical reference-independent utility. We assume that consumption utility is additively separable across dimensions, and that the consumption-utility function in dimension k, $m^k(\cdot)$, is differentiable and strictly increasing. The terms $N(F_{t,\tau}\,|\,F_{t-1,\tau})$ represent "gain-loss utility," and $\gamma_{\tau,\tau} \geq \gamma_{\tau-1,\tau} \geq \ldots \geq \gamma_{0,\tau} \geq 0$ are weights on these gain-loss utilities. We normalize $\gamma_{t,t} = 1$. For $\tau > t$, $N(F_{t,\tau}\,|\,F_{t-1,\tau})$ is "prospective gain-loss utility," which derives from changes between last period and this period in beliefs regarding future outcomes.[6,7] While notationally, substantively, and psychologically consistent with such prospective gain-loss utility, the functions $N(F_{t,t}\,|\,F_{t-1,t})$ can be usefully distinguished as "contemporaneous gain-loss utility." These derive from comparing the consumption

[5] While the within-period timing above is sufficient for all the applications in this paper, the model of preferences is fully compatible with other decision-making structures.

[6] While it would be more realistic to assume that comparisons to past beliefs matter beyond a one-period lag, it seems most essential intuitions of such preferences can be captured in our specification through assumptions about the within-period timing of information and decisions. Consider, for instance, the house-money effect predicted by Thaler and Eric J. Johnson (1990) and documented in a somewhat different form by Thierry Post et al. (2008), whereby people are more willing to spend or risk money when obtained unexpectedly than when anticipated. This can be modeled by assuming that a person receives money, and *in the same period* makes decisions on how to spend the money—with her old expectations still determining current preferences. Of course, because we assume the decision maker fully incorporates the news into her reference point by next period, our model does not accommodate a gradual adjustment of the reference point.

[7] By including anticipation of consumption in a given period in instantaneous utility in all prior periods, our formulation may appear to be "multiple-counting" consumption. Repeated changes in beliefs about the same future period can indeed have substantial utility effects—generating some interesting results below about aversion to incessant arrival of information. But it is *only* when beliefs are changing incessantly that perception of one period's consumption has a huge effect on prior periods' utility. In fact, when consumption in a given period implements plans that were always held, our specification below implies that prospective gain-loss utility regarding that period is zero in all prior periods.

outcome that occurs in period t to beliefs regarding that consumption the decision maker entered period t with.

The weights $\{\gamma_{t,\tau}\}$ determine the relative importance of news as a function of how far in advance of consumption the news is received. When $\gamma_{t,\tau} = 0$ for all $\tau > t$, changed expectations of future consumption do not affect a person's current well-being. By assuming that it is only the contrast between contemporaneous consumption and prior expectations that generates sensations of gain and loss, previous beliefs-based theories of reference-dependent utility in Stone (2005) and Kőszegi and Rabin (2006, 2007) have implicitly made exactly this parameter restriction, and as such have ignored considerations that are central in this paper. When, by contrast, $\gamma_{t,\tau} = 1$ for all $\tau \geq t$, getting news about an outcome before it happens resonates with a person just as much as learning it at the time. The in-between case we consider, where $0 < \gamma_{t,\tau} < 1$ for $\tau > t$ and $\gamma_{t,\tau}$ increasing in t, is that the impact of news about a period is greater the closer is the period. Although we find this last case most plausible, the role of temporal distance in the impact of learning news is (as far as we know) largely unexplored empirically. Our results identify ways that this affects behavior, so they can be used to guide empirical investigation in the area.[8]

Our formulation follows some previous research suggesting that changes in beliefs are carriers of utility. George F. Loewenstein (1988) finds that people are willing to pay more to avoid delaying the delivery of a durable good when they expected to receive it immediately than to speed up its delivery when they expected to receive it later. This is consistent with our assumption that updating beliefs about consumption utility in the interim period affects utility. Matthey (2005) discusses a wide range of factors that determine whether people might experience loss aversion over changes in different types of risk, including future risks. Most explicitly related to our formal specification of prospective gain-loss utility and Matthey's (2006) "reference-dependent utility from expectations," Hsee and Tsai (2008) find that the "news utility" from learning about near-term consumption can be stronger than the utility from consumption itself, and the model by Kimball and Willis (2006) assumes that news about future utility is one component of current happiness. Going beyond previous research saying that "news" is *a* source of reference-dependent utility, the premise of our model is that news is the *central* source. But while markedly different in formal terms, our model is in fact consistent with most previous theories of reference-dependent utility. Because past consumption is often a reliable indicator of future consumption, in many environments our model makes similar predictions to theories of habit formation and reference-dependent preferences starting from Ryder and Heal (1973). As we show below, the use of money in most classical applications of prospect theory can be understood in terms of people treating money as news about future consumption. Similarly, concern for changes in ownership of a durable good in models of the endowment effect is naturally interpretable as concern for the implied news about future consumption. In this light, the wealth and ownership-based models of reference dependence can be seen as simply not harping on details of how outcomes generate gain-loss utility rather than as reflecting a hypothesis that it really is the receipt of money or objects per se that people care about. Our formal elaboration that *does* harp on such details helps provide some unifying foundations for these approaches. More importantly, we find cases where the more explicit focus on beliefs about consumption modifies and extends insights.

We now turn to specifying the gain-loss utility function N. While our definition is notationally quite cumbersome, the basic idea is that the decision maker makes an "ordered comparison" between her previous beliefs $F^k_{t-1,\tau}(\cdot)$ and new beliefs $F^k_{t,\tau}(\cdot)$: she compares the worst percentile of outcomes under $F^k_{t,\tau}(\cdot)$ to the worst percentile of outcomes under $F^k_{t-1,\tau}(\cdot)$, the second-worst

[8] With similarly little evidence to go on, we conjecture that hyperbolic decay in $\gamma_{t,\tau}$ as $\tau - t$ increases is more plausible than exponential decay: bad news about tomorrow may be considerably less painful than bad news about today, while bad news about 117 days from now may resonate much the same as bad news about 116 days from now.

percentile of outcomes under $F^k_{t,\tau}(\cdot)$ to the second-worst percentile of outcomes under $F^k_{t-1,\tau}(\cdot)$, and so on, and experiences sensations of gain or loss for each of these comparisons. Suppose, for example, that she had believed she has a 50-50 chance of spending either 0 or 100 minutes with Johnny Depp. If she now believes that she has a 50-50 chance of spending either 20 or 120 minutes in Johnny's company, she experiences a gain of 20 minutes. If she now believes that she has a 50-50 chance of having either 20 or 80 minutes with Johnny, she experiences "mixed feelings" of a gain from comparing 20 minutes to 0 and a loss from comparing 80 minutes to 100. If she now believes she has a 60-40 chance at either 0 or 100 minutes, she experiences a loss equal to a 10 percent chance of losing 100 minutes. And if her beliefs were correct and uncertainty is resolved in period t, with probability ½ she compares a deterministic 0 minutes to the fifty-fifty 0/100 lottery—which feels like a one-half chance of losing 100 minutes with Johnny—and with probability ½ she compares a deterministic 100 minutes to the lottery—which feels like a one-half chance of gaining 100 minutes with Johnny.

The ordered comparison above is a substantive assumption about what drives gain-loss utility, and at least two alternative assumptions would also be relatively natural. One alternative is that the decision maker compares the means of her new and old beliefs. Except for a caveat to the results of Section III, this would make little difference for findings in the current paper.[9] Under the other alternative formulation, a decision maker compares each possible outcome under her new beliefs to each possible outcome under her old beliefs, rather than just the outcome at the same percentile. This would imply that she experiences gain-loss utility even if she does not change her beliefs, which seems unrealistic.

Formally, for any distribution F over $\mathbb{R}$ and any $p \in (0,1)$, let $c_F(p)$ be the consumption level at percentile p, defined implicitly by the conditions that $F(c_F(p)) \geq p$ and $F(c) < p$ for all $c < c_F(p)$. We define gain-loss utility from the change in beliefs in dimension k as

$$N^k(F^k_{t,\tau} \mid F^k_{t-1,\tau}) = \int_0^1 \mu\Big(m^k(c_{F^k_{t,\tau}}(p)) - m^k(c_{F^k_{t-1,\tau}}(p))\Big)\,dp,$$

where $\mu(\cdot)$ is a "universal gain-loss utility function" with the following properties:

(A0) $\mu(x)$ is continuous for all x, twice differentiable for $x \neq 0$, and $\mu(0) = 0$.

(A1) $\mu(x)$ is strictly increasing.

(A2) If $y > x \geq 0$, then $\mu(y) + \mu(-y) < \mu(x) + \mu(-x)$.

(A3) $\mu''(x) \leq 0$ for $x > 0$ and $\mu''(x) \geq 0$ for $x < 0$.

(A4) $[\mu'_-(0)]/[\mu'_+(0)] \equiv \lambda > 1$, where $\mu'_+(0) \equiv \lim_{x\to 0}\mu'(|x|)$ and $\mu'_-(0) \equiv \lim_{x\to 0}\mu'(-|x|)$.

Properties (A0)–(A4), stated by Bowman, Minehart, and Rabin (1999), correspond to Kahneman and Tversky's (1979) explicit or implicit assumptions about their "value function" defined on the difference between an outcome and the reference point. Loss aversion is captured by (A2) for large stakes and (A4) for small stakes, and diminishing sensitivity is captured by (A3). While the

[9] The means-based formulation would not affect the key conclusion in Section III, that a person often behaves as if she narrowly brackets isolated monetary risk, but it would affect the form of implied risk preferences we derive in Web Appendix B.

inequalities in (A3) are most realistically considered strict, to characterize the implications of loss aversion without diminishing sensitivity as a force on behavior, we define a subcase of (A3):

(A3′) For all $x \neq 0, \mu''(x) = 0$.

