Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 162 (2016) 305–351

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

# Goals and bracketing under mental accounting

Alexander K. Koch [a,b], Julia Nafziger [a,*,1]

[a] *Department of Economics and Business Economics, Aarhus University, Fuglesangs Allee 4, 8210 Aarhus V, Denmark*
[b] *IZA, Germany*

Received 3 September 2015; final version received 17 December 2015; accepted 2 January 2016

Available online 8 January 2016

**Abstract**

Behavioral economics struggles to explain why people sometimes evaluate outcomes separately (narrow bracketing of mental accounts) and sometimes jointly (broad bracketing). We develop a theory of endogenous bracketing, where people set goals to tackle self-control problems. Goals induce reference points that make substandard performance painful. Evaluating goals in a broadly bracketed mental account insulates an individual from exogenous risk of failure; but because decisions or risks in different tasks become substitutes there are incentives to deviate from goals that are absent under narrow bracketing. Extensions include goal revision, naïveté about self-control, income targeting, and firms' bundling strategies.


*JEL classification:* D03; D81; D91

*Keywords:* Quasi-hyperbolic discounting; Reference-dependent preferences; Self-control; Mental accounting; Choice bracketing; Goals

* Corresponding author.
*E-mail addresses:* akoch@econ.au.dk (A.K. Koch), jnafziger@econ.au.dk (J. Nafziger).

[1] We thank an associate editor, two anonymous referees, as well as Johannes Abeler, Kathrin Breuer, Stefano DellaVigna, Kfir Eliaz, Dirk Engelmann, Guido Friebel, Leonie Gerhards, Lorenz Götte, Paul Heidhues, Botond Kőszegi, Georg Nöldeke, Marco Ottaviani, Larry Samuelson, Wendelin Schnedler, Palle Sörensen, Heiner Schumacher, Norov Tumennasan, Yoram Weiss, participants at seminars at the Universities of Aarhus, Berlin (Humboldt), Bern, Copenhagen, Copenhagen Business School, East Anglia, Frankfurt, Heidelberg, Mainz, Munich, Stockholm, Zürich, ZEW Mannheim and Copenhagen Business School. Financial support from the Danish Council for Independent Research | Social Sciences (FSE) under grant 12-124835 is gratefully acknowledged.

http://dx.doi.org/10.1016/j.jet.2016.01.001

## 1. Introduction

We investigate how people set goals and evaluate these in mental accounts to achieve self-control, thereby making novel predictions about a central question in behavioral economics: mental accounting.[2] An important and still poorly understood aspect of mental accounting is how people "set the brackets". Do they evaluate outcomes in a narrowly or in a broadly bracketed mental account? Camerer et al. (1997) and Read et al. (1999) informally discuss that narrowly evaluated goals, such as daily work goals, may facilitate self-control. Heath and Soll (1996) document how people control their expenditures in narrowly bracketed mental accounts, such as entertainment, clothing, or food. And many diet programs, such as Weight Watchers PointsPlus™, involve daily nutrition goals. But at the same time, not all goals are narrow. People do not have a mental account for every item they buy, or for every possible consumption category. And diet programs typically combine daily nutrition goals with the recommendation to weigh yourself not daily, but only at weekly intervals, or to set a weekly exercise goal (e.g., UK National Health Service, 2012).

To explain such puzzling evidence, we explicitly model the processes through which mental accounts impose constraints on future behavior, and how these constraints depend on the bracket of the account. By asking what goals are self-enforcing under a certain type of bracket, we derive boundary conditions for self-regulation and a theory of endogenous bracketing. Extending our model to allow for revision of goals and brackets, we address the question why it is not optimal for an individual to deviate from an originally adopted bracket for mental accounts by transforming narrow brackets into a broad bracket.

The idea that narrow bracketing is a means of overcoming self-control problems figures prominently in behavioral economics and consumer research.[3] Yet, this literature has the important limitation that the brackets of a mental account are imposed rather than endogenous and that models directly assume non-fungibility between accounts without spelling out how the brackets of a mental account actually constrain behavior.[4]

We consider an individual who faces two decisions with uncertain productivity (for instance, how hard to work on two different tasks or how much to consume of two different goods) and who has a demand for self-control stemming from a present bias. The individual's present bias implies that, when making a decision, he puts too little weight on the future benefits, or the harm that working on the task brings along. To motivate his future self, the individual therefore sets goals and specifies the brackets of his mental accounts. Goals induce reference points in a particular mental account, and the individual is loss averse regarding goal achievement.[5]

---

[2] Mental accounting is often associated with how people organize their financial activities (cf. Thaler, 1999). We adopt the broader perspective of Tversky and Kahneman (1981).

[3] Examples are Thaler and Shefrin (1981), Shefrin and Thaler (1988), Thaler (1985, 1999), Heath and Soll (1996), Prelec and Loewenstein (1998), Read et al. (1999), and Fudenberg and Levine (2006, 2011).

[4] Prominent examples are Thaler and Shefrin (1981) and Shefrin and Thaler (1988), who assume a different marginal propensity to consume out of wealth for each account; and Fudenberg and Levine (2006, 2011), who model mental accounts as "pocket cash constraints" on a short-run self with a life-time of one period. Non-fungibility between accounts is imposed by assuming that the short-run self has no access to other accounts.

[5] Goals and mental accounts are internal commitment devices. Relatedly, Bénabou and Tirole (2004) explain why internal commitment devices can actually work if an individual has imperfect knowledge about his willpower. Our approach applies to different informational environments (perfect vs. imperfect self-knowledge) and relies on different mechanisms.

Modeling goals as reference points builds on the psychology research on goals, which Locke and Latham (2002, p. 709–710) summarize as follows: "goals serve as the inflection point or reference standard for satisfaction versus dissatisfaction [. . . ] For any given trial, exceeding the goal provides increasing satisfaction as the positive discrepancy grows, and not reaching the goal creates increasing dissatisfaction as the negative discrepancy grows." Heath et al. (1999) provide evidence for people evaluating goals in a value function as the one used in Prospect Theory (Kahneman and Tversky, 1979).

We use the definition of mental accounts of Tversky and Kahneman (1981, p. 456): "an outcome frame which specifies (i) the set of elementary outcomes that are evaluated jointly and the manner in which they are combined and (ii) a reference outcome that is considered neutral or normal." Studies in economics and consumer research suggest tight links between mental accounting and loss aversion (Tversky and Kahneman, 1981; Kahneman and Tversky, 1984; Thaler, 1985, 1999; Prelec and Loewenstein, 1998; Soman, 2004). Direct evidence for such links comes from two studies in particular. In a field experiment by Fehr and Götte (2007) only those workers who exhibit loss aversion show a labor supply pattern consistent with a mental account around a daily income goal. Crawford and Meng (2011) empirically explain taxi drivers' labor supply with expectation based loss aversion and mental accounts that are bracketed daily.

A goal is a plan, such as "I want to study 8 hours on Monday for the exam". Goals give rise to expectations about outcomes, such as "I will get a good grade". These expectations serve as reference points for future selves (Kőszegi and Rabin, 2006, 2007). If the outcome exceeds expectations under a given bracket, the individual feels a gain. If the outcome falls short of expectations, the individual feels a loss. In line with the above evidence, we assume that the individual is loss averse.

For goals to translate into expectations, the individual must believe that his goals can be accomplished.[6] To capture this idea, we build upon the concept of *personal equilibrium* (Kőszegi and Rabin, 2006) and assume that goals, along with the expectations they induce, are rational. If the individual has the goal to work 8 hours, it must indeed be optimal to do so given the induced reference points for related costs and benefits. This also rules out that the individual can make himself arbitrarily happy by setting arbitrarily low goals.[7]

There is a wide-spread view that narrow bracketing of decisions is an error (e.g., Kahneman et al., 1993; Benartzi and Thaler, 1995; Read et al., 1999; Rabin and Weizsäcker, 2009). In our context this is captured by the result (Proposition 1) that, all else equal (i.e., for *fixed* decision levels), it is optimal for a loss-averse individual to evaluate outcomes together under a broad bracket to pool risks (e.g., Thaler et al., 1997; Kőszegi and Rabin, 2007). The reason is that *risk*

[6] The psychology literature finds that goals must be "realistic" and "attainable" (e.g., Hollenbeck et al. 1989). Popular self-help guides stress that goals should be "SMART" – specific, measurable, attainable, realistic, and timely.

[7] We extend previous work on goals as reference points in single-task settings (Carrillo and Dewatripont, 2008; Suvorov and van de Ven, 2008; Koch and Nafziger, 2011b; Hsiaw, 2013) to a continuous action space and stochastic reference points (cf. Proposition 8 in appendix A.2). Using an equilibrium concept similar to ours, Hsiaw (2015) studies two sequential, continuous-time optimal stopping problems. The environment and intuitions however are different from ours. Goals help counter the tendency of a present-biased individual to stop projects too early and to forego the option value of waiting for a potentially higher payoff. A narrow goal is evaluated as soon as a project stops, whereas broad goals postpone the evaluation until the second project is stopped, which weakens incentives. Because gain–loss utility is discounted more strongly, a deviation from the goal in the first task is less painful under broad relative to narrow goals. This discounting effect plays no role for our results. The incentive effects of bracketing from our paper do not appear in her setting.

*pooling* reduces the probability of falling into the loss domain. These are the well-known benefits of broad bracketing.

Our contribution is to show that broad bracketing can have costs, because the bracket affects the incentives of the individual to stick to his goals. We establish that for an individual with sufficiently severe self-control problems, narrow bracketing is optimal if there is not too much uncertainty about the productivity of a decision (Proposition 2).[8] With narrow bracketing, the individual evaluates, say, the health benefits and costs of one particular meal, or he evaluates the study effort and grade for one particular exam. If he deviates from his goal, say he works less than his goal prescribes, he will face a loss from a lower than expected outcome. This fear of a loss helps the individual stick to his goal. In contrast, under broad bracketing the individual is partly insured against experiencing a loss. This insurance weakens incentives. A shortfall in one task can be offset (with some probability) by a larger than expected outcome in the other task. This weakens incentives under broad bracketing relative to narrow bracketing. These negative incentive effects can counterbalance the benefits from risk pooling.

Our findings predict that for present-biased individuals who face tasks with little uncertainty about productivity, such as routine work, it is optimal to bracket narrowly (e.g., specify a daily work goal), whereas for individuals who face considerable uncertainty about the productivity of their daily effort it is optimal to bracket broadly (set, e.g., a monthly goal). Think of teaching preparation versus research. The former involves tasks with a close relationship between effort and outcomes. Hence, our model predicts that people set tight goals for teaching preparation, such as "every day, spend one hour preparing", or "prepare $x$ slides on a given day". In comparison, research involves tasks where success on a given day might be uncertain – despite high effort (think of proving a theorem). Hence, our model predicts that it is optimal to evaluate research outcomes over a longer period rather than evaluating the quality of a day's work. Similarly, we predict that it is suboptimal for dieters to weigh themselves daily, because day-to-day weight is subject to considerable fluctuations that are outside the control of the individual. This prediction is consistent with related advice in popular guides on goal setting.

Applying our analysis to situations where a present-biased individual makes an investment decision while facing some exogenous background risk, we show that it can be optimal to 'ignore' the background risk, i.e., to evaluate the investment decision and the background risk narrowly (Proposition 3). As we will point out, our predictions provide a new perspective on how to interpret some of the, at first glance, contradictory findings in recent experiments regarding whether or not there is a correlation between background shocks and the experimentally elicited intertemporal rate of substitution (Harrison et al., 2005; Meier and Sprenger, 2015; Tanaka et al., 2010; Giné et al., 2012; Krupka and Stephens, 2013; Dean and Sautmann, 2015).

In deriving our first set of results, we kept the incentives under broad and narrow bracketing as comparable as possible by considering solely the incentive to deviate from the goal in a single task, while sticking to the goal in the other task. Yet, unlike narrow brackets, a broad bracket offers the possibility to jointly deviate from the goals in both tasks. For instance, the individual can lower the decision in one task and compensate with a higher decision in the other task. That is, broad brackets provide fewer instruments of control than narrow brackets. Considering the full range of possible deviations under broad bracketing obviously can only strengthen our

[8] Kőszegi and Rabin (2009) show that a decision maker who experiences gain–loss utility over changes in beliefs about future consumption will behave as if he narrowly bracketed risks whose resolution is spread out over time. In contrast to our setting, such "news utility" however cannot explain why people might evaluate separately lotteries whose outcomes occur simultaneously or are deterministic.

findings on the optimality of narrow bracketing. Nevertheless, looking at joint deviations offers interesting additional insights into when narrow bracketing can be optimal (Proposition 4). What matters for the optimality of narrow bracketing is whether the individual is relatively biased towards one of the tasks. The occurrence of such a relative bias depends on the type of tasks or goods the individual is facing. For example, many people are tempted to overeat on chips and chocolate, viewing them as similarly tempting. Consequently, the individual's incentives to stick to his goals do not differ depending on whether these two goods are evaluated in a narrowly or broadly bracketed mental account. If, however, chips consumption was evaluated together with ordinary meals, the individual would deviate from his goal to not overeat chips because he balances the relatively unhealthy snack with a relatively healthy meal in his broad account. This is not possible if he evaluates consumption narrowly, like in 'snack' vs. 'ordinary meal' accounts. Thus, our finding helps understand why people evaluate mundane goods and tempting vice goods separately, i.e., adopt narrowly bracketed accounts for 'household expenses' and 'entertainment' – but not for every single item they buy; or why many diet programs advocate narrow, daily goals for calorie intake.

We consider several extensions and robustness checks to our model. Our main results are robust to introducing partial naïveté about self-control and we discuss the assumption of an additively separable utility function. Further, we extend our model to settings where the referent adapts to changes in the individual's expectations caused by the arrival of new information or by a revision of goals and brackets. Moreover, we apply our theory to a debate, triggered by Camerer et al. (1997), on why individuals who can choose their working hours (such as taxi drivers) often appear to have narrow, daily income targets. All of this literature exogenously imposes a daily evaluation horizon (see DellaVigna, 2009, p. 326). This is an important gap that we close by endogenizing the driver's evaluation horizon. We show (in Proposition 6) that it is optimal for a present-biased taxi driver to adopt narrow, single-day evaluation brackets in a setting that captures in a stylized way the patterns in the data of Farber (2005). Our second application illustrates the broader applicability of our framework. Taking up evidence that firms' marketing strategies influence how consumers' bracket their mental accounts, we show (in Proposition 7) how firms can unhinge consumers' self-regulation by bundling products together, such as is common in fast food menus.