When we apply (A3′) below, we will parameterize μ as $\mu'_+(0) = \eta$ and $\mu'_-(0) = \lambda\eta > \eta$, so that η can be interpreted as the weight attached to gain-loss utility, and λ as the coefficient of loss aversion.[10]

We assume that total gain-loss utility in period t is simply the sum of gain-loss utilities in each dimension: $N(F_{t,\tau} | F_{t-1,\tau}) = \sum_{k=1}^{K} N^k(F^k_{t,\tau} | F^k_{t-1,\tau})$. Finally, in period t the person wishes to maximize the expectation of the sum of instantaneous utilities,[11]

$$U^t \equiv \sum_{\tau=t}^{T} u_\tau. \tag{2}$$

For notational simplicity as well as substantive reasons, our formulation of total utility assumes no discounting. In a classical model, the discount factor can be seen as representing uncertainty-based "heuristic discounting," where the decision maker puts lower weight on a future date because something might render her modeled decisions for that date irrelevant. The same shortcut is inappropriate in our model because optimal planning for the contingency of interest could depend on what happens in other contingencies, so that the uncertainty must be modeled directly. Although unlikely to be of any calibrational interest, to the extent that there is time-consistent "hedonic discounting"—whereby a person simply cares less about how happy she is in the future—this will not substantively affect most of our results.[12]

Our formulation above of the person's utility as a function of consumption and beliefs can be combined with any theory of how those beliefs are formed. In this paper, we ignore any possible errors of belief formation, and as a conceptually useful starting point base our theory on consistency with rationality: the person correctly anticipates the implications of her plans, and cannot make plans she knows she will not carry through. Irrespective of initial beliefs at the beginning of period 0, such rationality implies that the beliefs the decision maker ends each period with, F_0 to F_T, must stochastically match behavior and outcomes in periods 1 to T. In Web Appendix A, we discuss sundry solution concepts that follow from this, which correspond to various assumptions about the timing, hedonic consequences, and correctness of the decision maker's initial beliefs about the choice situation being modeled. In the text, however, we apply a stronger solution concept, *preferred personal equilibrium (PPE),* under which the decision maker chooses the plan for state-contingent behavior that in each period is consistent with her future behavior, and that maximizes expected reference-dependent utility going forward. While in some natural types of circumstances we discuss in Appendix A, PPE is likely to be the appropriate prediction for a rational person with the preferences modeled above, in many other circumstances it is likely to be too strong, and Appendix A provides a guide for how to make predictions in such cases.

[10] Our model follows other models of reference-dependent utility in making the clearly correct assumption that $\lambda > 1$, but it is worth noting that our theory would generate some interesting results even with $\lambda = 1$. In the model of Section IV, for instance, the decision maker overconsumes even when $\lambda = 1$.

[11] Welfare statements below, including examples of suboptimal behavior, are made with respect to these time-consistent preferences.

[12] For some of our results assuming that $\gamma_{t,\tau} < 1$, this is true only so long as the person's discount factor is not too small relative to the rate with which $\gamma_{t,\tau}$ decreases as the outcome moves further into the future. It is more likely that incorporating a time-inconsistent taste for immediate gratification would substantially change some of our results, but we have not explored how.

Our theory, then, requires that $F_0, \ldots, F_T$ be determined by a plan that in each period maximizes expected reference-dependent utility given the expectations generated by the plan, with the constraint that future behavior and plans must be similarly optimal. To formalize this while still suppressing a fair amount of cumbersome notation, we denote by d_t a state-contingent strategy for behavior starting in period t, or "plan" for short. In a consumption-savings model such as that considered in Section IV, for example, d_t could be the sequence of decisions regarding how much to consume in each period starting in period t, as a function of the available information about wealth at the time.[13] Given the environment, any plan induces some expectations F_{t-1} over future outcomes. For instance, a consumption plan defining consumption as a function of wealth, combined with a distribution of wealth levels, induces beliefs about the distribution of consumption in future periods. Let D_t be a set of feasible plans beginning in period t, which can depend on both prior choices and exogenous stochastic events. In a consumption-savings problem, this would be the set of feasible state-contingent consumption plans. This of course can depend on past decisions and uncertainty because these can affect current wealth: if the decision maker consumed more in the past or lost money in the stock market, some otherwise feasible consumption paths will now be unavailable. We define a preferred personal equilibrium as follows:

DEFINITION 1: *Define the sets* $\{D_t^*\}_{t=0}^{T}$ *in the following backward-recursive way. A plan* $d_t \in D_t$ *is in* D_t^* *if, given the expectations generated by* d_t*, in any contingency, (i) it prescribes a continuation plan in* D_{t+1}^* *that maximizes the expectation of* U^t*; and (ii) it prescribes an action in period t that maximizes the expectation of* U^t*, assuming that future plans are made according to (i). A plan* $d_1 \in D_1$ *is a preferred personal equilibrium (PPE) if* $d_1 \in D_1^*$ *and it maximizes the expectation of* U^1 *among plans in* D_1^*.

To solve for PPE, the decision maker has to think backward. First, she must figure out, for each possible choice set D_T she might face in the last period, which plans in D_T are credible. Any plan $d_T \in D_T$ induces beliefs F_{T-1} regarding consumption outcomes in period T. To be consistent with rationality, d_T must then satisfy the constraint that, given the expectations generated by d_T, self T (the period-T incarnation of the individual) will be willing to follow d_T. Knowing the set of credible plans D_T^* for each D_T, self $T-1$ can evaluate any action and any realization of uncertainty by making the best credible continuation plan for each such contingency. Note that there may be multiple credible plans in D_T^*; that the person can choose the best one amounts to assuming that she can form any credible expectations at will—and change future preferences as a result. Then, we can define the set of credible plans D_{T-1}^* for any D_{T-1} analogously to the above, and continue similarly to period 1. In period 0, the decision maker chooses her favorite plan among those that are credible—that is, her favorite plan in D_1^*.

II. Information Preferences

In this section, we explore some significant implications of the most basic premise of our model—loss-averse preferences over changes in beliefs—for how a person feels about information regarding fixed but unknown future consumption. As a complement to models—such as Caplin and Leahy (2001, 2004), Caplin and Kfir Eliaz (2003), and Kőszegi (2006a)—that assume directly a taste or distaste for information, we derive such tastes from the same preferences toward good and bad news that (as we interpret it) underlie prospect theory, and emphasize how a person's

[13] Our notation and definition suppresses that (i) the beliefs F_t depend on the available information and are generated by the entire state-contingent plan; (ii) actions can depend on the entire history, including past beliefs; and (iii) expected utility is taken over the entire sequence of outcomes and beliefs.

like or dislike of information may depend on features of the information and the environment. Our results mostly identify and elaborate on two key principles: that people prefer to get information clumped together rather than apart, and that—if prospective gain-loss utility weakens with time lag to the outcome—people prefer to get information sooner rather than later.[14]

To isolate the implications of loss aversion, we assume that $\mu(\cdot)$ satisfies (A3′), with $\mu(x) = \eta x$ for $x \geq 0$ and $\mu(x) = \eta\lambda x$ for $x < 0$. Because the general statements below are notationally somewhat cumbersome, we illustrate the intuition for many of our results in a simple example. Suppose $T = 2$, there is no consumption in period 1 and one dimension of consumption in period 2, and $m(c_2) = c_2$. There are two equiprobable possible consumption levels, $c_2 = 0$ and $c_2 = 1$, and the decision maker has no control over this outcome. She may, however, receive information about c_2 in period 1. Specifically, she may observe a signal $s \in \{0, 1\}$, where the signal is accurate $(s = c_2)$ with probability $q > 1/2$. We investigate how observing the early signal affects the decision maker's expected utility as a function of q and the strength of her concern for prospective gain-loss utility, $\gamma \equiv \gamma_{1,2}$.

If the person observes the signal, her expected gain-loss utility is[15]

$$-\frac{1}{2}\gamma\eta(\lambda - 1)\left(q - \frac{1}{2}\right) - q(1 - q)\,\eta(\lambda - 1). \tag{3}$$

The first term captures expected prospective gain-loss utility in period 1. After observing the early signal, the decision maker will either believe the high outcome $c_2 = 1$ happens with probability q—leading to a gain of $q - 1/2$ as compared to her prior—or she will believe it happens with probability $1 - q$—leading to a loss of $q - 1/2$ as compared to her prior. Because the loss is more heavily felt, expected prospective gain-loss utility is negative. The second term in expression (3) captures expected gain-loss utility in period 2. With probability 1/2, the person leaves period 1 believing consumption will be high with probability q. In that case, with probability q she later learns that $c_2 = 1$—leading to a gain of $1 - q$—and with probability $1 - q$ she learns that $c_2 = 0$—leading to a loss of q. The expected utility from this possibility is therefore $-q(1 - q)\eta(\lambda - 1)/2$. Similar considerations apply if she leaves period 1 assigning probability $1 - q$ to the high outcome.

If the decision maker does not observe the signal, she experiences gain-loss utility only in period 2, which in expectation is $-\eta(\lambda - 1)/4$. Hence, observing the signal generates strictly more expected utility than not observing it if and only if

$$\gamma < 2\left(q - \frac{1}{2}\right). \tag{4}$$

Suppose first that $q = 1$, so that the signal provides the same information the decision maker would otherwise learn in period 2. By Inequality (4), if $\gamma < 1$, the person strictly prefers to

[14] Closely related to the intuition for our decision maker's dislike of piecemeal information, Thaler (1999) argues that loss aversion in combination with narrow bracketing explains why a person would take multiple plays but not a single play of a bet—when she does not have to watch how the multiple bets play out. As we explain in the next section, we reinterpret narrow bracketing as a preference over fluctuations in beliefs. Without elaborating a general model, Ignacio Palacios-Huerta (1999) develops an example based on a natural extension of Gul's (1991) model of disappointment aversion that also features a preference for clumped information. And although he does not explicitly formulate a model of utility from changing beliefs, David Dillenberger (2008) also explores a set of related intuitions. In a class of recursive preferences over compound lotteries in which the decision maker does not care when uncertainty is fully resolved, Dillenberger shows that the preference for one-shot resolution of uncertainty—a special case of what we call the preference for clumped information—is closely related to the static concept of the certainty effect by Kahneman and Tversky (1979). Similar to our findings, Dillenberger points out that some forms of narrow bracketing can be interpreted in terms of preferences rather than as a mistake.

[15] Since consumption utility is independent of the arrival of information, we focus on gain-loss utility.

receive the information early, and if $\gamma = 1$, she is indifferent. Intuitively, given that there is an equal chance of information moving beliefs up and down, loss aversion implies that the decision maker finds it unpleasant in expected terms to learn information. If $\gamma < 1$, the sense of loss for non-immediate outcomes is not as large, so the person is better off receiving the information early. If $\gamma = 1$, the sense of loss is exactly as aversive in period 1 as in period 2, so the person is indifferent to the timing of information.

In addition to liking early information, the decision maker dislikes partial information. For the key intuition, suppose that $\gamma = 1$, so that the decision maker does not care when she experiences a given change in beliefs. Nevertheless, for any $q < 1$ she prefers not to receive the signal. Intuitively, piecemeal information exposes her to fluctuations in her beliefs: there is a possibility, for instance, that she receives a positive signal and her hopes rise, but then she is all the more disappointed by getting the worse of the two outcomes. Since the pleasure from receiving positive news is smaller than the pain from finding out that the news was incorrect, such fluctuations in beliefs decrease utility.

Although we are unaware of any empirical or experimental research asking directly whether people dislike piecemeal information, this prediction of our model is indirectly supported by experimental evidence on myopic loss aversion documenting that people pay less for assets that provide more interim information. In Gneezy and Potters (1997), subjects sequentially decided how much of \$2 to bet on a lottery with a two-thirds chance of losing the amount bet and a one-third chance of multiplying the amount bet by 2.5. In the "high-frequency" condition, subjects made a new decision and learned that period's outcome in every period. In the "low-frequency" condition, subjects made a decision every three periods for the next three periods, and learned only the aggregate outcome. Consistent with our theory, subjects bet much less in the high-frequency than in the low-frequency condition, and do so already starting in the first period. In this experiment, however, subjects in the high-frequency condition both received more information and could change their decisions more frequently. Charles Bellemare et al. (2005) disentangled these two effects by including a condition where subjects received information every period but could change their decisions only every three periods, and found that Gneezy and Potter's results are largely due to feedback frequency rather than investment flexibility. Michael S. Haigh and John A. List (2005) repeated essentially the same experiment as Gneezy and Potters (1997) with undergraduates as well as professional traders, and found a greater difference between conditions in the latter group than in the former group.[16]

When $\gamma < 1$ and $q < 1$, the decision maker faces a trade-off between her taste for early information and her distaste for partial information. For "weak" news, the latter effect dominates: for any $\gamma > 0$, the decision maker dislikes sufficiently weak early information. Intuitively, a very weak signal causes a small but first-order immediate change in beliefs, which loss aversion makes aversive in expected terms. While this early information decreases the expected change in beliefs in the future, it does so only by a second-order effect: it moves beliefs a small amount toward what eventual beliefs will be with slightly higher probability than it moves beliefs a small amount away from what eventual beliefs will be.

We now generalize these intuitions. Suppose that consumption occurs solely in period T, but that a person is potentially receiving information about this consumption in periods 1 through $T - 1$. Consider a sequence σ of signals about consumption, $s_1, s_2, \ldots, s_J$, and let $t(s_j \mid \sigma)$ denote the time that signal s_j is received under σ, with $t(s_j) \leq t(s_{j+1})$ for all j. We assume that each signal s_j is nontrivial in that it provides extra information relative to previous signals: there is a realization of

[16] Thaler et al. (1997) also obtain similar results, but in their experiment subjects were not told the distribution of returns from the different possible bets, so the difference between the high-frequency and low-frequency conditions arises only once subjects learn the riskiness of investments through experience.

$s_1, \ldots, s_{j-1}$ such that some realization of s_j changes the decision maker's beliefs. This specification fits many information-acquisition scenarios, and for the first two results we do not even impose that signals are conditionally independent. Also note that in applying the propositions below, it is useful that "two" signals arriving in the same period can be labeled as either one signal or two. We assume, however, that consumption in period T is binary and the signals are discrete, and as above the two possible outcomes are $c_T = 0$ and $c_T = 1$. We conjecture that our results in essence extend to more general settings, but we have not found an appropriate way to characterize and prove them.

To be able to compare information structures, we say that an information structure σ' is (t_a, t_b, j)–equivalent to σ if (i) σ and σ' involve the same sequence of signals, (ii) in both σ and σ' exactly the two signals s_j and s_{j+1} arrive between periods t_a and $t_b > t_a$ (inclusive), and (iii) for all $i \neq j, j+1$, $t(s_i \mid \sigma') = t(s_i \mid \sigma)$. That is, σ' and σ differ solely in the timing of the two signals s_j and s_{j+1}. Under a given set of preferences, let $U(\sigma)$ and $U(\sigma')$ be the discounted expected utilities for the two information structures.

Proposition 1 is the key result generalizing our insight above that piecemeal information is utility-decreasing. It says that collapsing two signals into one, so long as that does not delay the signals, always strictly improves welfare:

PROPOSITION 1: *Suppose that σ' is (t_a, t_b, j)-equivalent to σ with $t(s_{j+1} \mid \sigma') = t(s_j \mid \sigma') \leq t(s_j \mid \sigma) < t(s_{j+1} \mid \sigma)$. Then $U(\sigma') > U(\sigma)$ for any $\gamma_{t,T} > 0$ nondecreasing in t.*

By iteratively applying this proposition, it is clear that any change in information structure that collapses different signals without delaying any of them will raise utility for the person.

Our second proposition generalizes the point that receiving information earlier increases welfare if $\gamma < 1$, and does not affect welfare if $\gamma = 1$:

PROPOSITION 2: *Suppose that σ' is (t_a, t_b, j)-equivalent to σ with $t(s_j \mid \sigma') < t(s_j \mid \sigma)$, $t(s_{j+1} \mid \sigma') \leq t(s_{j+1} \mid \sigma)$, and $t(s_{j+1} \mid \sigma') = t(s_j \mid \sigma')$ if and only if $t(s_{j+1} \mid \sigma) = t(s_j \mid \sigma)$. Then, $U(\sigma') > U(\sigma)$ when $\gamma_{t,T}$ is strictly increasing in t and $U(\sigma') = U(\sigma)$ when $\gamma_{t,T} = 1$ for all t.*

This proposition says that, as long as it does not change the order of signals (including which signals arrive simultaneously), getting signals earlier rather than later is better if prospective gain-loss utility weakens with time to the outcome, and is equally good if prospective gain-loss utility does not weaken with time to the outcome. Combined with the previous proposition, the latter point implies that when prospective gain-loss utility is always as strong as contemporaneous, all the person cares about is to avoid the "dribbling in" of information—collapsing signals, even when this involves delaying the signals, is always a good thing. The two propositions also imply that learning everything right away is always at least as good as any other information structure.

For the final two propositions in this section, we assume that the signals s_1 through s_J are independent conditional on c_T. To state the next proposition, we call a signal *always informative* if for any realization s of the signal, $\Pr[s \mid c_T = 1] \neq \Pr[s \mid c_T = 0]$. The proposition establishes that, no matter how weak is prospective gain-loss utility, a sufficiently weak signal always harms welfare by a little. Although the proposition could be (more clumsily) stated for any small signal, for simplicity we consider symmetric binary signals of the form in our two-period example above.

PROPOSITION 3: *Choose any $\gamma_{t,T} > 0$, j, t_a, t_b, $\{s_i\}_{i \neq j}$, and (t_a, t_b, j)-equivalent σ and σ' with $t(s_j \mid \sigma) < t(s_{j+1} \mid \sigma) = t(s_j \mid \sigma') = t(s_{j+1} \mid \sigma')$. Suppose s_j is a binary signal with accuracy $1/2 + \epsilon$ and s_{j+1} is always informative. Then, if ϵ is sufficiently small, $U(\sigma') > U(\sigma)$.*

While Proposition 3 implies a dislike of *isolated* inaccurate information, the same logic that underlies the proposition in fact predicts a *preference* for weak information when a person has just been or is about to be exposed to other information. To state this result, say that a signal is of size $\epsilon > 0$ if the maximum change in the probability assigned to the high outcome resulting from the signal is ϵ. Proposition 4 says that a person always strictly prefers information of a size smaller than her recent change in beliefs, even if $\gamma = 1$, the information is very inaccurate, and receiving it means splitting it away from a later signal:

PROPOSITION 4: *Let σ and σ' be $(t_a, t_b, 1)$-equivalent with $t(s_1 \mid \sigma') = t_a < t(s_1 \mid \sigma)$ and $t(s_2 \mid \sigma') = t(s_2 \mid \sigma)$. Suppose the decision maker has received information in period t_a that changed her subjective probability of getting $c_T = 1$ from p to p'. If s_1 is of size less than $|p - p'|$, $U(\sigma') > U(\sigma)$.*

For an intuition, suppose, say, that the person has just received information that has decreased her beliefs by a given amount. Any further information that is of smaller size will not move her beliefs back to their original level. Hence, since any news will change only the degree of loss she suffers, but not whether she suffers a loss or a gain, further positive news is evaluated just as strongly as further negative news. Therefore, this news generates zero immediate utility in expectation. Yet this news decreases expected future fluctuations in beliefs, increasing expected utility.

An implication of the logic of Proposition 3 is that if a decision maker receives information in the form of a sequence of small news, and she experiences the full force of gain-loss utility for the changes in beliefs associated with each item of news, then in the limit as the number of such episodes becomes arbitrarily large, her utility converges to negative infinity. While our basic prediction that piecemeal information lowers utility seems sensible and is supported by evidence cited above, this extreme prediction contradicts some potentially common behaviors: many investors follow the performance of their stock portfolios on a day-to-day basis, and many sports fans follow the online ticker during a game rather than looking only at the final score. But this extreme version of our prediction is, for reasons both within and outside our model, unrealistic. Most important, to capture that the reference point is lagged expectations, we have assumed for convenience that a belief change *within* a period does not generate gain-loss utility, but a belief change *between* periods does. Exactly because the reference point does not immediately adjust to new expectations, taking this simplified lag structure to vanishingly short time periods is calibrationally unrealistic. In such a setting, it is reasonable to make the reference point a weighted average of past beliefs, significantly decreasing aversion to small news. In addition—although this is not captured by our theory—it is unlikely that individuals would or could pay attention to every small piece of information. And in the few cases that they do, this seems partly to be due to motivations, such as a curiosity-induced inability to avoid available information or the pure enjoyment of following a sports game, our model ignores. Finally, Proposition 4 clarifies that when a decision maker is already receiving information, she may like rather than dislike small news. Our theory predicts, for example, that if an investor can avoid all information regarding her retirement wealth, she prefers not to monitor her prospects too closely. But if she receives some unavoidable information, she will immediately look for additional information herself. Even if the first of these predictions turns out to be inaccurate because of alternative motivations not included in our model, the comparative-static prediction—that the more information a person receives in a given period, the more willing she is to receive additional information—is likely to be robust to other motives.

A simple extension of our results in this section can be used to formalize and extend a key point by Kimball and Willis (2006). They argue that since news about consumption affects immediate happiness, changes in happiness following news can be used to infer a person's consumption utility. In our model, not only happiness, but information-seeking *behavior*, can be used to identify consumption utility. While we have for simplicity set $m(0) = 0$ and $m(1) = 1$, in each of the instances above it is clear that the decision maker's like or dislike of information is proportional to $m(1) - m(0)$. Hence, how much the person is willing to pay to receive or avoid information about a particular outcome reveals how much she cares about that outcome, even if she never makes any choice that affects the probability of the outcome. This means, in principle, that after identifying the general nature of a person's prospective and contemporaneous gain-loss utilities, we may be able to use revealed preference over information to identify her consumption utility for outcomes over which she has no control.