## 2. The model

*The decisions* An individual faces two symmetric activities $i \in \{1, 2\}$, each calling for a decision $x_i \in \mathbb{R}_0^+$ at date $t = 1$ that involves a cost $c(x_i) > 0$ ($c' > 0$, $c'' > 0$) and provides a benefit. There is uncertainty about how productive each decision will be. For a given decision level, the benefit in activity $i$ can be either high or low: $b_{s_i}(x_i)$, $s_i \in \{L, H\}$, where $b_H(x_i) > b_L(x_i) \geq 0$, $b'_H(x_i) > b'_L(x_i) \geq 0$, $b''_H(x_i) \leq b''_L(x_i) \leq 0$ for all $x_i \in \mathbb{R}_0^+$, and the probability of a high-productivity state is $\Pr(s_i = H) = \pi \in [0, 1]$. Productivity draws for both activities are independent. To ensure unique interior solutions, we assume that $c(0) = b_{s_i}(0) = 0$, $b'_{s_i}(0) - c'(0) > 0$, and $\lim_{x_i \to \infty} b'_{s_i}(x_i) - c'(x_i) < 0$, $s_i \in \{L, H\}$. Let $y_{tik}(x_i)$ be the outcome that occurs at date $t \in \{1, 2\}$ for task $i \in \{1, 2\}$ in the cost dimension ($k = c$) or benefit dimension ($k = b$), and let $y_{tik}(x_i) = 0$ if there is no outcome. Denote a profile of outcomes at date $t$ by $y_t(x_1, x_2) = (y_{t1b}(x_1), y_{t2b}(x_2), y_{t1c}(x_1), y_{t2c}(x_2)) \in \mathbb{R}^4$. Our lead case has two investment activities, where costs arise at $t = 1$, while benefits realize at $t = 2$. That is, $y_{1ic}(x_i) = -c(x_i)$ and $y_{2ib}(x_i) = b_{s_i}(x_i)$.

*Design of goals and brackets* At $t = 0$, self 0 (the date-0 incarnation of the individual) sets goals $\hat{x}_i \in \mathbb{R}_0^+$ for the decisions in the two activities and sets the bracket of the mental account in which the goal is evaluated. Goals induce reference points. For deterministic outcomes, the referent is the outcome associated with goal $\hat{x}_i$, giving rise to a profile of reference standards $r_t(\hat{x}_1, \hat{x}_2) = (r_{t1b}(\hat{x}_1), r_{t2b}(\hat{x}_2), r_{t1c}(\hat{x}_1), r_{t2c}(\hat{x}_2)) \in \mathbb{R}^4$. For example, $r_{1ic}(\hat{x}_i) = -c(\hat{x}_i)$. Stochastic outcomes induce stochastic reference points, which we introduce below. How the individual evaluates the different outcomes from the two decisions relative to the reference standards is governed by the type of bracket, broad $(B)$ or narrow $(N)$ as outlined in the next paragraph. When making his decisions at $t = 1$, self 1 (the date-1 incarnation of the individual) takes these goals, brackets, and reference standards as given (we relax this in section 4.2).

*Instantaneous utility* The individual is expectation based loss averse in the sense of Kőszegi and Rabin (2006). Consider first the special case where outcomes are deterministic. The instantaneous utility under goal bracket $A \in \{B, N\}$ at date $t$ for outcomes $y_t(x_1, x_2) \in \mathbb{R}^4$ and goal dependent reference points $r_t(\hat{x}_1, \hat{x}_2) \in \mathbb{R}^4$ for the benefit and cost dimensions ($k \in \{b, c\}$) is given by (suppressing arguments in $y_t(x_1, x_2)$ and $r_t(\hat{x}_1, \hat{x}_2)$):

$$u_t^A(x_1, x_2|\hat{x}_1, \hat{x}_2) \equiv \tilde{u}_t^A(y_t|r_t) = \sum_{k\in\{b,c\}} \left( \sum_{i\in\{1,2\}} y_{tik} + n_k^A(y_t|r_t) \right), \quad \text{where}$$

$$n_k^N(y_t|r_t) = \sum_{i=1}^{2} \mu(y_{tik} - r_{tik}), \quad n_k^B(y_t|r_t) = \mu\left( \sum_{i=1}^{2} y_{tik} - \sum_{i=1}^{2} r_{tik} \right).$$

The instantaneous utility is composed of two components. The first component $(y_{tik})$ is outcome-based consumption utility from costs and benefits. Utility is separable across tasks and dimensions.[9] The second component $(n_k^A(y_t|r_t))$ is gain–loss utility. It compares, for each consumption dimension, the consumption utility with its reference point and takes the form of Kahneman and Tversky's (1979) value function. For tractability we assume a piece-wise linear gain–loss utility function.[10] If the consumption utility differs from its reference point by $z$, the corresponding gain–loss utility is $\mu(z) = \eta z$ for $z \geq 0$ and $\mu(z) = \eta \lambda z$ for $z < 0$, where $\eta > 0$ is the weight of gain–loss utility in the utility function, and $\lambda > 1$ is the coefficient of loss aversion.

Table 1 summarizes the instantaneous utilities under narrow and broad bracketing at the different dates. At $t = 0$, no payoff-relevant events occur, and we normalize $u_0 = 0$. Under narrow bracketing, each goal $\hat{x}_i$ induces a cost reference point $c(\hat{x}_i)$ in the respective activity $i \in \{1, 2\}$. At $t = 1$, the individual evaluates the actual costs by comparing, for each activity separately, the actual cost $c(x_i)$ with the expected cost $c(\hat{x}_i)$. Under broad bracketing, goals $(\hat{x}_1, \hat{x}_2)$ induce the cost reference point $c(\hat{x}_1) + c(\hat{x}_2)$ to which, at $t = 1$, the individual compares the actual joint costs of the two decisions $c(x_1) + c(x_2)$. At $t = 2$, benefits realize, where – more generally than in Table 1 – we allow for uncertainty about the productivity of the decisions. Thus, we need to introduce how the individual evaluates stochastic outcomes.

We assume that the individual has stochastic reference points (Kőszegi and Rabin, 2006, 2007). That is, stochastic outcomes are evaluated according to their expected utility, with the utility of each outcome being the average of how it feels relative to each possible realization of

[9] We discuss the assumption of linear separability in consumption utility and CARA preferences in section 4.1.

[10] Diminishing sensitivity is often considered an important determinant of decision making under risk. Our related Proposition 2 is robust to introducing diminishing sensitivity (details available upon request).

Table 1
Instantaneous utilities for deterministic outcomes ($\pi = 1$).

| Date | Broad bracketing | Narrow bracketing |
|---|---|---|
| $t=0$ | $0$ | $0$ |
| $t=1$ | $-c(x_1)-c(x_2)$ <br> $+$ <br> $\mu(c(\hat{x}_1)+c(\hat{x}_2)-(c(x_1)+c(x_2)))$ | $-c(x_1)-c(x_2)$ <br> $+$ <br> $\mu(c(\hat{x}_1)-c(x_1))+\mu(c(\hat{x}_2)-c(x_2))$ |
| $t=2$ | $b_H(x_1)+b_H(x_2)$ <br> $+$ <br> $\mu(b_H(x_1)+b_H(x_2)-(b_H(\hat{x}_1)+b_H(\hat{x}_2)))$ | $b_H(x_1)+b_H(x_2)$ <br> $+$ <br> $\mu(b_H(x_1)-b_H(\hat{x}_1))+\mu(b_H(x_2)-b_H(\hat{x}_2))$ |

the reference lottery. Denote by $G_t^A$ the probability measure from which the profile of reference points $r_t(\hat{x}_1, \hat{x}_2)$ is drawn and by $F_t^A$ the probability measure from which the profile of outcomes $y_t(x_1, x_2)$ is drawn, where $A \in \{B, N\}$. Then a given outcome realization $y_t$ is evaluated as $\tilde{u}_t^A(y_t|G_t^A) = \int \tilde{u}_t^A(y_t|r)dG_t^A(r)$. And from an ex ante perspective the instantaneous utility under bracket $A \in \{B, N\}$ is given by:

$$u_t^A(x_1, x_2|\hat{x}_1, \hat{x}_2) \equiv \tilde{u}_t^A(F_t^A|G_t^A) = \int \left[ \int \tilde{u}_t^A(y|r)dG_t^A(r) \right] dF_t^A(y).$$

Under narrow bracketing, a goal $\hat{x}_i$ induces a reference lottery for the benefit at $t = 2$ in activity $i$ of $(\pi \circ b_H(\hat{x}_i); (1-\pi) \circ b_L(\hat{x}_i))$. For each activity, the individual then evaluates the outcome $b_{s_i}(x_i)$ relative to each possible realization of the reference lottery. Similarly, under broad bracketing, the individual evaluates the realization of joint benefits $b_{s_1}(x_1) + b_{s_2}(x_2)$ relative to each possible realization of the reference lottery for the joint outcome $(\pi^2 \circ [b_H(\hat{x}_1) + b_H(\hat{x}_2)]$; $\pi(1-\pi) \circ [b_H(\hat{x}_1) + b_L(\hat{x}_2)]$; $\pi(1-\pi) \circ [b_L(\hat{x}_1) + b_H(\hat{x}_2)]$; $(1-\pi)^2 \circ [b_L(\hat{x}_1) + b_L(\hat{x}_2)])$.

*Present bias* The individual has a present bias, modeled using $(\beta, \delta)$-*preferences* (cf. Laibson, 1997). The parameter $\delta$ corresponds to the standard exponential discount factor (for simplicity, $\delta \equiv 1$). The parameter $\beta \in (0, 1)$ captures the extent to which the individual overemphasizes immediate utility flows relative to more distant utility flows. The expected utility at date $t \in \{0, 1, 2\}$ for decisions $(x_1, x_2)$, goals $(\hat{x}_1, \hat{x}_2)$, and bracket $A \in \{N, B\}$ is given by

$$U_t^A(x_1, x_2|\hat{x}_1, \hat{x}_2) = u_t^A(x_1, x_2|\hat{x}_1, \hat{x}_2) + \beta \left[ \sum_{\tau=t+1}^{2} u_\tau^A(x_1, x_2|\hat{x}_1, \hat{x}_2) \right].$$

For instance, self 0 weighs future utility flows $u_1$ and $u_2$ equally; but self 1 puts a larger relative weight on $u_1$ because he discounts $u_2$ with $\beta < 1$. As a consequence, an intra-personal conflict of interest arises. In our lead case with investment activities, the *preferred decisions of self 0* under bracket $A \in \{B, N\}$, $(x_0^A, x_0^A)$, exceed the *preferred decisions of self 1*, $(x_1^A, x_1^A)$, where $(x_t^A, x_t^A) = \arg\max_{x_1, x_2} U_t^A(x_1, x_2|x_1, x_2)$. Because our model is about deliberate self-regulation, it is natural to assume that the individual knows about his present-biased preferences (relaxed in section 4.1).

*Equilibrium* Goals, the expectations that they induce, and the decisions of the individual must constitute a *personal equilibrium* (Kőszegi and Rabin, 2006). On the equilibrium path, given the expectations induced by the goal $\hat{x}_i$, the actual decision $x_i$ must correspond to the goal.[11] That is, goals are rational. Any decision or goal that is consistent with a personal equilibrium under narrow (broad) bracketing is said to be *narrow- (broad-) bracketing implementable*. The aim of self 0 is to maximize his utility by choosing goals $(\hat{x}_1, \hat{x}_2)$ and bracket $A \in \{B, N\}$ subject to the constraint that the goals are implementable in the respective bracket, i.e., that it is optimal for self 1 to stick to these goals[12]:

$$\max_{\{A\in\{B,N\},(\hat{x}_1,\hat{x}_2)\in\mathbb{R}_0^+\times\mathbb{R}_0^+\}} U_0^A(\hat{x}_1, \hat{x}_2|\hat{x}_1, \hat{x}_2) \quad \text{s.t.} \tag{P}$$

$$\mathbb{1}_{\{A=N\}} \times \left[U_1^N(\hat{x}_i, \hat{x}_{-i}|\hat{x}_i, \hat{x}_{-i}) - U_1^N(x_i, \hat{x}_{-i}|\hat{x}_i, \hat{x}_{-i})\right] \geq 0,\ \forall x_i \in \mathbb{R}_0^+, i \in \{1, 2\}, \tag{1}$$

$$\mathbb{1}_{\{A=B\}} \times \left[U_1^B(\hat{x}_1, \hat{x}_2|\hat{x}_1, \hat{x}_2) - U_1^B(x_1, x_2|\hat{x}_1, \hat{x}_2)\right] \geq 0,\ \forall (x_1, x_2) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+. \tag{2}$$

## 3. Analysis

There is a wide-spread view that narrow bracketing of decisions is an error. In our context, this error is expressed in the result that it is optimal for a loss-averse individual to bracket broadly because this allows the individual to pool risks (e.g., Thaler et al., 1997; Kőszegi and Rabin, 2007). To illustrate the idea of risk pooling, consider a simple gamble with equal odds of winning \$200 or losing \$100 (based on Thaler et al., 1997). Assuming piecewise linear gain–loss utility with a reference point of zero and coefficient of loss aversion $\lambda = 2.5$, gain–loss utility from a narrowly evaluated gamble is $1/2\,(200) + 1/2\,(-250) < 0$, whereas with broad evaluation of two independent gambles it is $1/4\,(400) + 1/2\,(100) + 1/4\,(-500) > 0$. *Risk pooling* allows a gain of \$200 in one gamble to offset a loss of \$100 in the other gamble, reducing the chances of falling into the loss domain from $1/2$ to $1/4$. Utility under broad bracketing exceeds that under narrow bracketing because avoided losses count more than foregone benefits. The following result shows that this intuition carries over to our setting. (All proofs are in Appendix B.) Specifically, if there was no incentive problem (because decision levels are fixed or self 0 could implement his preferred decisions), then self 0 would achieve a higher utility under broad than under narrow bracketing.

**Proposition 1.**

1. *Suppose the decision levels in both tasks $i \in \{1, 2\}$ are fixed at $x_i = x > 0$. Then for all $\pi \in (0, 1)$ the utility of self 0 is strictly higher when he brackets broadly than when he brackets narrowly. If benefits are deterministic ($\pi \in \{0, 1\}$) the utility of self 0 is the same under broad and narrow bracketing.*
2. *The preferred decision of self 0 is (weakly) higher under broad bracketing than under narrow bracketing: $x_0^B > x_0^N$ for all $\pi \in (0, 1)$ and $x_0^B = x_0^N$ for $\pi \in \{0, 1\}$.*

An immediate implication of Proposition 1 is that the maximized utility of self 0 would be higher under broad than under narrow bracketing if self 0 was able to commit to his preferred decision: $U_0^B(x_0^B, x_0^B|x_0^B, x_0^B) \geq U_0^N(x_0^N, x_0^N|x_0^N, x_0^N)$ with strict inequality for $\pi \in (0, 1)$. That

11 Allowing to freely choose the reference point yields similar results (Koch and Nafziger, 2009).

12 Because utility under narrow bracketing is additive separable in the two tasks, one only needs to consider unilateral deviations in constraint (1).

is, in the absence of an incentive problem narrow bracketing would be suboptimal. However, for a present-biased individual one or both of the two incentive constraints for self 1 (inequalities (1) and (2)) may bind. The novel aspect that our model emphasizes is that broad bracketing may weaken the incentives of self 1, as compared to narrow bracketing. These *incentive effects* of broad bracketing can counteract the benefit from risk pooling as the following result shows.

**Proposition 2.** *Suppose that under broad bracketing self 1 has an incentive to deviate in one task from the goal $\hat{x}_i = x_0^B$ to a lower decision, i.e., the incentive constraint (2) is violated at goals $\hat{x}_i = x_0^B$, $i \in \{1, 2\}$, because there exists an $x_i \in (0, x_0^B)$ such that $U_1^B(x_0^B, x_0^B | x_0^B, x_0^B) < U_1^B(x_i, x_0^B | x_0^B, x_0^B)$. Then there exist thresholds $\underline{\pi}, \bar{\pi} \in [0, 1)$, where $\underline{\pi} \leq \bar{\pi}$, such that narrow bracketing is strictly optimal if $\pi \in (0, \underline{\pi}) \cup (\bar{\pi}, 1)$.*

One can always find a self-control problem 'severe enough' such that self 0 cannot incentivize self 1 under broad bracketing to provide the preferred decision from the perspective of self 0, i.e., such that the incentive constraint under broad bracketing (2) is violated.[13] In such cases, the best implementable decision under broad bracketing may be worse than the best implementable decision under narrow bracketing. Proposition 2 provides conditions on $\pi$, the probability of the high-productivity state, such that the ability to implement a better decision under narrow bracketing trumps the benefits from risk pooling under broad bracketing. One can easily construct examples where narrow bracketing is strictly optimal for all $\pi \in (0, 1)$.[14]

To understand the driving forces behind Proposition 2, we now examine how goal bracketing shapes incentives to live up to one's goals. Note that one can establish sufficient conditions for the optimality of narrow bracketing solely by considering unilateral deviations from one of the goals under broad bracketing, rather than considering the full range of possible deviations from goals.[15] To facilitate exposition, discussions center around the special case where $b_L(x_i) = 0$ for all decisions ("a failure") and $b(x_i) \equiv b_H(x_i)$ ("a success"). We relegate details for the general case to Appendix A.

*Incentives under narrow bracketing* To know what decisions self 0 can implement in, say, task 1, we need to check for what goal levels $\hat{x}_1$ self 1 has no incentive to deviate from the goal, given the reference points that the goal induces. We will show that the incentive constraint (1) holds under narrow bracketing for any goal $\hat{x}_1 \in [x_{min}^N, x_{max}^N]$, where for our purposes the *maximal implementable goal*, $x_{max}^N$, is most important ($x_{min}^N$ is defined in appendix A.2). To illustrate the main ideas, let us first assume that there is no uncertainty about task outcomes ($\pi = 1$). By sticking to the goal, self 1 meets expectations, and there will be neither a gain nor a loss when outcomes from the task are evaluated. The utility of self 1 is $U_1^N(\hat{x}_1 | \hat{x}_1) = \beta\, b(\hat{x}_1) - c(\hat{x}_1)$.[16] If self 1 however deviates from the goal to a lower decision level, this causes a gain be-

[13] For any $\pi$ and finite $\lambda$ one can find a $\beta$ such that $U_1^B(x_0^B, x_0^B | x_0^B, x_0^B) < U_1^B(x_i, x_0^B | x_0^B, x_0^B)$. Further, the existence of the thresholds $\underline{\pi}, \bar{\pi} \in [0, 1)$ does not rely on a specific $\beta$ (or $\lambda$).

[14] For example, we get $\underline{\pi} = \bar{\pi} = 0$ for $b_H(x_i) = x_i$, $b_L(x_i) = \alpha\, x_i$, $\alpha \in [0, 1)$, and $c(x_i) = (x_i^2\, c)/2$, $c \in \mathbb{R}^+$ with parameterization $\eta = 1/2$, $c = 1/8$, $\alpha = 3/4$, $\lambda = 2$, $\beta = 0.7$.

[15] Checking whether self 1 has an incentive to deviate from his goal in a single activity while sticking to the goal in the other activity, places an upper bound on what decisions are implementable under broad bracketing, which in turn places an upper bound on the maximized utility that then can be compared with the maximized utility under narrow bracketing.

[16] To simplify exposition, we slightly abuse notation here: Without loss we can focus on the utility from one task by writing $U_1^N(\hat{x}_1, \hat{x}_2 | \hat{x}_1, \hat{x}_2) = U_1^N(\hat{x}_1 | \hat{x}_1) + U_1^N(\hat{x}_2 | \hat{x}_2)$.

cause costs are lower than expected; and it causes a loss from falling short of the expected benefits: $U_1^N(x_1|\hat{x}_1) = \beta\, b(x_1) - c(x_1) + \eta\,\beta\,(c(\hat{x}_1) - c(x_1)) - \eta\,\beta\,\lambda\,(b(\hat{x}_1) - b(x_1))$. Consequently, for goal $\hat{x}_1$ to be implementable, self 1 should have no incentive to lower his decision: $U_1^N(\hat{x}_1|\hat{x}_1) \geq U_1^N(x_1|\hat{x}_1)$ for all $x_1 < \hat{x}_1$. To ensure that the utility from sticking to the goal, $U_1^N(\hat{x}_1|\hat{x}_1)$, exceeds the utility from falling short of it, $U_1^N(x_1|\hat{x}_1)$, the latter has to be increasing in $x_1$ for any $x_1 < \hat{x}_1$. This is the case for any goal less or equal to the *maximal implementable goal*, $x_{max}^N$, which for $\pi = 1$ is defined by

$$\beta\,(1 + \eta\,\lambda)\,b'(x_{max}^N) = (1 + \eta)\,c'(x_{max}^N). \tag{3}$$

What motivates self 1 is the fear of facing a loss in the benefit dimension in case he falls short of the goal. Thereby, loss aversion can counterbalance the present bias. Whenever $x_0^N \leq x_{max}^N$, self 0 can even implement his preferred decision. The larger the present bias parameter $\beta$, the larger the maximal implementable goal, i.e., the more likely it is that self 0 can fully overcome his self-control problem. Self-regulation however is constrained if the individual faces a more severe self-control problem such that $x_0^N > x_{max}^N$. In this case, self 0 sets $x_{max}^N$ as his goal. Still, self 0 can implement a higher decision level than self 1 would want on his own. It can be shown that $x_{max}^N$ exceeds the preferred decision of self 1, $x_1^N$.

The discussion makes evident that the incentive effects of a goal stem from the fear of facing a loss in case self 1 lowers his decision relative to the goal. With deterministic outcomes, the individual faces a loss with probability one once he deviates. With stochastic outcomes, a deviation from the goal only matters if the decision actually has an impact on the outcome, i.e., if the outcome is a success. This happens with probability $\pi$. The outcome $b(x_i)$ then only feels like a loss when the reference standard is $b(\hat{x}_i)$, to which the reference lottery assigns probability $\pi$. That is, the probability that a deviation from the goal causes a loss is only $\pi^2$. For $\pi \in [0, 1]$, the maximal implementable goal $x_{max}^N$ is defined by:

$$\beta\,\{\pi + \eta\,[\pi + (\lambda - 1)\,\pi^2]\}\,b'(x_{max}^N) = (1 + \eta)\,c'(x_{max}^N). \tag{4}$$

*Incentives under broad bracketing* When considering solely unilateral deviations from the goal in one of the activities, the incentives of self 1 under broad bracketing exactly mirror the ones under narrow bracketing if outcomes are deterministic. With stochastic outcomes, however, risk pooling kicks in. A broad bracket cancels out some of the losses that occur under narrow bracketing, and thus the probability that a deviation causes a loss is different under broad and narrow bracketing. Checking for unilateral deviations provides us with an upper bound on implementable goals under broad bracketing, which we denote by $x_{ub}^B$ and which is given by[17]:

$$\beta\left[\pi + \eta\left\{\pi + \pi^2\,(\lambda - 1)\,(1 + (1 - \pi)\,(1 - 2\pi))\right\}\right]b'(x_{ub}^B) = (1 + \eta)\,c'(x_{ub}^B). \tag{5}$$

How does $x_{ub}^B$ compare to $x_{max}^N$? Risk pooling offsets a lower than expected outcome in one task with a higher than expected outcome in the other task. This has a positive and a negative effect on incentives under broad in comparison to narrow bracketing. To explain these effects, it is convenient to label the success probabilities of the two decisions by $\pi_i$, still assuming $\pi_1 = \pi_2$.

---

[17] Inequality (2) imposes the additional constraint that self 1 has no incentive to deviate from his goal in the other task either, which can only shrink the set of broad-bracketing implementable goals, i.e., the maximal implementable goal $x_{max}^B \leq x_{ub}^B$.

Let us start with the negative effect on incentives. Consider a unilateral deviation from the goal in task 1 while sticking to the goal in task 2: $x_1 < x_2 = \hat{x}$. Under narrow bracketing, the probability of a loss in task 1 is $\pi_1^2$. With probability $\pi_1^2$, the individual expects to succeed in task 1, and the decision actually matters. Under broad bracketing, the individual however does not always incur a loss when he would do so under narrow bracketing. With probability $\pi_2(1-\pi_2)$, the individual expects to fail in task 2, but actually succeeds. And this success provides a buffer against a loss. Even if self 1 deviates from the goal for task 1, he can meet the reference state $b(\hat{x}_1) + 0$ with the outcome $b(\hat{x})$ in task 2. Overall, compared to narrow bracketing, broad bracketing reduces the probability that a deviation from the goal in task 1 causes a loss by

$$\pi_1^2 \pi_2 (1 - \pi_2). \tag{6}$$

Let us next consider the positive effect on incentives.[18] Decision $x_1$ not only affects losses in task 1, but it can also provide a buffer against losses in task 2. With probability $\pi_1(1-\pi_1)$, the individual expects to fail in task 1, but the outcome is actually $b(x_1)$. In this case, a deviation to $x_1 < \hat{x}$ does not cause a loss under narrow bracketing. The reason is that any $b(x_1) \geq 0$ exceeds the reference state of 0. But with broad bracketing such a deviation causes a loss if, in addition, the individual expects to succeed in task 2 and is unlucky. That is, with probability $\pi_2(1-\pi_2)$, the joint outcome $b(x_1) + 0$ falls short of the joint reference state $0 + b(\hat{x})$ as soon as $x_1 < \hat{x}$. Compared to narrow bracketing, the probability that a deviation from the goal in task 1 causes a loss increases by

$$\pi_1 (1 - \pi_1) \pi_2 (1 - \pi_2). \tag{7}$$

Comparing (6) and (7) shows that the negative incentive effect dominates, i.e., $x_{max}^N \geq x_{ub}^B$, if and only if $\pi \geq 1/2 \equiv \bar{\nu}$ (and $\underline{\nu} = 0$), which bounds the thresholds in Proposition 2: $\bar{\pi} \in [\bar{\nu}, 1)$. More generally, allowing for $b_L(x_i) > 0$, we obtain the following result:

**Lemma 1.** *There exist thresholds $\underline{\nu}, \bar{\nu} \in [0, 1)$, where $\underline{\nu} \leq \bar{\nu}$, such that $x_{max}^N > x_{ub}^B \geq x_{max}^B$ if $\pi \in (0, \underline{\nu}) \cup (\bar{\nu}, 1)$.*

In the example above with $b_L(x_i) = 0$, we have seen that $\bar{\nu} = 1/2$ and $\underline{\nu} = 0$, such that $x_{max}^N > x_{ub}^B \geq x_{max}^B$ for $\pi > 1/2$. It is not difficult to construct more extreme examples, namely examples where $x_{max}^N > x_{ub}^B \geq x_{max}^B$ for all $\pi \in (0, 1)$. For instance, with $b_L(x_i) = \alpha\, b_H(x_i)$, $\alpha > 1/4$ is a sufficient condition for the thresholds $\underline{\nu}, \bar{\nu}$ to coincide and to take the lowest possible value, $\underline{\nu} = \bar{\nu} = 0$.

*Background risk* The insights from Proposition 2 can be extended to settings where one activity is replaced by exogenously given background risk (a risk that cannot be avoided). The following result addresses how an individual brackets such a background risk when making a consumption-investment decision.

**Proposition 3.** *Consider an individual who faces exogenously given background risk that yields $y \in \mathbb{R}^+$ with probability $\pi_2$ and zero with probability $(1 - \pi_2)$, where $\pi_2 \in (0, 1)$. Suppose he can*

[18] The positive incentive effect relies on stochastic reference points. With deterministic reference points, only the negative incentive effect arises (Koch and Nafziger, 2009).

*invest $x$ at cost $c(x)$ today to obtain a stochastic future reward that yields $b(x)$ with probability $\pi_1$ and zero with probability $(1-\pi_1)$. Further, suppose the incentive constraint (2) is violated at $(x_0^B, y)$, because there exists an $x \in (0, x_0^B)$ such that $U_1^B(x_0^B, y|x_0^B, y) < U_1^B(x, y|x_0^B, y)$. Then there exists some threshold $\check{\pi}_1 < 1$, such that the individual brackets the background risk and the investment activity narrowly if $\pi_1 \in (\check{\pi}_1, 1]$.*

Proposition 3 reveals that ignoring favorable background risk may be optimal for a present-biased individual who faces an investment decision. Narrow bracketing of the investment decision makes self 1 more sensitive to losses created by a deviation from the goal than when the investment decision and the background risk are bracketed together. With broad bracketing losses that result from a deviation from the investment goal are sometimes buffered by the background risk. Thus, they are less likely to result in the experience of an overall loss than with a narrowly bracketed investment goal. That is why narrow bracketing increases the individual's incentives to stick to the investment goal relative to broad bracketing.

The proposition states that present-biased people who make an investment decision that does not involve too much uncertainty ($\pi_1 \in (\check{x}_1, 1]$) do not take into account background risk so as to stay motivated for the investment decision. Consider, for example, young versus old people making an investment decision. For many such decisions, such as searching for a new job, the investment is more risky for older people, because they have a shorter investment horizon or because it is more difficult for older people to find a new job. So Proposition 3 predicts that older people do take background risk (like health shocks or mortality) into account when making investment decisions. In contrast, young people may not take into account background risk so as to stay motivated for the investment decision. This is consistent with the finding that investment decisions of individuals aged 50 and above are affected by anticipated future health risk (Atella et al., 2012), whereas studies including younger individuals typically find little impact of health on investment decisions (e.g., Love and Smith, 2010).

For investment decisions that have a certain return ($\pi_1 = 1$), the predictions of Proposition 3 are consistent with some (puzzling) observations made in experiments on the elicitation of time preferences. In these experiments, subjects typically choose between a lower amount to be paid "today" and a higher amount to be paid "in $x$ months" (these choices yield the intertemporal marginal rate of substitution). The task in these experiments hence mirrors an investment activity with no uncertainty in our model. The question is whether the marginal rate of substitution correlates with background risk or not. If not, this can be seen as an indication of narrow bracketing. If yes, it is an indication of broad bracketing. Indeed, Harrison et al. (2005), Meier and Sprenger (2015), Tanaka et al. (2010) and Giné et al. (2012) all observe no, or no robust correlation between some background risk and the intertemporal marginal rate of substitution. This is consistent with our result. However, Krupka and Stephens (2013) and Dean and Sautmann (2015) find a correlation between the marginal rate of substitution and exogenous background risk. Such a correlation, as well as the absence of a correlation in the other studies, is consistent with our model once we allow for non-CARA preferences. Specifically, we show that endogenous narrow bracketing is robust to the introduction of non-CARA preferences (see appendix B.4). It is optimal for the present-biased individual to evaluate an investment decision separately from a background risk with adverse shocks that occur with low probability. And indeed, in the mentioned experiments, whenever a correlation between the marginal rate of substitution and shocks stemming from background risk is observed, the shocks seem to occur with a rather high probability. Conversely, when no correlation is observed, the shocks seem to occur with a rather low probability (see appendix B.4 for details on the specific studies).

In contrast to the time-preference elicitation tasks, experimental elicitation of risk preferences does not involve an investment task (costs and benefits occur at the same date). Hence, the incentive considerations that drive the optimality of narrow bracketing play no role, and we predict that the choices in the risk-preference elicitation task are correlated with background risk. This prediction is consistent with the observations of Guiso and Paiella (2008) and Tanaka et al. (2010).