III. Monetary Preferences

Many previous models of reference-dependent utility assume—presumably as a shortcut—that individuals care directly about receiving money or experiencing changes in wealth. Although we cannot analyze a completely general model and we clearly do not engage all issues related to the complex psychology of money, in this section we show how our model might provide consumption-based foundations for gain-loss utility over money, and use these foundations to derive some new predictions about monetary preferences. To motivate the issues, we begin with a puzzle regarding the psychology of money that has been noted by researchers, and was elaborated most clearly by Nicholas Barberis, Ming Huang, and Thaler (2006) and Kőszegi and Rabin (2007). Because the typical person holds wealth primarily for future consumption and faces substantial uncertainty regarding future wealth—so that modest changes in current wealth are unlikely to determine whether she ends up above or below her reference point in the future—it would seem that even for a loss-averse person, the optimal strategy for gambles over modest stakes is close to expected-value maximization. Hence, a nonneutral attitude toward small risks seems to require "narrow bracketing"—ignoring that the current risk will be integrated with substantial other risk—in a way that has been interpreted as an error by Kahneman and Dan Lovallo (1993), Shlomo Benartzi and Thaler (1995), Daniel Read, Loewenstein and Rabin (1999), and others. This behavior is, in fact, especially puzzling when viewed as a cognitive error: by maximizing a value function meeting assumptions (A0) to (A4) in a piecemeal way rather than simply maximizing expected value, people are exhibiting a complicated suboptimal pattern of reference-dependent behavior rather than a simple near-optimal one.

By defining gain-loss utility over changes in beliefs, our approach provides a new perspective on these issues. We predict that people may care about small changes in wealth even if they recognize that the changes contribute negligible risk to the consumption ultimately determined by their wealth. The reason is simple: gains and losses in money are news about future consumption, and this news generates immediate prospective gain-loss utility. But beyond this basic insight, our specification predicts how the timing of news about a risk affects a person's attitude toward that risk. This means that our model extends some of the arguments in Kőszegi and Rabin (2007) in providing a unifying framework for determining whether and when various existing reduced-form models of monetary preferences apply. The model also predicts when rationally accounting for background risk would in fact eliminate any significant aversion to moderate amounts of additional risk, rendering narrow bracketing a mistake.

To demonstrate that a person may exhibit reference-dependent preferences with respect to gains and losses in money even if she will not use the money for immediate consumption, we develop our formal results for situations where all consumption occurs in the future. For simplicity, we also assume that there is a single period of future consumption. This captures in a reduced form a setting where the decision maker does not think through when in the future she will absorb current changes in wealth.[17] Suppose $T = 2$, $K = 1$, nontrivial consumption occurs only in period 2, and $m(c_2) = c_2$. The decision maker's consumption in period 2 is the sum of two components: "background risk" with a fixed and known distribution, and the outcome resulting from her choice from the set D of independent small risks to period-2 consumption. The decision maker might have known since period 0 that she would make a choice from D, or—in what can be thought of as a "surprise" situation—she might have believed she would be facing only the background risk, and then find in period 1 or 2 that she must choose from D.[18] In addition, the decision maker may have to implement her choice in period 0, 1, or 2, and the uncertainty in the lotteries in D might be realized in period 1 or 2.[19] Appendix B characterizes behavior formally as a function of all these features of the environment. Here we discuss intuitively some key features of the results.

A stark case illustrating some implications of the model is when the person is confronted with D, implements her decision, and learns the outcome all in period 1. Then, between periods 0 and 1 the distribution of future consumption shifts by exactly the realized outcome of the chosen lottery, generating a gain or loss equal to that realized outcome. This means that *independently of the background risk*, the decision maker chooses from the set D as postulated by prospect theory, maximizing reference-dependent utility from money receipts with a reference point of zero. In this sense, our model says that some forms of "narrow bracketing" are not necessarily errors. Whatever the background risk, losing $10 conveys the bad news that one will have less to consume in one's lifetime, and in our theory this fully and rationally generates a sense of loss.

Although the exact form of behavior is more complicated, the decision maker also attends to the gains and losses resulting from her choice if she learns D in period 0, implements her decision in period 1, and uncertainty in D is also resolved in period 1. Then, both a deviation of her own behavior from expected, and the resolution of uncertainty in period 1, generate news that shifts the distribution of period-2 consumption and induces immediate gain-loss utility. Hence, when choosing from D in period 1, the person attends to gains and losses relative to her expectations regarding her choice. This means that she behaves according to a "static" version of PPE (defined originally in Kőszegi and Rabin (2006) and also in Appendix B) *as applied to D*: she makes the best plan she knows she will carry through. The complication is that the way she evaluates outcomes relative to the reference lottery depends on the background risk. For some forms of background risk, the decision maker evaluates each outcome relative to all possibilities in the reference lottery, as proposed by Kőszegi and Rabin (2006, 2007). But for other forms of background risk, she evaluates an outcome relative to the mean of the reference lottery, as in the disappointment-aversion models of Graham Loomes and Robert Sugden (1982) and David E. Bell (1985). In either case, she is first-order risk averse with respect to the options in D.

Similar considerations arise if the decision maker learns D and implements her decision in period 0, and uncertainty is resolved in period 1. In this case, she anticipates that the resolution of uncertainty will generate utility-inducing news in period 1, so whether or not there is background risk she chooses from D according to something like the CPE concept defined in Kőszegi and Rabin (2007) and in Appendix B: she maximizes expected reference-dependent

[17] We speculate at the end of the section how endogenous choices of the type we explore in Section IV, where the person allocates wealth between immediate and future consumption, would affect loss aversion over money. But because the path of consumption in dynamic settings is quite complicated, we do not know to what extent our results extend to a more general model.

[18] This latter environment is the limiting case of situations where the decision maker had expected to face only the background risk with probability $1 - \epsilon$ and also to choose from D with probability ϵ, and the low-probability event is realized. Since the reference point is then largely determined by the background risk, unless there is indifference the person makes the same choice from D as when D is a complete surprise.

[19] Technically, our model introduced in Section I does not allow for a decision to be made in period 0. The more general setting in Appendix A incorporates, among other things, this simple extension.

utility taking into account that both the reference lottery and the outcome lottery are determined by her choice from D. As above, however, how she evaluates an outcome relative to the reference lottery depends on the nature of the background risk.

In all these situations, the uncertainty in D is resolved in period 1, and in various ways the decision maker cares about gains and losses in money. The common thread is clear: although the decision maker fully understands the fungibility of money and that realizations of wealth in period 1 have no immediate consumption implications, since she cares about the changes in beliefs induced, she attends to those realizations. Notice that this applies not only to current risk, but also to future risk: even in period 0, the decision maker attends to gains and losses in period 1—because she anticipates they will generate utility-inducing news in period 1.

In contrast, we show in Appendix B that our theory typically predicts approximate risk neutrality when uncertainty in D is resolved in period 2 and is integrated with large background uncertainty.[20] Intuitively, our model says that it is only when the decision maker learns the outcomes of a risk separately that she treats gains and losses asymmetrically—as only such outcomes generate asymmetric effects on gain-loss utility. When the risk in question is drowned in other risk that is resolved contemporaneously, any additional news from it will likely just affect the *degree* of good or bad news—not whether the overall news is good or bad—so the person will be approximately neutral to the risk. Once again, this logic applies whether or not the person makes her decision from D in the same period as when the risk is resolved (period 2) or earlier (period 0 or 1). Therefore, while the comparative-statics prediction that a person will be more averse to more isolated risk seems to us broadly right, our theory fails to explain important instances of narrow bracketing that have been observed. For instance, Tversky and Kahneman (1981) and Rabin and Georg Weizsäcker (forthcoming) found that subjects narrowly bracket two separate pairwise choices even when they know the two chosen lotteries would be played out and reported at the same time. But while inconsistent with the fully rational framework we develop, these mistakes seem more interpretable when viewed through the lens of our model than with either classical reference-independent preferences or other reference-dependent theories. Since we learn about the consequences of many or most of our decisions separately, and "narrow bracketing" in these situations is *not* a mistake in our model, mistaken narrow bracketing may naturally result when people heuristically treat even simultaneous decisions as if they would be informed about the outcomes in isolation.

Due to the many cases and possibilities, our model may appear to be an overly complicated and unwieldy theory of monetary preferences. We feel, however, that the complexity reflects part of the complicated psychology of money. In fact, in almost all cases we consider in the Appendix, our theory reduces to a previous theory of monetary preferences—which researchers presumably introduced because they believed it was realistic in the contexts where they were applying it. Furthermore, our theory not only says that all these previous theories may be right in some circumstances, it predicts in *which* circumstances each of them is appropriate.

By dint of providing consumption-based foundations for loss aversion in money, our model also establishes an endogenous relationship between loss aversion over goods and loss aversion over money. If $\gamma_{t,\tau} < 1$ for $\tau > t$, a person will be less loss averse over money that will be used for future consumption than over immediately consumed goods, as suggested for example by Nathan Novemsky and Kahneman (2005). This contrasts with our single-decision model in Kőszegi and Rabin (2006), where we have assumed equal loss aversion across dimensions, including money. While underlying gain-loss utility in our dynamic setting is still a single gain-loss utility

[20] As we explain in Appendix B, the only exception can happen when the decision maker learns D in period 1. In that case, the news of learning D can itself generate departures from risk neutrality.

function $\mu(\cdot)$ applying to all dimensions, the dynamic aspect introduces differences in loss aversion according to differences in the timing of consumption.

Incorporating an endogenous choice regarding the timing of consumption can further weaken loss aversion over money, and may also affect other results in this section. As we show in the next section, when $\gamma < 1$ a person may respond to increases in wealth by increasing consumption immediately, but to decreases in wealth by reducing *future* consumption. Because $\gamma < 1$, this decreases her sensitivity to losses but not gains, weakening and in some cases eliminating (but never reversing) loss aversion. We briefly return to this issue below.

IV. Wealth and Consumption in Intertemporal Choice

In this section we explore the pattern of consumption in simple two-period consumption-savings decisions with and without uncertainty regarding wealth. We elaborate on a new form of overconsumption relative to the optimal committed plan that was first identified by Stone (2005), and demonstrate that a stochastic plan for period-1 consumption, and hence ex ante uncertainty in wealth that is resolved in period 1, exacerbates overconsumption. Our theory predicts an asymmetric response to surprises regarding wealth. Finally, we provide a novel and intuitive explanation for precautionary savings, while also noting that whether a person increases or decreases savings in response to future uncertainty typically depends on when she learns about the uncertainty.

Suppose a consumer needs to decide how to allocate consumption spending between periods 1 and 2, given an intertemporal budget constraint $c_1 + c_2 = W$. Her consumption utility $m(\cdot)$ is strictly increasing and strictly concave.[21] As in Section II, for notational simplicity we let $\gamma \equiv \gamma_{1,2}$.