*Preventing joint deviations from the goals* So far we considered the incentives under broad bracketing to deviate from the goal in one task while sticking to the goal in the other task. This allowed to establish sufficient conditions for the optimality of narrow bracketing.[19] Yet, this approach ignores an additional disadvantage of broad bracketing. Namely, under broad bracketing the individual potentially has an incentive to jointly deviate from the goals in both tasks. In the special case used for exposition, we can show that Proposition 2 yields a necessary and sufficient condition:

**Corollary 1.** *Suppose $b_L(x_i) = 0$ for all $x_i$. Then $x^B_{max} = x^B_{ub}$ and $x^B_{min} = x^B_{lb}$. If $x^B_{max} < x^B_0$ there exists a cutoff $\hat{\pi} \in (1/2, 1)$, such that for $\pi \in (\hat{\pi}, 1)$ narrow bracketing is strictly optimal and for $\pi \in (0, \hat{\pi}]$ broad bracketing is strictly optimal.*

The individual has no incentive to jointly deviate in both tasks because he faces two so-called investment activities (both have immediate costs at $t = 1$ and delayed benefits at $t = 2$). To study the impact of joint deviations, we hence extend our model and consider different categories of activities. Each activity can either be an investment activity, or a leisure activity, such as consumption of unhealthy food (where benefits realize at $t = 1$ and costs realize at $t = 2$), or a neutral activity (where benefits and costs both realize at $t = 1$).[20]

**Proposition 4.** *Suppose $\pi \in \{0, 1\}$.*

1. *Suppose both activities belong to different categories (a neutral activity together with an investment or a leisure activity, or a leisure and an investment activity). Then narrow bracketing is strictly optimal.*
2. *Suppose both activities belong to the same category (investment, leisure, or neutral activities). Then narrow and broad bracketing are both optimal.*

The result makes transparent how narrow bracketing can be driven solely by the need to deter joint deviations. It complements Proposition 2 by turning off the risk effects that drive the (sub)optimality of broad bracketing there. With deterministic outcomes, the utility of self 0 from a given decision is the same under narrow and broad bracketing (Proposition 1), and the marginal incentives to deviate from the goal in a single task are the same, i.e., $x^N_{max} = x^B_{ub}$.

To build some intuition for part 1, suppose that task 1 is a neutral activity and task 2 an investment activity. Under broad bracketing, self 1 can increase his utility by reducing a bit consumption in the task which is relatively less attractive to him (the investment activity) and

[19] Considering possible joint deviations can only strengthen our result on the optimality of narrow bracketing in Proposition 2.

[20] We assume that gain–loss utility accrues once the individual is in a position to evaluate all outcomes in an bracket. So if the individual faces activities where the costs or benefits occur at different points in time, evaluation of a broad bracket in one or both dimensions $k \in \{b, c\}$ occurs in $t = 2$.

increase it a bit in the other activity (the neutral activity) in such a way that overall costs are kept at their reference level. The described joint deviation causes no gain–loss utility in the cost dimension, but self 1 gains $b'(\cdot)$ in consumption utility on the margin by 'borrowing' consumption from his future self at 'price' $\beta\, b'(\cdot)$. The proof establishes that these utility gains are of first order while the losses in the benefit dimension caused by the deviation are of second order.

We use the term *decision substitution* to refer to joint deviations such as the one just described, in which (relative to his goals) the individual increases the level in one activity to (partially) substitute for a decrease in the other activity.[21] Decision substitution may at first glance seem to be nothing else than the expression of the present bias we commonly see in self-control problems.[22] Part 2 of Proposition 4, however, reveals that the mere existence of a self-control problem in both tasks is not sufficient for broad bracketing to do strictly worse than narrow bracketing. Intuitively, if both decisions involve only delayed benefits and immediate costs (as in our lead case with two investment activities), self 1 and self 0 perceive in the same way the relative trade-off between the two activities.[23] Consequently, self 1 has no incentive to increase one decision and decrease the other. As a result, self 0 can implement the same decisions under broad and narrow bracketing.[24]

Proposition 4 helps to understand how an individual brackets his mental accounts. What matters for the optimality of broad versus narrow bracketing is whether the individual is relatively biased toward one of the decisions. For example, he might well group several leisure activities or several investment activities together. But it is not optimal to group leisure and investment activities together, as this would lead to overconsumption in the leisure activity and underconsumption in the investment activity. How narrowly defined the mental accounts need to be depends on the consumption preferences of the individual and the relative trade-offs that these preferences imply. For example, a bracket that subsumes all kinds of food can be optimal for an individual who is not tempted to overeat on unhealthy food (because, despite his present bias, he quickly reaches a point of satiation). But it would be suboptimal for a person who struggles with a diet. Such a person is better off with narrower bracketed mental accounts such as "fruits and vegetables" and "sweets", or "main meals" and "snacks". A goal could be to follow the American Cancer Society's advice to eat five servings of fruits and vegetables a day. A similar intuition applies to spending on food and drinks. When dining in a restaurant, most people are more biased to

[21] Two lab experiments provide evidence for decision substitution. Khan and Dhar (2007) offer participants a virtue-vice consumption decision, finding that 48 percent of participants chose the vice good over the virtue good in a one-shot treatment and that this share rises to 70 percent in a treatment were participants are told they can make the same choice again in the following week. Zhang et al. (2007) observe that anticipated progress toward a fitness goal increases the likelihood of indulgence in unhealthy food right now.

[22] Decision substitution requires a link between decisions that could also occur through a budget constraint, time constraint, or any other interdependence of rewards. Our additional assumption of loss aversion in fact makes it more difficult for decision substitution to arise because under broad bracketing joint deviations cause a loss in at least one dimension.

[23] In geometric terms, the iso-benefit and iso-cost curves for self 0 and self 1 overlap because the present-bias 'cancels out' from self 1's iso-benefit curve. As a consequence, at the preferred decision of self 0, the iso-benefit and iso-cost of self 1 are tangent, and he has no incentive to deviate.

[24] An immediate corollary of Proposition 4 is that goal setting under narrow bracketing is strictly optimal for a twice repeated investment (or leisure) activity if $x_0^N \in (x_{min}^N, x_{max}^N)$, because self 1 is tempted to postpone providing effort. Motivational slack in one task is required to create an incentive to fall short of the goal in the other task. If $x_0^N \geq x_{max}^N$ the individual's goals already push his future self to the limit and provide commitment not to go "over target" to compensate for an earlier shortfall. See our previous working paper (Koch and Nafziger, 2011a).

overconsume on drinks than on food, and indeed consumers appear to bracket food and drinks in separate mental accounts (cf. Abeler and Marklein, 2010).

## 4. Discussion and extensions

### 4.1. Discussion on robustness

*Partial naïveté about self-control* People often overestimate their future self-control. Suppose the individual holds an overly optimistic belief $\hat{\beta} \in (\beta, 1]$ about his present bias. Adjusting the equilibrium concept to assume "perception-perfect" strategies (O'Donoghue and Rabin, 1999), self 0 chooses goals to maximize utility given his (incorrect) beliefs about future behavior and forms corresponding expectations about future decisions. Our results still hold, essentially by replacing $\beta$ with the individual's belief $\hat{\beta}$ in the proofs.

*Linear, additively separable utility function* Our assumption of a linear, additively separable utility function amounts to imposing CARA preferences on the consumption utility part. We make the assumption of separability in consumption utility for two reasons. First, any narrowness in the evaluation is imposed endogenously by self 0 and is not imposed exogenously through properties of the utility function or as part of the constraint set. Second, it rules out that our predictions of narrow bracketing require the individual to ignore interdependencies between decisions that affect marginal consumption utility. Such a theory seems less convincing, because it would imply, say, that a thirsty person could narrowly bracket each sip of water he drinks to experience it as if it was the first sip.

However, the assumption implies that the non-linear aspects of utility are only in gain–loss utility, not in consumption utility. Specifically, risk pooling under broad bracketing has no advantage (relative to narrow bracketing) in terms of consumption utility, but only in terms of gain–loss utility. This advantage in gain–loss utility suffices already to show that broad bracketing does strictly better than narrow bracketing in the absence of a self-control problem (Proposition 1). Rabin and Weizsäcker (2009) show that a narrow bracketer with non-CARA utility makes combinations of choices that are first-order stochastically dominated. Thus, in Proposition 1, assuming non-CARA preferences in consumption utility would further increase self 0's utility under broad bracketing relative to narrow bracketing. Yet, as we outline in appendix B.4, such non-CARA preferences also imply an additional negative incentive effect under broad bracketing. In the specific set-up of Proposition 3, we show that narrow bracketing can still dominate broad bracketing for non-CARA preferences.

### 4.2. Goal revision and arrival of new information

People may learn about the performance in a past task before they start with a new task. For example, a student may get to know his grade in one exam before he studies for the next exam. Or, before working on a task, people may receive information about how productive their effort will be. For example, a cab driver may discover from observing the weather at the start of his shift whether it will be a busy or a slow day (we revisit this specific application in section 4.3). Moreover, people may revise their goals before a decision. The arrival of new information or a revision of goals change the individual's beliefs about the outcomes that will occur. In some settings, it may be plausible that the reference distribution adjusts quickly enough to such changes in beliefs to affect decisions. We extend the model to allow for such adjustments of the reference

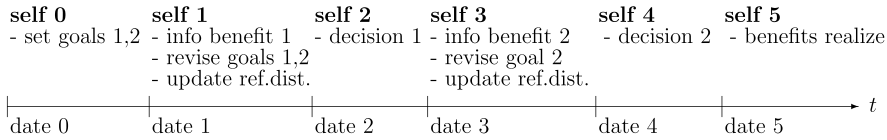


Fig. 1. Timing in the model extension of section 4.2.

distribution to new information and to goal revisions to examine whether narrow bracketing can still be optimal.

Fig. 1 summarizes the timing. We extend our lead case with two symmetric investment activities, assuming that $b_L(x_i) = 0$ for all $x_i \in \mathbb{R}_0^+$ and denoting $b(x_i) \equiv b_H(x_i)$. Self 0 designs goals $\hat{x}_{0,i}$ for activities $i \in \{1, 2\}$ that take place at $t = 2$ and $t = 4$, respectively. To allow for goal adjustment and arrival of information, we introduce pre-decision periods $t \in \{1, 3\}$. In these pre-decision periods, the individual learns the realization of the productivity for the activity at $t + 1$, possibly revises his goals for the activities, and updates the reference distribution. Given the (revised) goals and reference distribution, the individual then makes the decisions at $t = 2$ and $t = 4$, respectively. To make the problem of goal revision more interesting, we assume that the individual discounts outcomes in the same way at the pre-decision and the decision dates, i.e., we assume that selves 1 and 2 (selves 3 and 4) make the same intertemporal trade-offs. Costs are felt when decisions are made, and the delayed benefits from both tasks realize at $t = 5$.

We build on the ideas from Kőszegi and Rabin (2009) that changes in beliefs are carriers of utility which trigger "anticipatory" gain–loss utility. Anticipatory utility compares how an account will be evaluated under the revised reference distribution to how it would have been evaluated under the old reference distribution. We assume that news about current consumption being different from its expected level resonates more strongly than news that future consumption will differ from its expected level. That is, gain–loss utility from the evaluation of an account at date $\tau$ that is affected by a revision of reference distributions at date $t < \tau$ has weight $\gamma < 1$, whereas it has weight one if $t = \tau$.

To illustrate how to calculate anticipatory gain–loss utility, let us briefly consider deterministic task outcomes. If the individual revises his goal for task 1, he compares the outcome that arises under the new goal, $\hat{x}_1$, with the outcome that would have occurred under the past goal of self 0, $\hat{x}_{0,1}$. This causes anticipatory gain–loss utility $\gamma\,\mu(b(\hat{x}_1) - b(\hat{x}_{0,1})) + \gamma\,\mu(c(\hat{x}_{0,1}) - c(\hat{x}_1))$. With stochastic outcomes, following Kőszegi and Rabin (2009), the individual compares the worst percentile of outcomes under the new distribution with the worst percentile of outcomes under the old distribution, then the second worst percentile of outcomes, and so on. In the following, we describe results intuitively – relegating details to appendix B.7. There we also illustrate how to calculate the anticipatory utility with stochastic outcomes.

*Analysis* The optimization problem of self 0 differs in two ways from problem $(P)$ in section 2. The first difference is that goals can be revised, and that deviations from goals can occur at several dates. This means that implementable goals under narrow bracketing are pinned down by several incentive constraints of the form:

$$U_t^N(\hat{x}_{s,i}, \hat{x}_{s,-i}|\hat{x}_{s,i}, \hat{x}_{s,-i}) \geq U_t^N(x_{t,i}, \hat{x}_{s,-i}|\hat{x}_{s,i}, \hat{x}_{s,-i}), \quad i \in \{1, 2\}, s < t, t \in \{1, \ldots, 4\}.$$

At the decision dates ($t = 2$ and $t = 4$, respectively), the individual faces no uncertainty about the outcome in the activity. By arguments similar to those given for the derivation of equation (3), the individual has no incentive to deviate from the goal for the task at date $t \in \{2, 4\}$ if it does not exceed $x^N_{max,t}$, defined by

$$(\beta + \eta \gamma \lambda) \, b'(x^N_{max,t}) = (1 + \eta \gamma) \, c'(x^N_{max,t}). \tag{8}$$

Comparing equations (3) and (8), the only difference is that gain–loss utility in the benefit and cost dimension receives weight $\gamma$ (instead of $\beta$ and 1, respectively). Depending on whether $\gamma$ is smaller or larger than $\beta$, $x^N_{max,t}$ may be smaller or larger than the threshold defined by equation (3). At $t = 1$ ($t = 3$) the individual may revise his goal for task 2 (4) after observing the state. If the task is in the unproductive state, he has no incentive to do so. In the productive state, this information and a possible downward revision of his goal trigger anticipatory utility. Again the individual has no incentive to revise his goal downward as long as it does not exceed some threshold. A third threshold prevents self 1 from revising the goal for the productive state of task 2. These thresholds are derived in a similar fashion as $x^N_{max,t}$. In sum, the possibility of goal revision does not fundamentally alter the optimization problem, but just adds additional constraints on narrow-bracketing implementable goals. It is now the minimum of two (three) bounds that pins down the maximal implementable goal for activity 1 (2).

The second difference is that risk pooling no longer is possible under broad bracketing. More formally, we show that, for fixed decision levels, the utility of self 0 under narrow bracketing is the same as the one under broad bracketing. Intuitively, as soon as the individual experiences a bad outcome, this lowers his expectations for the overall outcome. The resulting adjustment of his reference distribution to lower expectations causes him negative anticipatory utility. If better outcomes occur at a later date, they come too late to avoid the sensation of a loss, because the reference point adjustment already has taken place.

The fact that for fixed decisions the utility under a broad bracket equals the one under narrow brackets implies, first, that self 0 has the same preferred decisions under broad and narrow bracketing and, second, that narrow bracketing is strictly optimal if self 0 cannot implement his preferred decisions under a broad bracket but can do so under narrow brackets. Related to the intuition for Proposition 4, this situation occurs if broad brackets induce decision substitution. The individual prefers to shift costs to the future under broad bracketing. That is, self 1 would lower his goal for task 1 and revise upward his goal for task 2 to (imperfectly) compensate. As, however, task 2 might fail, self 1 only likes to do so if task 2 has a sufficiently high success probability, i.e., only if $\pi$ exceeds some threshold $\hat{\pi} \in (0, 1)$.