Our first and biggest goal is to explore, in settings with and without uncertainty, whether the decision maker chooses the consumption path that maximizes her ex ante utility among the strategies available to her.[22] Suppose first that W is deterministic. Then, choosing $c_1 = c_2 = W/2$ maximizes both ex ante consumption utility and ex ante expected gain-loss utility (which is zero for deterministic plans and negative for nondeterministic plans), and so is the ex ante optimal committed strategy. But this strategy may not be consistent. If the consumer had planned $c_1 = W/2$, then her period-1 utility for $c_1 \geq W/2$ is

$$m(c_1) + \eta(m(c_1) - m(W/2)) - \gamma\eta\lambda(m(W/2) - m(W - c_1)) + m(W - c_1). \tag{5}$$

The first and last terms constitute consumption utility in periods 1 and 2, respectively. The second term is the period-1 contemporaneous gain from consuming above plans in that period, and the third term is the period-1 prospective loss from having to plan lower future consumption as a result. Whatever the person consumes in period 1, she forms new expectations that determine her reference point in period 2, so there is no gain-loss utility in that period.

[21] With reference-independent utility, it is generally appropriate to use an indirect utility function defined over spending as a reduced form for the solution to a decision maker's full optimization over multiple goods. The same is true in our model (under Assumption $(A3')$) in deterministic settings, but not, for instance, when there is uncertainty about prices: because a person's expectations can affect her preferences, a single indirect utility function may not capture her preferences across all different situations. We have little sense for the general implications or the calibrational significance of such examples, however. To derive results in the simplest possible setting, we use a single-dimensional utility function.

[22] As we have mentioned above, we do so under the assumption that preferences are time consistent. We do not believe that our main results, especially regarding the role of uncertainty in overconsumption, would be qualitatively different if a source of overconsumption was a time-inconsistent taste for immediate gratification.

The derivative of expression (5) with respect to c_1 evaluated at $W/2$ is

$$(1+\eta)m'(W/2) - (1+\gamma\eta\lambda)m'(W/2) = \eta(1-\gamma\lambda)m'(W/2). \tag{6}$$

Hence, if $\gamma < 1/\lambda$—if the consumer cares much more about contemporaneous gain-loss utility than about prospective gain-loss utility—deviating from the ex ante optimal plan increases utility ex post, so that this plan is not consistent, and any consistent plan must have $c_1 > W/2$. This pattern is behaviorally very similar to recent models of hyperbolic discounting and present bias (Laibson 1997; O'Donoghue and Rabin 1999) and temptation disutility (Gul and Pesendorfer 2001) in that consumption is higher than the ex ante optimal committed level, so that self 0 values commitment, and would commit to a lower consumption level than the one implemented in PPE.[23] But in our model behavior is suboptimal for a completely different reason than in these previous theories. Namely, whereas the decision maker takes the reference point as given in period 1, her awareness in period 0 that she would deviate raises her reference point for period 1, and consequently lowers her ex ante utility in that period. This beliefs-based time inconsistency arises even though preferences in our model are time consistent.

That expectations-based loss aversion can generate overconsumption replicates in our setting the same point made by Stone (2005). Stone assumes that deviations of current consumption from previous expectations induce gain-loss sensations, but there are no gain-loss sensations from the implied deviation of future consumption from previous expectations. Translated into our model, this means that Stone (2005) implicitly assumed $\gamma = 0$. Our theory says that Stone's implicit assumption is not without loss of generality: the weight on prospective gain-loss utility is crucial in determining whether overconsumption occurs. If $\gamma \geq 1/\lambda$, the derivative in expression (6) is negative, so increasing c_1 above c_2 does not increase utility ex post. Intuitively, the prospective loss from lowering future consumption tends to act as an internal commitment device that discourages the decision maker from deviating from her plan by increasing c_1. Summarizing:

PROPOSITION 5: *Suppose that wealth W is deterministic. If $\gamma \geq 1/\lambda$, the PPE consumption path is $(W/2, W/2)$. If $\gamma < 1/\lambda$, the PPE consumption path (c_1^*, c_2^*) has $c_1^* > c_2^*$ and satisfies*

$$(1+\eta)m'(c_1^*) = (1+\gamma\eta\lambda)m'(c_2^*). \tag{7}$$

We now show that uncertainty that is resolved in period 1 exacerbates the tendency to overconsume. Suppose that wealth W is distributed according to the continuous distribution $F(\cdot)$, that the uncertainty regarding wealth is resolved in period 1, and that the consumer makes a PPE plan that in both periods calls for strictly increasing consumption as a function of total wealth; a sufficiently large amount of wealth uncertainty will force the consumer to make such plans. Let these consumption functions be $c_1(W)$ and $c_2(W)$, respectively. Suppose that the realized wealth level is W, and the consumer is considering whether to change consumption from the planned level of $c_1(W)$. Since the probability that c_1 was going to be lower than $c_1(W)$ is $F(W)$, an increase in c_1 is evaluated as a gain in proportion to $F(W)$, and as a decreased loss in proportion to $1 - F(W)$. Hence, the marginal utility from an increase in c_1 is $m'(c_1(W))[1 + F(W)\eta + (1 - F(W))\eta\lambda]$. The same deviation also lowers c_2, lowering prospective gain-loss utility in period 1. As with period-1 consumption, the probability that c_2 was going to be lower than $c_2(W)$ is $F(W)$, so that a decrease

[23] In fact, if $\lambda = 1$, our model is, for both deterministic and stochastic wealth, observationally equivalent to the β-δ model presented in Laibson (1997), with $\beta = (1+\gamma\eta)/(1+\eta)$ and $\delta = 1$. As we show below, however, when $\lambda > 1$ our model predicts that uncertainty can exacerbate overconsumption, while previous models do not make the same prediction.

in c_2 is evaluated as a foregone prospective gain in proportion to $F(W)$, and as a prospective loss in proportion to $1 - F(W)$. Because the consumer learns her wealth in period 1, there is no gain-loss utility in period 2. Hence, the total marginal disutility from a decrease in c_2 is $m'(c_2(W))[1 + F(W)\gamma\eta + (1 - F(W))\gamma\eta\lambda]$. In order for $c_1(W)$ to be consistent, therefore, it must be that

$$(8)\ 0 = m'(c_1(W))[1 + F(W)\eta + (1 - F(W))\eta\lambda] - m'(c_2(W))[1 + F(W)\gamma\eta + (1 - F(W))\gamma\eta\lambda].$$

Equation (8) implies that if $\gamma < 1$, then $c_1(W) > c_2(W)$ for all W. Intuitively, the consumer cares more about surprises regarding c_1 than about surprises regarding c_2, so she is willing to surprise herself with an increase in consumption utility in period 1 if the cost is a similar or smaller decrease in consumption utility in period 2. Hence, no continuous plan with $c_1(W) \leq c_2(W)$ can be consistent. Although the ex ante optimal committed plan is itself not generally to set $c_1(W) = c_2(W)$, this logic also leads $c_1(W)$ to be higher than optimal.

PROPOSITION 6: *Suppose wealth is distributed continuously, the uncertainty regarding wealth is resolved in period* 1, *and PPE plans* $c_1^*(W), c_2^*(W)$ *are strictly increasing in W. If* $\gamma < 1$, *then* $c_1^*(W) > c_2^*(W)$, *and decreasing* $c_1^*(W)$ *in any neighborhood on the support of W would strictly increase ex ante expected utility.*

The analysis above implies a major difference between deterministic plans and stochastic plans. For $1/\lambda \leq \gamma < 1$, in fact, there is a striking result: if a wealth distribution induces the consumer to make stochastic plans, she overconsumes for *all* wealth realizations, but if she knew her wealth level in advance, she would not overconsume for *any* wealth realization. A prescribed consumption path that has high probability (or is deterministic) acts as a bright line that discourages overconsumption: the sense of gain from increasing consumption in period 1 is smaller than the sense of loss from the corresponding decrease in consumption in period 2. But when a prescribed period-2 consumption level has low (or zero) probability, the possibility of higher or lower consumption is already incorporated into period-2 expectations, so the two sensations from deviating are evaluated much more symmetrically. This diminishes the commitment power of the plan.

Note that if the consumer learns her wealth in period 2 instead of period 1, uncertainty does not similarly undermine self-control. In that case, a plan involves a deterministic consumption level in period 1, so that if the consumer deviates by increasing c_1, she experiences that as a contemporaneous gain combined with a prospective loss due to a downward shift in the distribution of period-2 consumption. In fact, as we show below, uncertainty that is resolved in period 2 leads the consumer to *decrease* consumption in period 1.

The next result shows that when $\gamma < 1$, the consumer responds asymmetrically to surprises about wealth, as documented by John Shea (1995) for unionized employees. To illustrate this most simply, suppose that she had been expecting to have lifetime wealth W with probability one, but at the beginning of period 1 learns (to her surprise) that her wealth is actually $W + y$. This is the limiting case of environments where she assigns a small probability to having wealth $W + y$.

PROPOSITION 7: *Suppose that the consumer had expected a deterministic wealth level W, made PPE plans to consume* c_1^*, c_2^* *in the two periods, and in period* 1 *learns that her wealth is* $W + y$. *If* $\gamma < 1$, *there are constants* $\underline{y} < 0$ *and* $\overline{y} > 0$ *with the following properties:*

(*i*) *If* $0 \leq y \leq \overline{y}$, *the resulting consumption pattern is* $c_1^* + y, c_2^*$. If $y > \overline{y}$, *consumption satisfies* $c_1 > c_1^*, c_2 > c_2^*$ *and is given by*

$$(1 + \eta)u'(c_1) = (1 + \gamma\eta)u'(c_2).$$

(*ii*) *If* $0 \geq y \geq \underline{y}$, *the resulting consumption pattern is* $c_1^*, c_2^* + y$. *If* $y < \underline{y}$, *consumption satisfies* $c_1 < c_1^*, c_2 < c_2^*$ *and is given by*

$$(1 + \eta\lambda)u'(c_1) = (1 + \gamma\eta\lambda)u'(c_2).$$

Proposition 7 says that (i) the person consumes sufficiently small windfalls entirely in period 1, leaving period-2 consumption unchanged; and (ii) she does not cut immediate consumption in response to small bad surprises, absorbing the entire shock in period 2. Larger increases or decreases in wealth are split between the two periods. Intuitively, $\gamma < 1$ implies that in period 1, consuming a windfall gain immediately is more pleasant than planning for higher future consumption, and unexpectedly lowering consumption immediately is more painful than planning for lower future consumption. Hence, the consumer immediately takes advantage of surprise improvements in circumstances, so as to be able to enjoy the pleasant surprise most; and she puts off absorbing negative news, so as to feel the unpleasant surprise least.