**Proposition 5.** *Consider the model extension where revisions of the reference distribution trigger anticipatory utility.*

1. *The risk-pooling effect disappears: Holding decision levels constant, the utility and preferred decisions of self 0 are equal under broad and narrow bracketing.*
2. *Suppose the preferred decisions of self 0 are narrow-accounting implementable. Then there exists a* $\hat{\pi} \in (0, 1)$*, such that for* $\pi \in (\hat{\pi}, 1)$ *narrow bracketing is strictly optimal.*

*Revising brackets* Being the architects of their own mental brackets, individuals may redefine these brackets. To illustrate the process, consider deterministic outcomes. Suppose self 0 adopted narrow brackets and set the goals $(\hat{x}_{0,1}, \hat{x}_{0,2})$. A later deviation to decisions $(x_1, x_2)$ would lead to gain–loss utility $\sum_i [\mu(b(x_i) - b(\hat{x}_{0,i})) + \mu(c(\hat{x}_{0,i}) - c(x_i))]$. The constraints that we described above guarantee that such a deviation does not pay off. Suppose now that we allow not only for goal revision, but also for a revision of brackets. Then self 1 may want to switch from narrow to broad bracketing to facilitate decision substitution that increases his consumption utility. But if self 1 anticipates that his future selves will deviate from the original goals, he must revise the goals to match these expectations. Along with the revision of the brackets, self 1 updates the goals from $(\hat{x}_{0,1}, \hat{x}_{0,2})$ to the new goals $(\hat{x}_{1,1}, \hat{x}_{1,2})$. This triggers anticipatory utility from comparing outcomes under the new goals with those that would have occurred under the past goals. Parallel to the assumption that past reference points still matter when expectations change, we assume that an individual who just started thinking about his new, broad bracket still experiences anticipatory gain–loss utility in the old, narrow brackets. So the revision of goals and brackets causes anticipatory utility of $\gamma\,\{\sum_i[\mu(b(\hat{x}_{1,i}) - b(\hat{x}_{0,i})) + \mu(c(\hat{x}_{0,i}) - c(\hat{x}_{1,i}))]\}$. The only way to avoid the (negative) anticipatory feelings is not to revise the brackets. The anticipatory utility that self 1 feels after such a revision of his brackets (and as a consequence also of his goals) exactly mirrors the anticipatory utility that he feels when he solely revises his goals. The bracket revision problem therefore boils down to the goal adjustment problem under narrow bracketing discussed above. Thus, the conditions under which the individual has no incentive to revise his brackets are the same as the ones that guarantee that he has no incentive to revise his goals. In other words, neither goal revision nor switching from narrow to broad bracketing constitute a problem for goals that satisfy the conditions for minimal and maximal narrow-bracketing implementable goals discussed in the previous section.

### *4.3. Taxi drivers*

Camerer et al. (1997) find a negative wage elasticity of hours worked for taxi drivers and propose that drivers pursue daily income targets which make them work less on days when earnings per hour are high. This triggered a lively debate about the extent to which the labor supply data support reference dependence (summarized in DellaVigna, 2009). Yet, a theoretical foundation of narrow income targets is missing. Contributions to that literature all exogenously impose a narrow evaluation horizon to derive the empirical implications of daily income targets.[25] However, it seems puzzling why taxi drivers, and others free to choose their working hours, adopt narrow daily targets. As emphasized in the literature, broader targets may allow drivers to increase earnings and leisure by working more on high wage days and less on low wage days.

Our model extension from section 4.2 helps close this gap in the literature by considering a present-biased taxi driver with a two-day horizon. He decides whether to evaluate goals for the number of hours to drive on each day narrowly (on a daily basis) or broadly (for the two days combined). We assume that goals are conditional on the wage rate on a given day. Each morning, the driver observes the current day's wage rate $w$ and then chooses how long to drive $x \in \mathbb{R}_0^+$. With probability $\pi$ it is a high-wage day ($w = w_H$) and with probability $1 - \pi$ it is a low-wage day ($w = w_L < w_H$). In line with the literature, we assume that the wage is uncorrelated across days. The driver has the goal to work $x_H$ hours on a high-wage day and $x_L$ hours on a low-wage day. He evaluates outcomes against the corresponding reference distribution, which adjusts to the observed wage and actually chosen working hours (cf. section 4.3). To capture in a stylized way

[25] For example, Kőszegi and Rabin (2006) show that their model can generate a negative wage elasticity of labor supply with daily targets, and Crawford and Meng (2011) show empirically that this is a useful model of taxi drivers' labor supply for the data of Farber (2005).

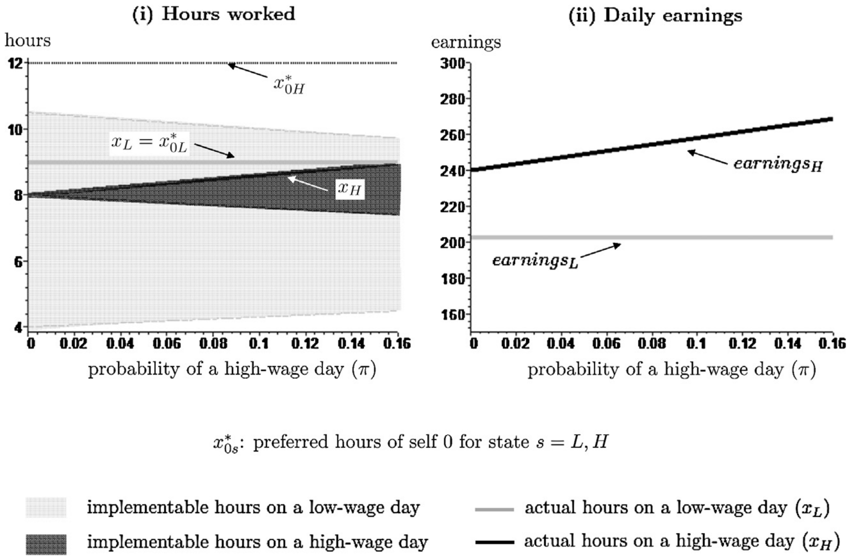


Fig. 2. Illustration of hours worked and earnings in the taxi driver model.

patterns in the data of Farber (2005), we let $w_H = \$30$ and $w_L = \$22.5$.[26] We get the following result (proved in Appendix C):

**Proposition 6.** *In the parametric taxi driver model:*

1. *There is a negative wage elasticity of labor supply, i.e., a driver works fewer hours on high-wage days than on low-wage days ($x_H < x_L$).*
2. *It is optimal for a taxi driver to adopt narrow, single-day goals.*

The model produces a similar pattern of negative wage elasticity of labor supply as Kőszegi and Rabin (2006) do in a binary-decision model, but at the same time we endogenize the bracket of the mental account. The optimality of narrow bracketing is a corollary of Proposition 5. The problem with broad goals is that the driver knows that with some chance he will make up for shirking today. Specifically, on low-wage days there is motivational slack to make up for a previous short fall in effort. For that reason, broad bracketing leads to decision substitution and makes it impossible to implement the same range of decisions as under narrow bracketing.

Fig. 2 illustrates that the driver optimally works fewer hours on high-wage days than on low-wage days ($x_H < x_L$), but income on high-wage days still is higher ($x_H\, w_H > x_L\, w_L$). Why does the driver not work longer on high-wage days, to enjoy extra income or more leisure on low-wage days? He is already pushing himself as hard as he can on a high-wage day, given his present-bias. In contrast, on a low-wage day it is easier for the driver to motivate himself, because

[26] Farber (2005, figure 1c) provides a kernel density estimate of the shift average hourly wage. We chose as $w_L$ the mid-point of the 15–30 dollar range on which it puts almost the entire mass for a normal working day. Further, we chose $w_H$ to capture that with a small probability the wage falls into a long tail with little mass above 30 dollars. Further, to produce Fig. 2, we set $\eta = 1$, $\lambda = 2.5$, $\beta = \gamma = 0.5$ and quadratic costs such that in the absence of uncertainty it would be optimal for self 0 to work 12 hours for $w_H$.

it feels painful to compare the low earnings with what he makes on one of those high-wage days. Working longer on low-wage days helps close the painful gap between the earnings on high- and low-wage days. As a result, the maximal implementable goal is higher on low-wage than on high-wage days.

### *4.4. Bundling strategies that unhinge consumers' self-regulation*

We now present an application of our model to self-control problems related to food choice. Such self-control problems are considered to be a major cause of the overweight problems witnessed in many countries. In a recent survey, Chandon and Wansink (2011) outline how food marketing contributes to overeating. One such marketing strategy is bundling. When healthy and unhealthy foods are presented together, as menus often do, people are more likely to express a positive preference for tempting items (cf., Fishbach and Zhang, 2008). This suggests that bundling induces a broadly evaluated mental account. That is, the consumer balances the adverse effects of eating unhealthy food against the benefits from consuming healthy food in a menu.[27]

It is easy to adapt our framework to situations where the bracketing of a mental account can be exogenously influenced. We show that a bundling strategy that induces a broadly bracketed mental account in a consumer may render infeasible the goal of not buying an "undesired" but tempting product, even though such a goal would work if products were sold individually and bracketed narrowly. Further, we show that bundling is profitable if the consumer is very present-biased. That is, those consumers who are particularly prone to self-control problems, like overeating, are induced to do so even more by the firm's marketing strategy.

Consider a consumer who has a demand of at most one unit for each of two products $i = 1, 2$. If nothing is consumed, the individual has consumption utility normalized to zero. Product 1 is a 'vice good'. Consuming one unit of it leads to immediate benefits $b_1$ and to one-period-delayed costs $c_1$. We assume that $b_1 - c_1 < 0$ and $b_1 - \beta c_1 > 0$. That is, self 1 would like to consume the product, while self 0 considers it undesirable. Both selves agree that product 2 is desirable. One unit provides one-period-delayed benefits $b_2$, where we assume that $b_1 + b_2 > c_1$.

We consider a monopolist who sets a price $p_i$ for each individual product $i = 1, 2$. It also can offer a bundle that contains one unit of each product with overall price $p_B$. The cost of producing a unit of product $i$ is $k_i \geq 0$. Specifically, at date 0 the firm advertises prices $p_1$, $p_2$, $p_B$, and self 0 sets goals for his purchasing decision. We assume that the advertising strategy influences how the consumer brackets his mental accounts. Self 0 forms a broad bracket if products are sold as a bundle, and narrow brackets if products are sold separately. Self 1 decides on purchases, experiences consumption utility from prices paid and from any immediate benefits from purchasing and consuming products. Self 2 experiences consumption utility from any one-period delayed benefits and costs from products purchased and consumed by self 1. After all outcomes have realized, the consumer evaluates outcomes against expectations.

*Selling products separately* Suppose first the monopolist sells products separately so that they are bracketed narrowly by the consumer. Self 0 prefers not to consume product 1. A goal of no

[27] Peoples' estimates of a meal's calories are sometimes even lowered when unhealthy food items are displayed next to healthy foods. This indicates that people do not only have a broad bracket for a food menu, but that they are prone to additional biases (e.g., they evaluate items in a broad bracket and calculate the average calories). Such additional biases would strengthen our insights.

consumption is implementable if the product is not too cheap, i.e., if $p_1 \geq \underline{p}_1$ defined by (for derivations of such a lower price bound cf. Kőszegi and Rabin, 2006):

$$\underline{p}_1 = b_1 \frac{1+\eta}{1+\eta\lambda} - \beta c_1. \tag{9}$$

To illustrate how bundling undermines self-regulation, we consider the case where successful self-regulation is possible if the monopolist sells products separately. That is, we assume $\underline{p}_1 < 0$. This guarantees that, if self 0 sets the goal of not consuming the vice good, there exists no positive price at which the monopolist can attract self 1 to buy the product. Conversely, it is always in the interest of self 0 that his future self consumes product 2 as long as its price does not exceed $b_2$. Given a goal of consuming the product, self 1 has no incentive to deviate if $p_2 \leq \bar{p}_2$, defined by:

$$\bar{p}_2 = \beta b_2 \frac{1+\eta\lambda}{1+\eta}. \tag{10}$$

To extract the highest surplus from the consumer, the monopolist sets $p_2 = \min\{b_2, \bar{p}_2\}$ and some $p_1 \geq p_2$ (such that, given the goal of buying product 2, self 1 has no incentive to buy product 1 instead). Hence, the profit of the monopolist from selling products separately is

$$\Pi_S = \min\{b_2, \bar{p}_2\} - k_2. \tag{11}$$

*Bundling products* When offered a bundle of products 1 and 2, this induces a broad bracket in the consumer. If self 0 sets the goal of purchasing the bundle, he knows that he will evaluate all actual spending at date 1 against the expectation $p_B$, the realized benefits against the expectation $b_1 + b_2$, and the realized costs against the expectation $-c_1$ for date 2. The highest price self 0 is willing to pay for the bundle is $b_1 + b_2 - c_1 > 0$. Self 1 indeed buys the bundle if $p_B \leq \bar{p}_B$ defined by

$$\bar{p}_B = b_1 \frac{1+\beta\eta\lambda}{1+\eta} + \beta b_2 \frac{1+\eta\lambda}{1+\eta} - \beta c_1. \tag{12}$$

If the monopolist charges $p_B = \min\{b_1 + b_2 - c_1, \bar{p}_B\}$, and sets $p_1$ and $p_2$ high enough, self 0 expects to buy the bundle (rather than just a single product or nothing at all), and self 1 indeed buys the bundle. The profit of the monopolist from selling the products as a bundle is:

$$\Pi_B = \min\{b_1 + b_2 - c_1, \bar{p}_B\} - k_2 - k_1. \tag{13}$$

*Can bundling be optimal?* Comparing equations (11) and (13) shows that whenever the monopolist can extract the full surplus for the ex ante desirable good 2 ($b_2$) it is never optimal for him to bundle the products. But suppose this is not possible, because the consumer's present bias makes the willingness to pay for the ex-ante desirable good (good 2) too low for self 1: $\min\{b_2, \bar{p}_2\} = \bar{p}_2$ (one can show that the assumption $\underline{p}_1 < 0$ allows this condition to hold for a broad range of parameters). If in addition $\bar{p}_2 < \min\{b_1 + b_2 - c_1, \bar{p}_B\}$ holds, bundling can be optimal for the monopolist. Specifically, if $\beta$ is very low, then $\bar{p}_2$ is very small and can drop below $b_1 + b_2 - c_1$. Intuitively, a consumer with a severe present bias (low $\beta$) has a very low willingness to pay for the ex-ante desirable good 2 on its own. Bundling both products together increases the willingness to pay of self 1, because he now in addition gets the, from his perspective desirable, vice good (good 1). If products were sold separately, self 0 could prevent his future self from buying the vice good. Bundling unhinges self-control, allowing the monopolist to sell a product that is not desirable for self 0. Think of a fast food restaurant. If products are purchased separately, health concerned customers try to abstain from eating burgers, and the profit from selling water and carrot sticks might be very small. By bundling these items with the burger, these customers may be made to buy the burger as well, and thereby bundling increases profits. In Appendix D, we show that the result is robust to introducing Bertrand competition between $n \geq 2$ firms.