In combination with Proposition 5, Proposition 7 has an interesting implication for the relationship between overconsumption and the response to small surprise gambles. Proposition 5 says that with deterministic wealth, overconsumption occurs when $\gamma < 1/\lambda$, and that in this case, the decision maker chooses the lowest credible early consumption—setting it so that in period 1 a marginal contemporaneous gain is exactly as pleasant as a marginal prospective loss is painful. But given Proposition 7's result that the person absorbs small surprise gains immediately but small surprise losses only in the future, she will be neutral to small surprise gambles. In this sense, there is a negative relationship between overconsumption and loss aversion: the person is loss averse over small monetary risks if and only if she chooses the ex ante optimal consumption path. Intuitively, if a person is loss averse in a situation with deterministic plans, loss aversion gives her the power to implement a slightly different plan—which means that her plan cannot be suboptimal.

The negative relationship between overconsumption and loss aversion, however, is somewhat special for a number of reasons. Whenever the person cannot reasonably consume gains from a surprise gamble immediately, she will again be loss averse. Similarly, if she is liquidity constrained to the point where losing a gamble forces lower immediate consumption, she will display loss aversion. Finally, although we do not show this in the current paper, when there is uncertainty regarding wealth that is realized in period 1, a decision maker with $\gamma < 1$ often chooses a deterministic c_1, so that for most realizations of wealth she is loss averse even if she absorbs losses in period 2.[24]

Finally, we analyze how a person responds to future uncertainty. Suppose that y is a mean-zero nondeterministic lottery resolved in period 2, $s > 0$ is a scalar, and wealth is equal to $W_0 + s \times y$. We establish properties of the PPE level of savings as a function of the scale s of the risk. Although our formal results are more general, the main point can be illustrated easily by assuming that y is a binary lottery that gives 1 or -1 with probability ½ each. This means that wealth takes on two possible values, $W_0 + s$ and $W_0 - s$, with equal probabilities. To abstract from issues of overconsumption, for the illustration suppose that γ is sufficiently high for the optimal consumption path to be a PPE. Let $c_2 = W_0 + s - c_1$ and $c_2' = W_0 - s - c_1$ be the two

[24] There are also reasons going far beyond our model that weaken the negative relationship between overconsumption and loss aversion. When evaluating a small gamble, many people might not predict whether and how they will time the consumption of gains and losses, and assume heuristically that they will absorb the implications of the gamble at the same time. Also, as mentioned in Section III, they may simply narrowly bracket in ways outside our model.

possible consumption levels in period 2 as a function of c_1. Then, the consumer's expected utility in period 2 is

$$\frac{1}{2}m(c_2) + \frac{1}{2}m(c_2') - \frac{1}{4}\eta(\lambda - 1)[m(c_2) - m(c_2')]. \tag{9}$$

The first two terms represent expected consumption utility, while the last term represents expected contemporaneous gain-loss utility in period 2. Consuming c_2', which has a one-half chance of occurring, feels like a loss relative to c_2, which the consumer also expected to occur with probability ½. The expected disutility from this loss is $\eta\lambda(m(c_2) - m(c_2'))/4$. Similarly, c_2 feels like a gain relative to the possibility of getting c_2', generating utility $\eta(m(c_2) - m(c_2'))/4$. Because of loss aversion, the overall impact of uncertainty on gain-loss utility is negative. Since the consumer learns no new information in period 1, her prospective gain-loss utility in that period is zero.

Using expression (9), the optimal consumption path satisfies

$$m'(c_1) = \frac{1}{2}m'(c_2) + \frac{1}{2}\mathrm{m}'(c_2') + \frac{1}{4}\eta(\lambda - 1)[m'(c_2') - m'(c_2)]. \tag{10}$$

Equation (10) extends the standard Euler equation to loss-averse preferences, and collapses to the standard Euler equation for either $\eta = 0$ or $\lambda = 1$. The nonstandard term is the last term on the right-hand side, which reflects a loss-aversion-based incentive to increase savings. Intuitively, future risk exposes the decision maker to a sensation of loss should realized consumption utility be lower than other possible realizations. In order to decrease the pain from this loss, the consumer saves more to decrease the impact of any given wealth shock on consumption utility.

The reason above for precautionary savings is potentially more intuitive than the classical reference-independent one: whereas the classical account predicts that increased consumption meliorates the disutility of uncertainty only insofar as higher spending makes consumption utility more linear, our model predicts the melioration happens simply because more spending lowers the marginal utility of consumption. For the same reason, our model of precautionary savings may also be more robust than the standard one. To make this case formally, consider the Taylor-expansion approximation of the right-hand side of equation (10) around $s = 0$:

$$m'(c_2) + \frac{1}{2}m'''(c_2)s^2 + \frac{1}{2}\eta(\lambda - 1)(-m''(c_2))s.$$

When there is no loss aversion ($\eta = 0$ or $\lambda = 1$), uncertainty has a second-order effect on savings, and whether this effect is positive or negative depends on the third derivative of the utility function. With loss aversion, the effect is first-order and unambiguously positive for any strictly concave consumption-utility function. This feature generalizes to other situations:[25, 26]

[25] Technically speaking, the unambiguous prediction that uncertainty increases savings is true only for small amounts of risk. For large amounts of risk, the consumer's consumption utility can dominate gain-loss utility, and in that case—as in the standard model—savings in general depends on the third derivative of $m(\cdot)$. Since $m(\cdot)$ represents a global utility function, however, a small risk in our model can still be very large in practical terms.

[26] While our framework predicts a novel reason for precautionary savings, it also implies that the decision maker's response to future uncertainty depends on whether she had known about the uncertainty before the savings decision. Suppose she enters period 1 having made plans anticipating a deterministic wealth level W, but then learns that wealth is uncertain and will be resolved in period 2. Because building precautionary savings would require her to decrease period-1 consumption below the expected level, for sufficiently low amounts of uncertainty she does not decrease consumption. In fact, in some situations she increases consumption. For example, suppose that $2/(1 + \lambda) > \gamma \geq 1/\lambda$—so that the consumer plans to set $c_1 = c_2 = (W/2)$—and she learns that her wealth has a 50-50 chance of being either $W + \epsilon$ or $W - \epsilon$. For a sufficiently small ϵ, she responds by increasing consumption by ϵ, fully consuming the better realization of the uncertainty. To see this, note that the consumer's utility from the perspective of period 1 when $(W/2) < c_1 < (W/2) + \epsilon$ is:

PROPOSITION 8: *Suppose wealth is equal to* $W_0 + s \times y$, *where* y *is a nondeterministic mean-zero lottery that is resolved in period* 2. *For any strictly concave* $m(\cdot)$ *and any* $\eta > 0$, $\lambda > 1$, $\gamma \geq 0$, *the PPE consumption rule satisfies* $dc_1/ds\,|_{s=0} < 0$.

The predictions of our model differ in several ways from prominent existing dynamic specifications of reference-dependent utility and habit formation, such as Ryder and Heal (1973), Gary S. Becker and Kevin M. Murphy (1988), and John Y. Campbell and John H. Cochrane (1999), where the reference point for evaluating a consumption outcome is lagged consumption. The most striking difference is that, in our setting, increasing consumption today reduces tomorrow's reference point by reducing expectations of future consumption, whereas in lagged-consumption-based models the same raises tomorrow's reference point by raising the habitual level of consumption. Such habit formation is realistic and could be added to consumption utility in our model, but we have not done so. As a result, our model (unlike these others) does not predict the kind of ex ante preference for increasing consumption profiles that was found in surveys by Loewenstein and Nachum Sicherman (1991) and Loewenstein and Drazen Prelec (1993).[27] At the same time, our model generates many realistic predictions that lagged-consumption-based theories do not. For instance, because in these other theories a person typically plans for a strictly increasing consumption profile, her consumption is not at the reference point in most periods, so that she does not respond to news regarding wealth in the asymmetric way that our model predicts. In addition, none of these models predicts overconsumption, a role of prior uncertainty on behavior, or a preference over the timing of decision-irrelevant information.

V. Conclusion

In addition to unifying existing theories and intuitions about reference-dependent utility and making new predictions in the settings explored in Sections III–V, we hope the model developed can be applied in many other economic settings. Because our model predicts that deterministic plans are easier to stick to than stochastic plans, it may help provide some insights into the relationship between self-control, budgeting, and mental accounting. Intuitively, because deterministic plans accentuate the power of loss aversion to deter impatient behavior, a person might make rigid plans even in some uncertain environments where the rigidity would seem to be costly. Also, because our theory predicts that expectations to acquire a good are painful to give up, it may capture the common intuition that as an auction progresses, bidders who seem to have a chance of winning get excited and bid higher. And by emphasizing the importance of plans and beliefs that occur in a person's head before any observable decisions are made, our model may provide a foundation for the relationship between decisions and contemplation.

$$\begin{aligned}
&m(c_1) + \eta(m(c_1) - m(W/2)) + \frac{1}{2}m(W+\epsilon-c_1) + \frac{1}{2}m(W-\epsilon-c_1)\\
&\quad + \frac{1}{2}\gamma\eta(m(W+\epsilon-c_1) - m(W/2)) - \frac{1}{2}\gamma\eta\lambda(m(W/2) - m(W-\epsilon-c_1)) - \frac{1}{4}\eta(\lambda-1)(m(W+\epsilon-c_1)\\
&\quad - m(W-\epsilon-c_1)).
\end{aligned}$$

It is easy to show that for a sufficiently low ϵ, the derivative of this expression is positive. Intuitively, when wealth is deterministic, the sensation of loss that would result from lowering planned consumption for period 2 can keep a person from overconsuming. If she learns that her wealth may be higher, however, she evaluates a decrease in period-2 consumption partly as a foregone gain from the possible windfall. Hence, she does not find the decrease in future consumption so aversive.

[27] Note that our model (like other models of time-inconsistent behavior) calls into question whether such an ex ante expression of preference over a profile is what will manifest itself in a dynamic setting. Indeed, the "overconsumption" we might predict in a model that includes habit formation would presumably be relative to these ideal profiles.

But there are several important ways in which our model is both an incomplete and an inaccurate model of reference-dependent preferences. As discussed in Kőszegi and Rabin (2006, 2007), perhaps the greatest weakness of our theory—as well as other theories of reference-dependent utility put forward—is that it takes as a primitive the set of decisions and risks a person is focusing on. In this paper, for example, our theory predicts that a person is more likely to overconsume if she had been expecting a lot of uncertainty to be resolved just prior to her choice than if she had been expecting little. To the extent that it is often hard to know what people have been thinking before a particular decision, this kind of prediction may be hard to test.

Our theory also takes as a primitive input the set of dimensions on which gain-loss utility is evaluated separately. We discuss this issue in detail in the context of our static model in Kőszegi and Rabin (2004), but a dynamic setting introduces further complications. As an example of one unattractive feature, suppose a person revises her consumption in 101 periods down by one apple and her consumption in 100 periods up by one apple. In our formulation, the former is evaluated as a loss and the latter merely as a gain, generating negative gain-loss utility. It seems plausible, however, that the latter gain fully compensates for the former loss, so that the person experiences no gain-loss utility. That is, consumption 100 periods from now may be on the same psychological dimension as consumption 101 periods from now.