**Proposition 7.** *Suppose* $\underline{p}_1 < 0$ *(defined by (9)) and the unit costs of producing goods 1 and 2 satisfy* $k_1 + k_2 < \min\{b_1 + b_2 - c_1, \bar{p}_B\}$*, where* $\bar{p}_B$ *is defined by (12).*

1. *Suppose* $\min\{b_2, \bar{p}_2\} = b_2$*, where* $\bar{p}_2$ *is defined by (10). Then it never pays off to bundle products. The monopolist sets* $p_2 = b_2$*, and the consumer buys the ex-ante desirable good 2 only.*
2. *Suppose* $\min\{b_2, \bar{p}_2\} = \bar{p}_2$*. Then there exists a* $\check{\beta} \in (0, 1)$ *such that for all* $\beta < \check{\beta}$ *the monopolist sells the bundle including the ex-ante undesirable vice good and the consumer buys it. For all* $\beta \geq \check{\beta}$ *the monopolist sells the ex-ante desirable good 2 only.*
3. *Suppose* $k_1 < b_1 \frac{1+\beta\,\eta\,\lambda}{1+\eta} - \beta\, c_1$ *and* $k_2 < \beta\, b_2 \frac{1+\lambda\,\eta}{1+\eta}$*. Then the strategy of selling the bundle at price* $p_B = k_1 + k_2$ *and offering the individual products at* $p_i > k_i$ *survives competitive pressure. The consumer buys the bundle.*

## 5. Conclusion

Our theory provides a foundation of bracketing under mental accounting by explaining how self-control depends on the brackets in which a person evaluates goal related outcomes. We believe our framework can be usefully applied to a number of phenomena related to mental accounting. Our paper considers two applications. The first illustrates how our theory helps close an important gap in the literature that studies the labor supply of people who can choose their working hours, such as taxi drivers. All of this literature exogenously imposes a daily evaluation horizon and has been concerned with deriving negative wage elasticities. We show that it is indeed optimal for present-biased taxi drivers to set daily income targets. Our second application illustrates the broader applicability of our framework. Taking up evidence that firms' marketing strategies influence how consumers bracket decisions, we show how firms can unhinge consumers' self-regulation by bundling products together, such as is common in fast food menus. We conclude by outlining a few other concrete examples along these lines.

*Externally induced evaluation frames* Consistent with our theory, a number of studies find motivational benefits from externally inducing narrow brackets. In an online experiment by Koch and Nafziger (2016), participants were informed of an opportunity to count the number of zeros in tables from Monday to Friday in the following week. They could earn about 0.1 USD per table for up to 1000 tables, with the amount being transferred to their bank account two to six weeks later. Half of the participants were randomly assigned to set a broad goal for the week and the other half to set an individual goal for each weekday. Each weekday at midnight, participants received an email reminder of their goal for the day or week. Those in the daily goals treatment counted about 33 percent more tables than those in the weekly goal treatment.

In a field experiment with Indian laborers, Soman and Cheema (2011) earmarked a proportion of the wages handed out as savings according to an agreed-on goal, with some receiving the earmarked money in one envelope, and others receiving it in two separate envelopes. Actual savings were significantly higher for those who received separate envelopes. In a lab experiment by Cheema and Soman (2008), partitioning resources into more envelopes reduced the amount that those participants who viewed gambling as an undesirable activity spent on betting. Bundling also lowered ex ante desirable consumption in Soman and Gourville (2001), where multi-performance ticket holders were more likely to forgo a theatrical performance than single-performance ticket holders. Similarly, Gourville and Soman (2002) report more regular gym attendance for members who pay monthly compared to those who pay the same annualized membership fee at a lower frequency.

*Professional golfers* Pope and Schweitzer (2011) show that, compared to otherwise identical shots, professional golfers are about 2 percentage points more likely to score shots that put them at risk of performing worse than a normatively irrelevant[28] goal to score no more than "par" (the expected number of strokes a professional golfer needs to complete a given hole). Pope and Schweitzer (2011, p. 130) hypothesize that "players invest more focus when putting for par to avoid encoding a loss". While different in nature from a self-control problem, the mechanisms underlying such an intra-personal "focusing" problem are similar to the ones underlying an intra-personal motivation/self-control problem. The trade-offs described in our model fit well with a par goal as opposed to a narrower per-stroke goal, a broader 18-hole goal, or a hole-specific goal different from par. For high success probabilities our model predicts narrowly bracketed mental accounts, whereas it predicts a broadly bracketed account when there is substantial risk of missing a reference-point attached to a more narrowly defined outcome, such as with a simple reference point of "scoring".[29] The probability of missing a for-par goal is only around 17 percent. Having a goal of one stroke below par would increase the probability of missing the target to around 73 percent. A broad goal for all 18 holes of a tournament round has limited motivational power, because only after a few holes will a player ever get close to missing the goal, and only then be able to motivate extra focus on a shot. Pope and Schweitzer argue that golfers would have earned an additional $640,000 (18 percent) if they hit each of their one-stroke-under-par (birdie) shots as accurately as they hit otherwise similar for-par shots. Our model suggests that the appropriate counterfactual is how much golfers would have earned if they hit each of their *for-par* shots as *inaccurately* as otherwise similar birdie shots. With this interpretation, narrowly evaluated goals helped golfers increase their earnings by $468,293.[30]

*Other applications* We think that our theory can be applied in future research to better understand how policy makers and firms choose category or product attributes to influence decisions of individuals. Strategies to reduce overeating are an example. Chandon (2013) surveys studies showing that portion or package sizes influence how much people eat. Even simple nudges, such as coloring every $x$th chip in a tube package in red, considerably reduce consumption. An extended version of our model, where package sizes influence reference points and brackets, can explain such behavior and provide further insights for tools to fight obesity. Another example are labeling effects. By naming in-kind benefits "housing benefits" or "child benefits" policy makers

[28] Tournament placement depends on the relative number of strokes over 18 holes in four rounds.

[29] The probability of missing is around 91 percent for two strokes below par, 82 percent for one stroke below par, and 24 percent for par according to Pope and Schweitzer's table 1 (the last number deducts the 2-percentage-points "extra-focus" effect from for-par goals). Their figure 2 shows that the probability of missing for-par shots exceeds 50 percent for distances that are more than 75 inches from the target.

[30] According to Pope and Schweitzer, the average player attempts 45.1 birdie shots and misses 73.1 percent of them, which gives 33 attempted for-par shots. Hence, the effect is $33/45.1 \times \$640{,}000 = \$468{,}293$.

may be evoking narrowly bracketed mental accounts for these categories and thereby influence self-control. For example, Bertrand et al. (2006) discuss that bank accounts could be designed specifically to conform to people's mental-accounting schemes. Further applications are performance evaluations of firms.[31] The frequency of such evaluations can influence the evaluation brackets employees adopt and thereby their motivation. The risk effects we outlined may explain why such evaluations often are conducted only on a semi-annual or annual basis and not more frequently. Related is the question on how to evaluate school children or students. Is it better to have frequent tests that give narrow feedback, or is it be better to aggregate feedback in the form of an overall grade for a subject, or even average grades over several subjects?

## Appendix A. Additional results for the general case

The paper focuses on the case where $b_L(x_i) = 0$ for all $x_i \in \mathbb{R}_0^+$. We now state the preferred decisions as well as the maximal and minimal implementable goals for each type of bracket for the general case. The following assumption, also used in Herweg et al. (2010), guarantees that the preferred decision of self 0 is positive and strictly increasing in $\pi$:

**Assumption 1** *(No dominance of gain–loss utility).* $1 - \eta\,(\lambda - 1)\,(1 - \pi) > 0$.

### *A.1. The preferred decisions of self 0*

*Narrow bracketing* The utility of self 0 for task $i \in \{1, 2\}$ from decision $x_i$ is given by

$$\beta \left\{ b_L(x_i) + \pi \left[1 - \eta\,(\lambda - 1)\,(1 - \pi)\right] [b_H(x_i) - b_L(x_i)] - c(x_i) \right\}. \tag{14}$$

Hence, the preferred decisions of self 0, $x_0^N$, is characterized by the first-order condition

$$b_L'(x_0^N) + \pi \left[1 - \eta\,(\lambda - 1)\,(1 - \pi)\right] \left[b_H'(x_0^N) - b_L'(x_0^N)\right] = c'(x_0^N). \tag{15}$$

Assumption 1 and $b_H'(x_i) > b_L'(x_i)$ imply that $x_0^N > 0$ and that $x_0^N$ is strictly increasing in $\pi$. In addition, $b_H''(x_i) - b_L''(x_i) \le 0$ implies that the second-order condition is satisfied.

*Broad bracketing* The utility of self 0 from decisions $(x_1, x_2)$ is given by

$$\begin{aligned} \beta \Bigg\{ & \sum_{i=1}^{2} \Big[ b_L(x_i) + \pi \left[1 - \eta\,(\lambda - 1)\,(1 - \pi)\,(1 - \pi\,(1 - \pi))\right] [b_H(x_i) - b_L(x_i)] - c(x_i) \Big] \\ & + \eta\,\pi^2\,(1 - \pi)^2 \Big[ \mu\,(b_H(x_2) - b_H(x_1) + b_L(x_1) - b_L(x_2)) \\ & + \mu\,(b_H(x_1) - b_H(x_2) + b_L(x_2) - b_L(x_1)) \Big] \Bigg\}. \end{aligned} \tag{16}$$

We argue that the preferred decisions of self 0 under broad bracketing are symmetric. Note that for $x_1 \neq x_2$, $\mu(b_H(x_2) - b_H(x_1) + b_L(x_1) - b_L(x_2))$ is a gain if and only if $\mu(b_H(x_1) - b_H(x_2) + b_L(x_2) - b_L(x_1))$ is a loss. Asymmetric goals therefore reduce utility, because the last line of

[31] As pointed out by Brocas et al. (2004), one can transfer many of the insights from how people deal with *intra*-personal conflicts of interest (arising from time-inconsistent preferences) to principal-agent settings with *inter*-personal conflicts of interest (arising, e.g., from moral hazard and limits of contracting).

equation (16) is negative for $x_1 \neq x_2$ and equal to zero for $x_1 = x_2$. Hence, $x_0^B$ is characterized by the following first-order condition

$$b'_L(x_0^B) + \pi\left[1 - \eta\,(\lambda-1)\,(1-\pi)\,(1-\pi\,(1-\pi))\right]\left[b'_H(x_0^B) - b'_L(x_0^B)\right] = c'(x_0^B). \quad (17)$$

Assumption 1 and $b'_H(x_i) > b'_L(x_i)$ imply that $x_0^B > 0$ and $x_0^B$ is strictly increasing in $\pi$. In addition, $b''_H(x_i) - b''_L(x_i) \leq 0$ implies that the second-order conditions are satisfied.

### A.2. Derivation of implementable goals under narrow bracketing

**Lemma 2.** *The maximal narrow-bracketing implementable goal for an investment activity (immediate costs, delayed benefits) is defined by*

$$\frac{\beta}{1+\eta}\left\{\pi\,[1+\eta\,(1+\pi\,(\lambda-1))]\,b'_H(x^N_{max}) + (1-\pi)\,(1+\eta\,\lambda)\,b'_L(x^N_{max})\right\} = c'(x^N_{max}). \quad (18)$$

*The minimal narrow-bracketing implementable goal is defined by*

$$\frac{\beta}{1+\eta\,\lambda}\left\{\pi\,(1+\eta)\,b'_H(x^N_{min}) + (1-\pi)\left[1+\eta\,(1+\pi\,(\lambda-1))\right]b'_L(x^N_{min})\right\} = c'(x^N_{min}). \quad (19)$$

*Furthermore, $x^N_{max} > x^N_0$ for $\beta = 1$.*

**Proof.** We only present derivations for $x^N_{max}$, those for $x^N_{min}$ are analogous. The consumption utility of self 1 in task $i \in \{1, 2\}$, given decision $x_i$, is $\beta\,\{\pi\,b_H(x_i) + (1-\pi)\,b_L(x_i)\} - c(x_i)$. For a deviation to $x_i < \hat{x}_i$ the gain–loss utility in the cost dimension is $\eta\,(c(\hat{x}_i) - c(x_i))$. In the benefit dimension the gain–loss utility for the case $b_H(x_i) \geq b_L(\hat{x}_i)$ is

$$\beta\,\eta\,\{\pi\,[1+\pi\,(\lambda-1)]\,b_H(x_i) + \lambda\,(1-\pi)\,b_L(x_i) - \lambda\,\pi\,b_H(\hat{x}_i) \\ - (1-\pi)\,[\pi + \lambda\,(1-\pi)]\,b_L(\hat{x}_i)\};$$

and for the case $b_H(x_i) < b_L(\hat{x}_i)$ it is

$$\beta\,\eta\,\left\{\lambda\,\pi\,b_H(x_i) + \lambda\,(1-\pi)\,b_L(x_i) - \lambda\,\pi\,b_H(\hat{x}_i) - \lambda\,(1-\pi)\,b_L(\hat{x}_i)\right\}.$$

Adding up consumption and gain–loss utility and differentiating, yields the maximal implementable goal. We go through the case distinctions. Self 1 has no incentive to deviate and lower his decision level such that $b_H(x_i) \geq b_L(\hat{x}_i)$, as long as the goal does not exceed $x^N_{max}$, defined by[32]

$$\beta\left\{\pi\,[1+\eta\,(1+\pi(\lambda-1))]\,b'_H(x^N_{max}) + (1-\pi)\,(1+\eta\,\lambda)\,b'_L(x^N_{max})\right\} = (1+\eta)\,c'(x^N_{max}).$$

Self 1 has no incentive to deviate and lower his decision level such that $b_H(x_i) < b_L(\hat{x}_i)$, as long as the goal does not exceed $x^N_{max,2}$, defined by

$$\beta\left\{\pi\,(1+\eta\,\lambda)\,b'_H(x^N_{max,2}) + (1-\pi)\,(1+\eta\,\lambda)\,b'_L(x^N_{max,2})\right\} = (1+\eta)\,c'(x^N_{max,2}).$$

[32] Note that the left-hand-side is decreasing and the right-hand side is increasing in $x_i$.

Because $x^N_{max,2} \geq x^N_{max}$, the first condition pins down the maximal implementable goal. □

While for riskless activities we always have $x^N_{min} < x^N_0$, for $\pi \in (0, 1)$ it is possible that self 0 prefers a goal that is lower than the minimal implementable goal.[33] Replacing Assumption 1 with the following stronger condition rules out such corner solutions that are uninteresting for our purpose of studying time inconsistency arising from a present bias, where self 0 prefers a higher decision than self 1.

**Assumption 1'.** $1 - \eta\,(\lambda - 1)\,(1 - \pi) \geq \frac{\beta\,(1+\eta)}{1+\beta\,\eta\,\lambda}$.

To see this, rewrite equation (15) in the form $c'(x^N_0) = \rho^N_L\, b'_L(x^N_0) + \rho^N_H\, b'_H(x^N_0)$ and equation (19) in the form $c'(x^N_{min}) = \kappa^N_L(\beta)\, b'_L(x^N_{min}) + \kappa^N_H(\beta)\, b'_H(x^N_{min})$. The $\kappa^N_j$-terms, $j \in \{L, H\}$, are increasing in $\beta$, so it is enough to show that $\rho^N_j > \kappa^N_j(\beta = 1)$. Assumption 1' directly gives $\rho^N_H > \kappa^N_H(1)$ and $\rho^N_L - \kappa^N_L(1) = \eta\,(1 - \pi)\,(\lambda - 1)\,(1 + \eta\,\lambda\,\pi)/(1 + \eta\,\lambda) > 0$. Hence, $x^N_{min} < x^N_0$.