Finally, psychological evidence in the context of adaptation and other domains indicates that people often underestimate the extent to which changes in circumstances will change their preferences (for example, Kahneman 1991; Loewenstein, O'Donoghue, and Rabin 2003). In the context of expectations-based reference dependence, this could mean that a decision maker underestimates how much changes in expectations will change how she will feel in the future. Any results that are driven by a motive to manage future gain-loss utility by changing current expectations—e.g., the preference for early information in Section II—will presumably be weakened if such preference misprediction is incorporated into a model.

Appendix: Proofs

PROOF OF PROPOSITION 1:

The two sequences of signals generate the same expected utility up to s_{j-1} being received and after s_{j+1} is received, so we compare expected utilities for the two in-between signals. Notice that because the outcome is binary and μ is linear, for any updating of the probability of $c_T = 1$ from p_{t-1} to p_t generates gain-loss utility equal to $N_t(p_t \,|\, p_{t-1}) = \mu(p_t - p_{t-1})$.

Consider any realized signals $s_1, \ldots, s_{j-1}$. For simplicity, we introduce the following notation for the purposes of this proof and that of Proposition 2. Let p be the decision maker's posterior on $c_T = 1$ after $s_1, \ldots, s_{j-1}$. Let $p(s_j)$ and $p(s_j, s_{j+1})$ denote the updated probabilities after one or two additional signals. Finally, let $\gamma_1 = \gamma_{t(s_j|\sigma'),T} = \gamma_{t(s_{j+1}|\sigma'),T}, \gamma_2 = \gamma_{t(s_j|\sigma),T}, \gamma_3 = \gamma_{t(s_{j+1}|\sigma),T}$. Since $\gamma_{t,T}$ is nondecreasing in t, we have $\gamma_1 \leq \gamma_2 \leq \gamma_3$. Now expected utility with signal structure σ' is

$$\begin{aligned}\gamma_1 E_{s_j, s_{j+1}}[\mu(p(s_j, s_{j+1}) - p)] &= \gamma_1 E_{s_j, s_{j+1}}[\mu(p(s_j) - p + p(s_j, s_{j+1}) - p(s_j))] \\ &\geq \gamma_1 E_{s_j, s_{j+1}}[\mu(p(s_j) - p) + \mu(p(s_j, s_{j+1}) - p(s_j))],\end{aligned}$$

where the last inequality holds because μ is steeper for losses than for gains. Furthermore, because these signals are nontrivial, there are realizations of $s_1, \ldots, s_{j-1}$ such that the inequality above is strict. Now we can further rewrite the above as

$$\begin{gathered}\gamma_1 E_{s_j}[\mu(p(s_j) - p)] + E_{s_j}[\gamma_1 E_{s_{j+1}}[\mu(p(s_j, s_{j+1}) - p(s_j)) \,|\, s_j]] \geq \\ \gamma_2 E_{s_j}[\mu(p(s_j) - p)] + E_{s_j}[\gamma_3 E_{s_{j+1}}[\mu(p(s_j, s_{j+1}) - p(s_j)) \,|\, s_j]],\end{gathered}$$

since all these expectations are negative and $\gamma_1 \leq \gamma_2 \leq \gamma_3$. But the right-hand side above is exactly the decision maker's expected utility under σ.

PROOF OF PROPOSITION 2:

We use a similar method of proof to that of Proposition 1. We prove for the case $t(s_{j+1} \mid \sigma') > t(s_j \mid \sigma')$ and $t(s_{j+1} \mid \sigma) > t(s_j \mid \sigma)$. The other case, when $t(s_{j+1} \mid \sigma') = t(s_j \mid \sigma')$ and $t(s_{j+1} \mid \sigma) = t(s_j \mid \sigma)$, is similar. Let $\gamma_1 = \gamma_{t(s_j \mid \sigma),T}, \gamma_2 = \gamma_{t(s_{j+1} \mid \sigma),T}$ and $\gamma'_1 = \gamma_{t(s_j \mid \sigma'),T}, \gamma'_2 = \gamma_{t(s_{j+1} \mid \sigma'),T}$. Then, using the notation of Proposition 1, the expected gain-loss utility from these two signals with information structure σ' is

$$\gamma'_1 E_{s_j}[\mu(p(s_j) - p)] + E_{s_j}[\gamma'_2 E_{s_{j+1}}[\mu(p(s_j, s_{j+1}) - p(s_j)) \mid s_j]],$$

whereas with information structure σ it is

$$\gamma_1 E_{s_j}[\mu(p(s_j) - p)] + E_{s_j}[\gamma_2 E_{s_{j+1}}[\mu(p(s_j, s_{j+1}) - p(s_j)) \mid s_j]].$$

Since all these expectations are nonpositive, and they are strictly negative for some realizations of $s_1, \ldots, s_{j-1}$, the proposition immediately follows.

PROOF OF PROPOSITION 3:

Suppose that given the signals she has observed so far, the decision maker's posterior of $c_T = 1$ is p. Consider a signal realization $\tilde{s}$. For this proposition and the following one, we make use of the following lemma.

LEMMA 1: *The amount by which $\tilde{s}$ moves the decision maker's subjective probability of $c_T = 1$ is differentiable and strictly concave in p.*

PROOF OF LEMMA 1:

For notational simplicity, let $q = \Pr[\tilde{s} \mid c_T = 1]$ and $q' = \Pr[\tilde{s} \mid c_T = 0]$. By Bayes's rule, upon observing $\tilde{s}$, the decision maker updates her beliefs that $c_T = 1$ to

$$\frac{pq}{pq + (1-p)q'},$$

so the change in her beliefs is

$$\frac{pq}{pq + (1-p)q'} - p = (q - q')\,\frac{1}{\dfrac{q}{1-p} + \dfrac{q'}{p}}.$$

The denominator in the final expression is clearly positive, differentiable, and strictly convex in p. The reciprocal of such a function is always differentiable and strictly concave. This completes the proof of the lemma.

To prove the proposition, we consider any beliefs $p \in (0, 1)$ the decision maker holds after signals $s_1, \ldots, s_{j-1}$. By assumption, the probability that the decision maker ends up with such beliefs after $j - 1$ signals is positive. Similarly to the proof of Proposition 1, let $p(s_j)$, $p(s_{j+1})$, $p(s_j, s_{j+1})$ denote updated probabilities after the given signals, and let $\gamma_1 = \gamma_{t(s_j \mid \sigma),T}$, $\gamma_2 = \gamma_{t(s_{j+1} \mid \sigma),T}$. Notice that since s_{j+1} is always informative and has only finitely many possible realizations, for a sufficiently small ϵ no realization of s_j turns good news from s_{j+1} into bad news, or vice versa. This means that by virtue of the law of iterated expectations, for a sufficiently small ϵ we have $E_{s_j,s_{j+1}}[|p(s_j,s_{j+1}) - p|] = E_{s_{j+1}}[|p(s_{j+1}) - p|]$. Hence, under σ' the expected utility from signals s_j and s_{j+1} is $-1/2\ \gamma_2\eta(\lambda - 1)E_{s_{j+1}}[|p(s_{j+1}) - p|]$ where this expression reflects the utility of the expected change in beliefs, which is negative since losses are felt heavier than gains. The expected utility under σ is instead

$$-\frac{\gamma_1}{2}\eta(\lambda - 1)E_{s_j}[|p(s_j) - p|] - \frac{\gamma_2}{2}\eta(\lambda - 1)E_{s_j}[E_{s_{j+1}}[|p(s_j,s_{j+1}) - p(s_j)||s_j]].$$

By Lemma 1, the difference between $E_{s_j}[E_{s_{j+1}}[|p(s_j,s_{j+1}) - p(s_j)||s_j]]$ and $E_{s_{j+1}}[|p(s_{j+1}) - p|]$ is second-order in ϵ. But $E_{s_j}[|p(s_j) - p|]$ is first-order in ϵ, so for a sufficiently low ϵ the information structure σ yields strictly lower expected utility.

PROOF OF PROPOSITION 4:

Notice that under σ', the signal s_1 has zero expected utility impact in period t_a because its size is less than $|p' - p|$. Hence, the result is immediate if $t(s_1|\sigma) < t(s_2|\sigma)$ since in this case the utility impact of s_1 is negative under σ and the impact of s_2 is the same under σ and σ'. Now suppose $t(s_1|\sigma) = t(s_2|\sigma)$, and let $\gamma = \gamma_{t(s_1|\sigma),T} = \gamma_{t(s_2|\sigma),T} = \gamma_{t(s_2|\sigma'),T}$. Using the same notation as in the previous propositions, notice that, by virtue of the law of iterated expectation and the convexity of the absolute value function,

$$E_{s_1,s_2}[|p(s_1,s_2) - p|] \geq E_{s_2}[|p(s_2) - p|].$$

Notice that $-\gamma\eta(\lambda - 1)/2$ times the term on the left-hand side of the inequality above is the decision maker's expected utility from receiving the two signals under σ. By Lemma 1, the right-hand side of the inequality above is strictly greater than

$$E_{s_1}[E_{s_2}[|p(s_1,s_2) - p(s_1)||s_1]],$$

which is $-\gamma\eta(\lambda - 1)/2$ times the decision maker's expected utility under σ'. This completes the proof.

PROOF OF PROPOSITION 5:

Consider $\gamma \geq 1/\lambda$ first. Given the strict concavity of consumption utility, the ex ante optimal plan is $c_1 = c_2 = W/2$. We show that this plan is consistent. We have established in the text that self 1 does not want to deviate by locally increasing period-1 consumption, and it is easy to show that she does not want to deviate by locally decreasing consumption. Furthermore, since self 1's utility function is concave, this means that self 1 does not want to deviate from a plan to consume equally.

We use an extension of the same argument to derive the PPE for $\gamma < 1/\lambda$. In this case, self 0 makes plans to follow the smoothest consumption path from which self 1 will not deviate. That is, self 0 plans the lowest consumption level in period 1 that is consistent with period-1 behavior. We show that this is given by the equation (7). Again, since self 1's utility function is concave, it is sufficient to consider local deviations. Furthermore, for plans with $c_1 \geq c_2$, self 1 would clearly not deviate by decreasing period-1 consumption, so we consider only local increases in period-1 consumption. With plans to consume the amounts (c_1, c_2) in the two periods, a period-1 change of plans to slightly increase immediate consumption induces contemporaneous gain-loss utility of $\eta m'(c_1)$ and prospective gain-loss utility of $-\gamma\eta\lambda m'(c_2)$. For any $c_1 < c_1^*$, the net utility is positive, but it is zero for $c_1 = c_1^*$.

PROOF OF PROPOSITION 6:

The proof for the claim that $c_1^*(W) > c_2^*(W)$ is in the text. The marginal effect of changing $c_1^*(W)$ in a neighborhood of W on ex ante expected utility is

$$(11)\qquad f(W)m'(c_1^*(W))\,[\underbrace{1 + F(W)\eta + (1 - F(W))\eta\lambda}_{\text{direct effect}} - \underbrace{(1 - F(W))\eta - F(W)\eta\lambda}_{\text{indirect effect}}]$$
$$-f(W)m'(c_2^*(W))\,[\overbrace{1 + F(W)\gamma\eta + (1 - F(W))\gamma\eta\lambda} - \overbrace{(1 - F(W))\gamma\eta - F(W)\gamma\eta\lambda}].$$

Increasing $c_1(W)$ above $c_1^*(W)$ has two effects on ex ante expected utility. It has a *direct* effect through changing utility—both consumption utility and gain-loss utility—when realized wealth is W. This effect corresponds exactly to the right-hand of equation (8). But changing c_1 also affects ex ante utility *indirectly* through changing gain-loss utility when realized wealth is not W: it increases losses for lower wealth realizations and decreases gains for higher wealth realizations. Similar considerations hold for the ex ante effect of decreasing $c_2(W)$ on prospective-gain loss utility. Expression (11) can be rewritten as

$$(12)\quad f(W)m'(c_1(W))[1 + (1 - 2F(W))\eta(\lambda - 1)] - f(W)m'(c_2(W))[1 + (1 - 2F(W))\gamma\eta(\lambda - 1)].$$

Because

$$\frac{1 + F(W)\eta + (1 - F(W))\eta\lambda}{1 + F(W)\gamma\eta + (1 - F(W))\gamma\eta\lambda} > \frac{1 + (1 - 2F(W))\eta(\lambda - 1)}{1 + (1 - 2F(W))\gamma\eta(\lambda - 1)}$$

for any $\gamma < 1$, whenever equation (8) holds, expression (12) is negative.

PROOF OF PROPOSITION 7:

By Proposition 5, when $\gamma \geq 1/\lambda$, then $m'(c_1^*) = m'(c_2^*)$, and when $\gamma < 1/\lambda$, then $(1 + \eta)m'(c_1^*) = (1 + \gamma\eta\lambda)m'(c_2^*)$. Using that $\gamma < 1$ and $\lambda > 1$, this implies that in either case $(1 + \eta)m'(c_1^*) > (1 + \gamma\eta)m'(c_2^*)$. This immediately implies the statements regarding wealth increases. By similar considerations, $(1 + \eta\lambda)m'(c_1^*) > (1 + \gamma\eta\lambda)m'(c_2^*)$, implying the statements regarding wealth decreases.

PROOF OF PROPOSITION 8:

We prove that the derivative of the marginal utility of increasing saving with respect to s is positive. This implies that $dc_1/ds\,|_{s=0} < 0$ both when $\gamma > 1/\lambda$ (because the ex ante optimal plan, which the person follows, features a lower c_1) and when $\gamma \leq 1/\lambda$ (because a higher marginal utility in period 2 means that a lower c_1 becomes consistent).

Let F be the cumulative distribution function of the random variable y. Similarly to the argument in the text, the expected utility in period 2 is

$$\int m(c_2 + sy)\, dF(y) + \iint \mu(m(c_2 + sy) - m(c_2 + sy'))\, dF(y')\, dF(y)$$
$$= \int m(c_2 + sy)\, dF(y) - \frac{1}{2}\eta(\lambda - 1) \iint (m(c_2 + s\max\{y, y'\}) - m(c_2 + s\min\{y, y'\}))\, dF(y')\, dF(y).$$

The marginal utility from increasing savings is therefore

$$\int m'(c_2 + sy)\, dF(y) + \frac{1}{2}\eta(\lambda - 1) \iint (m'(c_2 + s\min\{y, y'\}) - m'(c_2 + s\max\{y, y'\}))\, dF(y')\, dF(y).$$

The derivative of the expression above with respect to s evaluated at $s = 0$ is

$$\frac{1}{2}\eta(\lambda - 1)(-m''(c_2)) \iint |y - y'|\, dF(y')\, dF(y).$$

This completes the proof.

REFERENCES

Akerlof, George A., and William T. Dickens. 1982. "The Economic Consequences of Cognitive Dissonance." *American Economic Review*, 72(3): 307–19.

Barberis, Nicholas, Ming Huang, and Richard H. Thaler. 2006. "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case of Narrow Framing." *American Economic Review*, 96(4): 1069–90.

Battigalli, Pierpaolo, and Martin Dufwenberg. 2009. "Dynamic Psychological Games." *Journal of Economics Theory*, 144(1): 1–35.

Becker, Gary S., and Kevin M. Murphy. 1988. "A Theory of Rational Addiction." *Journal of Political Economy*, 96(4): 675–700.

Bell, David E. 1985. "Disappointment in Decision Making Under Uncertainty." *Operations Research*, 33(1): 1–27.

Bellemare, Charles, Michaela Krause, Sabine Kroger, and Chendi Zhang. 2005. "Myopic Loss Aversion: Information Feedback vs. Investment Flexibility." *Economics Letters*, 87(3): 319–24.

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

Bowman, David, Deborah Minehart, and Matthew Rabin. 1999. "Loss Aversion in a Consumption-Savings Model." *Journal of Economic Behavior and Organization*, 38(2): 155–78.

Brunnermeier, Markus K., and Jonathan A. Parker. 2005. "Optimal Expectations." *American Economic Review*, 95(4): 1092–1118.

Campbell, John Y., and John H. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, 107(2): 205–51.

Caplin, Andrew, and Kfir Eliaz. 2003. "Aids Policy and Psychology: A Mechanism-Design Approach." *RAND Journal of Economics*, 34(4): 631–46.

Caplin, Andrew, and John Leahy. 2001. "Psychological Expected Utility Theory and Anticipatory Feelings." *Quarterly Journal of Economics*, 116(1): 55–79.

Caplin, Andrew, and John Leahy. 2004. "The Supply of Information by a Concerned Expert." *Economic Journal*, 114(497): 487–505.

Dillenberger, David. 2008. "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior." Unpublished.

Duesenberry, James S. 1952. *Income, Saving and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press.

Gneezy, Uri, and Jan Potters. 1997. "An Experiment on Risk Taking and Evaluation Periods." *Quarterly Journal of Economics*, 112(2): 631–45.

Gul, Faruk. 1991. "A Theory of Disappointment Aversion." *Econometrica*, 59(3): 667–86.

Gul, Faruk, and Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica*, 69(6): 1403–35.

Haigh, Michael S., and John A. List. 2005. "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis." *Journal of Finance*, 60(1): 523–34.

Hsee, Christopher K., and Claire I. Tsai. 2008. "Hedonomics in Consumer Behavior." In *Handbook of Consumer Psychology*, ed. Curtis P. Haugtvedt, Paul M. Herr, and Frank R. Kardes, 639–59. New York, NY: Psychology Press.

Kahneman, Daniel. 1991. "Judgement and Decision-Making: A Personal View." *Psychological Science*, 2(3): 142–45.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy*, 98(6): 1325–48.

Kahneman, Daniel, and Dan Lovallo. 1993. "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking." *Management Science*, 39(1): 17–31.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263–91.

Kimball, Miles, and Robert Willis. 2006. "Utility and Happiness." Unpublished.

Knetsch, Jack L. 1989. "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *American Economic Review*, 79(5): 1277–84.

Kőszegi, Botond. 2006a. "Ego Utility, Overconfidence, and Task Choice." *Journal of the European Economic Association*, 4(4): 673–707.

Kőszegi, Botond. 2006b. "Emotional Agency." *Quarterly Journal of Economics*, 121(1): 121–55.

Kőszegi, Botond, and Matthew Rabin. 2004. "A Model of Reference-Dependent Preferences." UC Berkeley Economics Department Working Paper E04–337.

Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133–65.

Kősszegi, Botond, and Matthew Rabin. 2007. "Reference-Dependent Risk Attitudes." *American Economic Review*, 97(4): 1047–73.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 112(2): 443–77.

Loewenstein, George F. 1988. "Frames of Mind in Intertemporal Choice." *Management Science*, 34(2): 200–14.

Loewenstein, George F., Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias in Predicting Future Utility." *Quarterly Journal of Economics*, 118(4): 1209–48.

Loewenstein, George F., and Drazen Prelec. 1993. "Preferences for Sequences of Outcomes." *Psychological Review,* 100(1): 91–108.

Loewenstein, George F., and Nachum Sicherman. 1991. "Do Workers Prefer Increasing Wage Profiles?" *Journal of Labor Economics*, 9(1): 67–84.

Loomes, Graham, and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty." *Economic Journal*, 92(368): 805–24.

Matthey, Astrid. 2005. "Getting Used to Risks: Reference Dependence and Risk Inclusion." Humboldt University Working Paper 2005–036.

Matthey, Astrid. 2006. "Yesterday's Expectation of Tomorrow Determines What You Do Today: The Role of Reference-Dependent Utility from Expectations." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=920997.

Novemsky, Nathan, and Daniel Kahneman. 2005. "The Boundaries of Loss Aversion." *Journal of Marketing Research*, 42(2): 119–28.

O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review*, 89(1): 103–24.

Palacios-Huerta, Ignacio. 1999. "The Aversion to the Sequential Resolution of Uncertainty." *Journal of Risk and Uncertainty*, 18(3): 249–69.

Post, Thierry, Martijn J. van den Assem, Guido Baltussen, and Richard H. Thaler. 2008. "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show." *American Economic Review*, 98(1): 38–71.

Rabin, Matthew, and Georg Weizsäcker. Forthcoming. "Narrow Bracketing and Dominated Choices." *American Economic Review.*

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

Ryder, Harl E., Jr., and Geoffrey M. Heal. 1973. "Optimum Growth with Intertemporally Dependent Preferences." *Review of Economic Studies*, 40(1): 1–33.

Scitovsky, Tibor. 1976. *The Joyless Economy: The Psychology of Human Satisfaction*. Oxford: Oxford University Press.

Shea, John. 1995. "Union Contracts and the Life-Cycle/Permanent-Income Hypothesis." *American Economic Review*, 85(1): 186–200.

Stone, Rebecca. 2005. "Loss Aversion and Self-Control." Unpublished.

Thaler, Richard H. 1999. "Mental Accounting Matters." *Journal of Behavioral Decision Making*, 12(3): 183–206.

Thaler, Richard H., and Eric J. Johnson. 1990. "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice." *Management Science*, 36(6): 643–60.

Thaler, Richard H., Amos Tversky, Daniel Kahneman, and Alan Schwartz. 1997. "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test." *Quarterly Journal of Economics*, 112(2): 647–61.

Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science*, 211(4481): 453–58.