From this one obtains the following result that extends previous work on goals as reference points in single-task settings (Carrillo and Dewatripont, 2008; Suvorov and van de Ven, 2008; Koch and Nafziger, 2011b; Hsiaw, 2013) to a decision with continuous action space in a stochastic environment were the individual has stochastic reference points:

**Proposition 8.** *Suppose Assumption 1' holds and consider an investment activity (immediate costs, delayed benefits).*

1. *There exists a continuum of narrow-bracketing implementable goals $[x^N_{min}, x^N_{max}]$. Goals offer commitment to exceed the preferred decision of self 1, $x^N_1$, because $x^N_1 < x^N_{max}$. Moreover, the preferred decision of self 0, $x^N_0$, satisfies $x^N_0 > x^N_{min}$.*
2. *There exists a cut-off $\tilde{\beta} \in (0, 1)$ such that $x^N_0 \leq x^N_{max}$ for $\beta \geq \tilde{\beta}$, i.e., the individual can fully overcome his self-control problem. Self 0 chooses as his goal $\hat{x} = \min\{x^N_0, x^N_{max}\}$.*

**Proof.** To see that $x^N_1 < x^N_{max}$, compare $x^N_{max}$ defined in equation (18) with $x^N_1$ defined by:

$$\beta \left\{ b'_L(x^N_1) + \pi\, \left[1 - \eta\,(\lambda - 1)\,(1 - \pi)\right] \left[ b'_H(x^N_1) - b'_L(x^N_1) \right] \right\} = c'(x^N_1). \tag{20}$$

Now, subtracting the respective left-hand-sides for fixed decision level $x$, one obtains

$$\begin{aligned} & LHS(18) - LHS(20) \\ &= \frac{\beta\,\eta\,(\lambda - 1)}{1 + \eta} \left[\pi\,(1 + \eta\,(1 - \pi))\, b'_H(x) + (1 - \pi)\,(1 - \pi\,(1 + \eta))\, b'_L(x)\right] \\ &\geq \frac{\beta\,\eta\,(\lambda - 1)}{1 + \eta}\,[1 - \pi\,(1 - \pi)]\, b'_L(x) > 0. \end{aligned}$$

[33] The reason is that a loss-averse individual with stochastic reference points is more risk averse when a decision is committed to well in advance than when the decision is made for fixed expectations (Kőszegi and Rabin, 2007). Specifically, self 1 takes expectations as given in a *preferred personal equilibrium*, whereas the preferred decision of self 0 accounts for the externality that a committed choice has on expectations ($x^N_0$ is a *choice-acclimating personal equilibrium*).

The existence of a unique cut-off $\tilde{\beta} \in (0, 1)$ follows from the intermediate value theorem because $x^N_{max}$ is strictly increasing in $\beta$ and it is straightforward that for $\beta = 1$ we have $x^N_{max} > x^N_0$ and for $\beta \to 0$ we have $x^N_{max} < x^N_0$. The remainder follows from Lemma 2. □

The larger the present bias parameter $\beta$, the more the selves agree on perceived delayed benefits. And the more they agree, the more likely it is that self 0 can implement his preferred decision. This result confirms Loewenstein (1999) conjecture that goal setting can explain why people with time-inconsistent preferences often behave in a way that makes them indistinguishable from time-consistent agents. Self-regulation however is constrained if the individual faces a more severe self-control problem so that $x^N_{\max} < x^N_0$. The best self 0 can do in such a case is to set as goal the maximal implementable goal. Nevertheless, self 0 still can nudge his future self to a more ambitious decision than self 1 would want on his own, because $x^N_{\max}$ exceeds the preferred decision of self 1, $x^N_1$.

### *A.3. Derivation of implementable goals under broad bracketing*

Checking that self 1 has no incentive to decrease (increase) his decision in one task, i.e., that

$$U^B_1(\hat{x}_i, \hat{x}_{-i}|\hat{x}_i, \hat{x}_{-i}) \geq U^B_1(x_i, \hat{x}_{-i}|\hat{x}_i, \hat{x}_{-i}), \ i \in \{1, 2\}, \ \forall x_i \in \mathbb{R}^+_0, \tag{21}$$

provides us with an upper and a lower bound on implementable goals, $x^B_{ub}$ and $x^B_{lb}$. Without loss of generality, we consider deviations in task 2 and fix the decision in task 1. The following result shows that the range of potentially implementable goals is larger with symmetric goals ($\hat{x}_1 = \hat{x}_2$) than with increasing goals ($\hat{x}_1 < \hat{x}_2$) or decreasing goals ($\hat{x}_1 > \hat{x}_2$). Together with the fact that self 0 has symmetric preferred decisions, this implies that asymmetric goals are never optimal.

**Lemma 3.** *Under broad bracketing, the upper bound on implementable goals for task $i \in \{1, 2\}$ is greatest with symmetric goals, $\hat{x}_1 = \hat{x}_2$, and satisfies*

$$\begin{aligned}\frac{\beta}{1+\eta}\Bigg\{&\pi\ (1+\eta\,[1+(\lambda-1)\,\pi\,(1+(1-\pi)\,(1-2\pi))])\ b'_H(x^B_{ub}) \\ &+(1-\pi)\left(1+\eta\left[1+(\lambda-1)\,(1-\pi\,(1-\pi)^2)\right]\right) b'_L(x^B_{ub})\Bigg\} = c'(x^B_{ub});\end{aligned} \tag{22}$$

*and the lower bound on implementable goals is lowest with symmetric goals and satisfies*

$$\begin{aligned}\frac{\beta}{1+\eta\lambda}\Bigg\{&\pi\left(1+\eta\left[1+(\lambda-1)\,\pi^2\,(1-\pi)\right]\right) b'_H(x^B_{lb}) \\ &+(1-\pi)\left(1+\eta\left[1+(\lambda-1)\,\pi\,(1+(1-\pi)\,(1-2\pi))\right]\right) b'_L(x^B_{lb})\Bigg\} = c'(x^B_{lb}).\end{aligned} \tag{23}$$

**Proof.** We only derive the upper bound on narrow-bracketing implementable goals because derivations for the minimal narrow-bracketing implementable goal are analogous.

*Utility of self 1* The consumption utility of self 1 from decisions $x_2$ and $x_1 = \hat{x}_1$ is

$$\beta\,\{\pi\, b_H(x_2) + (1-\pi)\, b_L(x_2) + \pi\, b_H(\hat{x}_1) + (1-\pi)\, b_L(\hat{x}_1)\} - c(\hat{x}_1) - c(x_2).$$

The gain–loss utility in the cost dimension for $x_2 < \hat{x}_2$ is $\eta\,(c(\hat{x}_2) - c(x_2))$. The only difference between the cases of symmetric, increasing, and decreasing goals is the gain–loss utility in the benefit dimension after a deviation of self 1 in task 2. To save space, we use $\kappa$ to summarize the constant terms in the gain–loss utility, such as $constant_H \times b_H(\hat{x}_i)$ and $constant_L \times b_L(\hat{x}_i)$; and we drop the proportionality factor $\beta\,\eta$.

*Symmetric goals* ($\hat{x}_1 = \hat{x}_2$) Consider first deviations to $x_2 < \hat{x}_2$, where $\hat{x}_2 - x_2$ is small in the sense that $b_H(x_2) \geq b_L(\hat{x}_2)$ (condition $A$), $b_H(\hat{x}_1) + b_L(x_2) \geq b_L(\hat{x}_1) + b_L(\hat{x}_2)$ (condition $B$), and $b_H(\hat{x}_1) + b_H(x_2) \geq b_L(\hat{x}_1) + b_L(\hat{x}_2)$ (condition $C$). The gain–loss utility in the benefit dimension is proportional to

$$
\begin{aligned}
&(1-\pi)^2 \left\{ \underbrace{\pi^2 b_H(x_2)}_{c} + \underbrace{\pi\,(1-\pi)\,b_H(x_2)}_{a} + \underbrace{\pi\,(1-\pi)\,b_L(x_2)}_{b} + \lambda\,(1-\pi)^2\,b_L(x_2) \right\} \\
&+ 2\,(1-\pi)\,\pi \left\{ \underbrace{\pi^2 b_H(x_2)}_{a} + \overbrace{\lambda\,\pi\,(1-\pi)\,[b_H(x_2) + b_L(x_2)]}^{I} + \lambda\,(1-\pi)^2\,b_L(x_2)] \right\} \\
&+ \lambda\,\pi^2 \left\{ \pi^2 b_H(x_2) + \pi\,(1-\pi)\,[b_H(x_2) + b_L(x_2)] + (1-\pi)^2\,b_L(x_2)] \right\} + \kappa \\
= {}& \pi\,\{1 + (\lambda - 1)\,\pi\,[1 + (1-\pi)\,(1-2\pi)]\}\,b_H(x_2) \\
&+ (1-\pi)\,\left\{1 + (\lambda - 1)\,[1 - \pi\,(1-\pi)^2]\right\}\,b_L(x_2) + \kappa. \qquad (24)
\end{aligned}
$$

The terms underbraced by $a$–$c$ are gains because conditions $A$–$C$ hold. Adding up the gain–loss utility in the benefit and cost dimensions as well as consumption utility, and differentiating yields the definition of $x^B_{ub}$ in equation (22). For larger deviations, self 1 is more likely to suffer a loss compared to the case above. For example, if condition $A$ is violated, the terms underbraced by $a$ in equation (24) are multiplied by $\lambda$, and this case gives a bound on implementable goals that is larger than $x^B_{ub}$. Similar arguments apply for deviations such that conditions $B$ and $C$ fail. Overall, $x^B_{ub}$ therefore is the lowest of these bounds.

*Increasing goals* ($\hat{x}_1 < \hat{x}_2$) *or decreasing goals* ($\hat{x}_1 > \hat{x}_2$) Consider deviations to $x_2 < \hat{x}_2$, where $\hat{x}_2 - x_2$ is small in the sense that conditions $A$ - $C$ from above hold and, in addition, $b_H(\hat{x}_1) + b_H(x_2) > b_L(\hat{x}_1) + b_H(\hat{x}_2)$ (condition $D$), and $b_L(\hat{x}_1) + b_H(x_2) > b_H(\hat{x}_1) + b_L(\hat{x}_2)$ (condition $E$), or $b_H(\hat{x}_1) + b_L(x_2) > b_L(\hat{x}_1) + b_H(\hat{x}_2)$ (condition $F$). Note that at most one of the conditions ($E$) and ($F$) can hold for small deviations, because either $b_L(\hat{x}_1) + b_H(\hat{x}_2)$ is greater, equal, or less than $b_H(\hat{x}_1) + b_L(\hat{x}_2)$. Start from the following expression:

$$
\begin{aligned}
&(1-\pi)^2 \left\{ \underbrace{\pi^2 b_H(x_2)}_{c} + \underbrace{\pi\,(1-\pi)\,b_H(x_2)}_{a} + \underbrace{\pi\,(1-\pi)\,b_L(x_2)}_{b} + \lambda\,(1-\pi)^2\,b_L(x_2) \right\} \\
&+ (1-\pi)\,\pi \left\{ \underbrace{\pi^2 b_H(x_2)}_{a} + \overbrace{\lambda\,\pi\,(1-\pi)\,b_L(x_2) + \underbrace{\pi\,(1-\pi)\,b_H(x_2)}_{e}}^{II} + \lambda\,(1-\pi)^2\,b_L(x_2) \right\} \\
&+ (1-\pi)\,\pi \left\{ \underbrace{\pi^2 b_H(x_2)}_{d} + \overbrace{\underbrace{\pi\,(1-\pi)\,b_L(x_2)}_{f} + \lambda\,\pi\,(1-\pi)\,b_H(x_2)}^{III} + \lambda\,(1-\pi)^2\,b_L(x_2) \right\} \\
&+ \lambda\,\pi^2 \left\{ \pi^2 b_H(x_2) + \pi\,(1-\pi)\,[b_H(x_2) + b_L(x_2)] + (1-\pi)^2\,b_L(x_2) \right\} + \kappa. \qquad (25)
\end{aligned}
$$

Similar to above, some of the gains in the terms underbraced by $a$–$f$ turn into losses if some of the conditions $A$–$F$ fail. Compared to the case with symmetric goals, losses are less likely. To see this, note first that condition $D$ is weaker than condition $A$. Next, compare term $I$ in equation (24) with terms $II$ and $III$ in equation (25). Because conditions $E$ and $F$ cannot both hold, $2\,I \geq II + III$ in all cases. All other terms in equations (24) and (25) coincide. Note that nowhere in the derivations did we use $\hat{x}_2 > \hat{x}_1$ or $\hat{x}_2 < \hat{x}_1$. Hence, $x^B_{ub} \geq x^B_{ub,incr}$ and $x^B_{ub} > x^B_{ub,decr}$. $\square$

## Appendix B. Proofs

### *B.1. Proof of Proposition 1*

The result follows directly from equations (14)–(17). Note that for $\pi \in (0, 1)$ we have $\left[1 - \eta\,(\lambda - 1)\,(1 - \pi)\,\{1 - \pi\,(1 - \pi)\}\right] - \left[1 - \eta\,(\lambda - 1)\,(1 - \pi)\right] = \eta\,(\lambda - 1)\,\pi(1 - \pi) < 0$, and that the last line of equation (16) is equal to zero for $x_1 = x_2$.

### *B.2. Proof of Proposition 2*

*Comment* To facilitate exposition we present Lemma 1 after Proposition 2 in the main text. The order of proofs in the appendix follows the order in the main text. Yet, for a better understanding, the reader is advised to work through the proof of the Lemma 1 before the proof of Proposition 2 because the latter refers to some notation introduced in the lemma.

*Overview* The utility of self 0 under bracket $A \in \{B, N\}$ is strictly increasing in $x$ for decision levels below the preferred decision of self 0, $x^A_0$. Proposition 1 hence implies that narrow brackets can be optimal if they allow to implement a higher decision (i.e., one that is closer to the preferred decision of self 0) than with broad brackets. For $\pi \in (0, \underline{\nu}) \cup (\bar{\nu}, 1)$ we have $x^N_{max} > x^B_{ub} \geq x^B_{max}$ (Lemma 1), while otherwise broad bracketing does better than narrow bracketing (if no profitable joint deviation exists). For $\pi \in \{0, 1\}$, broad and narrow brackets do equally well, because $x^B_{ub} = x^N_{max}$ and risk pooling plays no role. Below we show that, starting from $\pi = 1$ ($\pi = 0$), a marginal decrease (increase) in $\pi$ leads to a larger drop (smaller gain) in utility under broad than under narrow bracketing. That is, for $\pi$ close to zero or one, narrow bracketing yield strictly larger utility than broad bracketing. In showing this, we also rely on the fact that the utility functions are continuous and differentiable – something that we do not mention separately in the proofs.

*Broad bracketing* Note that for $x^B_{ub} < x^B_0$, self 0 implements $x^B_{ub}$ (or possibly an even lower decision). Hence, the maximized utility of self 0 under broad bracketing cannot exceed $U^B_0(x^B_{ub}, x^B_{ub}|x^B_{ub}, x^B_{ub}) = 2\,\beta\,\left\{b_L(x^B_{ub}) + [b_H(x^B_{ub}) - b_L(x^B_{ub})]\,B(\pi) - c(x^B_{ub})\right\}$, where $B(\pi) = \pi\,\left[1 - \eta\,(\lambda - 1)\,(1 - \pi)\,(1 - \pi\,(1 - \pi))\right]$.[34] Our goal is to derive a formula for how this maximized utility of self 0 varies with $\pi$, at $\pi \in \{0, 1\}$, and we use the following interim result.

$$\frac{d\,U^B_0(x^B_{ub}, x^B_{ub}|x^B_{ub}, x^B_{ub})}{d\,\pi} = 2\,\beta\,[b_H(x^B_{ub}) - b_L(x^B_{ub})]\,B'(\pi)$$

[34] $U^B_0(x^B_{ub}, x^B_{ub}|x^B_{ub}, x^B_{ub})$ bounds from above the utility of self 0, because utility is strictly increasing for decision levels $x_i < x^B_0$. For our purpose it does not matter whether or not this bound is tight (considering joint deviations would only reduce the maximized utility by further restricting the set of decisions that self 0 can implement). One can however show that the bounds $x^B_{ub}$ and $x^B_{lb}$ are tight in the special case where $b_L(x) = 0$ for all $x$ (see Corollary 1).

$$+2\beta\left\{[b'_H(x^B_{ub})-b'_L(x^B_{ub})]\,B(\pi)+b'_L(x^B_{ub})-c'(x^B_{ub})\right\}\frac{d\,x^B_{ub}}{d\,\pi},$$

where $B'(\pi)=1-\eta\,(1-2\pi)\;[1-2\pi\,(1-\pi)]\;(\lambda-1)$. Next, we derive $\frac{d\,x^B_{ub}}{d\,\pi}$. Rewriting equation (22) as $\Phi^B(x,\pi)=0$ implicitly defines $x^B_{ub}(\pi)$.[35] Implicit differentiation gives $\frac{d\,x^B_{ub}}{d\,\pi}=-\Phi^B_\pi(x,\pi)/\Phi^B_x(x,\pi)$, where

$$\Phi^B_\pi(x,\pi)=\beta\left\{b'_H(x)\left[1+\eta\,(1+(\lambda-1)\,\pi\,(4-9\pi+8\pi^2))\right]\right.$$
$$\left.-\,b'_L(x)\left[1+\eta\left(1+(\lambda-1)\left\{1+(1-\pi)^2\,(1-4\pi)\right\}\right)\right]\right\},$$
$$\Phi^B_x(x,\pi)=\beta\left\{\pi\,b''_H(x)\left[1+\eta\left[1+(\lambda-1)\,\pi\,(2-3\pi+2\pi^2)\right]\right]\right.$$
$$\left.+\,(1-\pi)\,b''_L(x)\left[1+\eta\left[1+(\lambda-1)\,(1-\pi\,(1-\pi)^2)\right]\right]\right\}-(1+\eta)\,c''(x).$$

To show our result, we need as an interim step $\left.\frac{d\,U^B_0(x^B_{ub},x^B_{ub}|x^B_{ub},x^B_{ub})}{d\,\pi}\right|_{\pi=1}$:

$$\left.\frac{d\,U^B_0(\cdot)}{d\,\pi}\right|_{\pi=1}=\beta\left\{\overbrace{2\,[b_H(x^B_{ub})-b_L(x^B_{ub})]\,\left[1+\eta\,(\lambda-1)\right]}^{\equiv\kappa^B(x^B_{ub})}+\overbrace{2\,[b'_H(x^B_{ub})-c'(x^B_{ub})]\;\left.\frac{d\,x^B_{ub}}{d\,\pi}\right|_{\pi=1}}^{\equiv\psi^B(x^B_{ub})}\right\},\tag{26}$$

where, using our above derivations,

$$\left.\frac{d\,x^B_{ub}}{d\,\pi}\right|_{\pi=1}=\beta\,\frac{b'_H(x^B_{ub})\,\left[1+\eta\,(1+3\,(\lambda-1))\right]-b'_L(x^B_{ub})\;(1+\eta\,\lambda)}{(1+\eta)\,c''(x^B_{ub})-\beta\,(1+\eta\,\lambda)\,b''_H(x^B_{ub})}.$$

And we need

$$\left.\frac{d\,U^B_0(\cdot)}{d\,\pi}\right|_{\pi=0}=\beta\left\{\overbrace{2\,[b_H(x^B_{ub})-b_L(x^B_{ub})]\,\left[1-\eta\,(\lambda-1)\right]}^{\equiv\underline{\kappa}^B(x^B_{ub})}+\overbrace{2\,[b'_L(x^B_{ub})-c'(x^B_{ub})]\;\left.\frac{d\,x^B_{ub}}{d\,\pi}\right|_{\pi=0}}^{\equiv\underline{\psi}^B(x^B_{ub})}\right\},\tag{27}$$

where, using our above derivations,

$$\left.\frac{d\,x^B_{ub}}{d\,\pi}\right|_{\pi=0}=\beta\,\frac{b'_H(x^B_{ub})\,(1+\eta)-b'_L(x^B_{ub})\;(1+\eta\,(1+2\,(\lambda-1)))}{(1+\eta)\,c''(x^B_{ub})-\beta\,(1+\eta\,\lambda)\,b''_L(x^B_{ub})}.$$

[35] Note that $1+(1-\pi)\,(1-2\pi)=2-3\pi+2\pi^2$. The latter representation is easier to work with here.

*Narrow bracketing* It is sufficient to show that $U_0^N(x_{max}^N, x_{max}^N | x_{max}^N, x_{max}^N) > U_0^B(x_{ub}^B, x_{ub}^B | x_{ub}^B, x_{ub}^B)$, because by assumption, at $\pi \in \{0, 1\}$, $x_0^N = x_0^B > x_{max}^N = x_{max}^B$ and hence, locally, $U_0^N(x_{max}^N, x_{max}^N | x_{max}^N, x_{max}^N) > U_0^B(x_{ub}^B, x_{ub}^B | x_{ub}^B, x_{ub}^B)$ implies $U_0^N(x_0^N, x_0^N | x_0^N, x_0^N) > U_0^B(x_{ub}^B, x_{ub}^B | x_{ub}^B, x_{ub}^B)$. Setting $x_i = x_{max}^N$ in equation (14) for both tasks $i = \{1, 2\}$ and differentiating, gives

$$\frac{d\,U_0^N(x_{max}^N, x_{max}^N | x_{max}^N, x_{max}^N)}{d\pi} = 2\beta\,[b_H(x_{max}^N) - b_L(x_{max}^N)]\left[1 - \eta\,(\lambda - 1)\,(1 - 2\pi)\right]$$
$$+\beta\left\{[b'_H(x_{max}^N) - b'_L(x_{max}^N)]\,\pi\,[1 - \eta\,(1 - \pi)\,(\lambda - 1)] + b'_L(x_{max}^N) - c'(x_{max}^N)\right\}\frac{d\,x_{max}^N}{d\pi}.$$

Rewriting equation (18) as $\Phi^N(x, \pi) = 0$ implicitly defines $x_{max}^N(\pi)$. Implicit differentiation gives $\frac{d\,x_{max}^N}{d\pi} = -\Phi_\pi^N(x, \pi)/\Phi_x^N(x, \pi)$, where

$$\Phi_\pi^N(x, \pi) = \beta\left\{b'_H(x)\left[1 + \eta\,(1 + 2\,(\lambda - 1)\,\pi)\right] - b'_L(x)\,(1 + \eta\,\lambda)\right\},$$
$$\Phi_x^N(x, \pi) = \beta\left\{\pi\,b''_H(x)\left[1 + \eta\,(1 + (\lambda - 1)\,\pi)\right] + (1 - \pi)\,b''_L(x)\,(1 + \eta\,\lambda)\right\} - (1 + \eta)\,c''(x).$$

To show our result, we need as an interim step $\frac{d\,U_0^N(x_{max}^N, x_{max}^N | x_{max}^N, x_{max}^N)}{d\pi}\Big|_{\pi=1}$:

$$\left.\frac{d\,U_0^N(\cdot)}{d\pi}\right|_{\pi=1} = 2\beta\left\{\overbrace{[b_H(x_{max}^N) - b_L(x_{max}^N)]\left[1 + \eta\,(\lambda - 1)\right]}^{\equiv\kappa^N(x_{max}^N)} + \overbrace{[b'_H(x_{max}^N) - c'(x_{max}^N)]\left.\frac{d\,x_{\mathrm{max}}^N}{d\pi}\right|_{\pi=1}}^{\equiv\psi^N(x_{max}^N)}\right\}, \tag{28}$$

where, using our above derivations,

$$\left.\frac{d\,x_{\mathrm{max}}^N}{d\pi}\right|_{\pi=1} = \beta\,\frac{b'_H(x_{max}^N)\left[1 + \eta\,(1 + 2\,(\lambda - 1))\right] - b'_L(x_{max}^N)\,(1 + \eta\,\lambda)}{(1 + \eta)\,c''(x_{max}^N) - \beta\,(1 + \eta\,\lambda)\,b''_H(x_{max}^N)}.$$

Similarly, we need

$$\left.\frac{d\,U_0^N(\cdot)}{d\pi}\right|_{\pi=0} = 2\beta\left\{\overbrace{2\,[b_H(x_{max}^N) - b_L(x_{max}^N)]\left[1 - \eta\,(\lambda - 1)\right]}^{\equiv\underline{\kappa}^N(x_{max}^N)} + \overbrace{2\,[b'_L(x_{max}^N) - c'(x_{max}^N)]\left.\frac{d\,x_{max}^N}{d\pi}\right|_{\pi=0}}^{\equiv\underline{\psi}^N(x_{max}^N)}\right\}, \tag{29}$$

where, using our above derivations,

$$\left.\frac{d\,x_{max}^N}{d\pi}\right|_{\pi=0} = \beta\,\frac{b'_H(x_{max}^N)\,(1 + \eta) - b'_L(x_{max}^N)\,(1 + \eta\,\lambda)}{(1 + \eta)\,c''(x_{max}^N) - \beta\,(1 + \eta\,\lambda)\,b''_L(x_{max}^N)}.$$

*Comparison* For $\pi \in \{0, 1\}$, $x^N_{max} = x^B_{ub} \equiv x_{max}$ and we have $U^N_0(x_{max}, x_{max}|x_{max}, x_{max}) = U^B_0(x_{max}, x_{max}|x_{max}, x_{max})$. Introducing a little bit of uncertainty, narrow bracketing thus does strictly better than broad bracketing if at $\pi = 1$ ($\pi = 0$) utility drops by less (increases by more) under narrow bracketing. So for $\pi = 1$ we need to show that

$$\left.\frac{d\, U^N_0(x^N_{max}, x_{max}|x_{max}, x_{max})}{d\pi}\right|_{\pi=1} < \left.\frac{d\, U^B_0(x^B_{ub}, x^B_{ub}|x^B_{ub}, x^B_{ub})}{d\pi}\right|_{\pi=1}. \tag{30}$$

Comparing equations (26) and (28), note that $\kappa^N(x_{max}) = \kappa^B(x_{max})$ and that all terms in $\psi^B(x_{max})$ and $\psi^N(x_{max})$ coincide, except for $\eta\,(1 + 3\,(\lambda - 1))$ in $\psi^B(x_{max})$ and $\eta\,(1 + 2\,(\lambda - 1))$ in $\psi^N(x_{max})$. Note further that at $\pi = 1$, $x^N_{max} = x^B_{ub} \equiv x_{max} < x^B_0 = x^N_0$, and therefore $b'_H(x_{max}) - c'(x_{max}) > 0$. Finally, $\eta\,(1 + 3\,(\lambda - 1)) > \eta\,(1 + 2\,(\lambda - 1))$ implies that inequality (30) holds. Similarly, for $\pi = 0$ we need to show that

$$\left.\frac{d\, U^N_0(x^N_{max}, x^N_{max}|x^N_{max}, x^N_{max})}{d\pi}\right|_{\pi=0} > \left.\frac{d\, U^B_0(x^B_{ub}, x^B_{ub}|x^B_{ub}, x^B_{ub})}{d\pi}\right|_{\pi=0}. \tag{31}$$

All terms in equations (27) and (29) coincide, except for $-[1 + \eta\,(1 + 2\,(\lambda - 1))]$ in $\underline{\psi}^B(x_{max})$ and $-[1 + \eta\,\lambda]$ in $\underline{\psi}^N(x_{max})$, which implies that inequality (31) holds.

Finally, either $U^N_0(x^N_{max}, x^N_{max}|x^N_{max}, x^N_{max}) > U^B_0(x^B_{ub}, x^B_{ub}|x^B_{ub}, x^B_{ub})$ for all $\pi \in (0, 1)$, in which case $\underline{\pi} = \bar{\pi} = 0$, or – because $U^N_0(\cdot)$ and $U^B_0(\cdot)$, and hence their difference, are continuous in $\pi$ – the intermediate value theorem guarantees the existence of interior cutoffs $\underline{\pi} \le \bar{\pi}$. Because $x^N_{max} > x^B_{ub}$ is a necessary condition for narrow bracketing to be optimal, $\underline{\nu}$ and $\bar{\nu}$ defined by Lemma 1 bound the cutoffs: $\bar{\pi} \in [\bar{\nu}, 1)$ and $\underline{\pi} \in [0, \underline{\nu})$, such that for $\pi \in (0, \underline{\pi}) \cup (\bar{\pi}, 1)$ narrow brackets are strictly optimal.

### *B.3. Proof of Lemma 1*

We want to compare the maximal implementable goal under broad bracketing with the one under narrow bracketing. To do so we set $x^N_{max} = x^B_{ub} = x_i$ and subtract the left-hand sides (note that the right-hand sides of the two equations are identical) of equations (18) and (22):

$$LHS(18) - LHS(22) = \beta\,\eta\,(\lambda - 1)\,\pi\,(1 - \pi)\,[A(\pi)\,b'_H(x_i) + B(\pi)\,b'_L(x_i)], \tag{32}$$

where $A(\pi) \equiv \pi\,(2\pi - 1)$ and $B(\pi) \equiv (1 - \pi)^2$. For the special case where $b_L(x_i) = 0$ for all $x_i \in \mathbb{R}^+_0$, we get $\bar{\nu} = 1/2$ because $A(\pi) > 0$ if and only if $\pi \in (1/2, 1)$ and $B(\pi) = 0$. If $b'_L(x_i) > 0$ we have $A(\pi)\,b'_H(x_i) + B(\pi)\,b'_L(x_i) > 0$ if and only if $\frac{b'_L(x_i)}{b'_H(x_i)} > -\frac{A(\pi)}{B(\pi)} = \frac{\pi\,(1-2\,\pi)}{1-\pi}^2 \equiv r(\pi)$. Note that our assumption $b''_H(x_i) \le b''_L(x_i) \le 0$ implies that $\frac{b'_L(x_i)}{b'_H(x_i)}$ is weakly decreasing in $x_i$.[36] So if equation (32) has a positive sign at $x_i = x^N_{max}$ then it also has a positive sign for all $x_i < x^N_{max}$, and it implies that $x^N_{max} > x^B_{ub} \ge x^B_{max}$. The expression $r(\pi)$ describes a parabola with maximum 1/4 at $\pi = 1/3$, which yields the sufficient condition $\alpha > 1/4$ for a

36

$$\frac{d}{d\,x_i}\frac{b'_L(x_i)}{b'_H(x_i)} = \frac{b''_L(x_i)b'_H(x_i) - b'_L(x_i)b''_H(x_i)}{(b'_H(x_i))^2} \le \frac{b''_L(x_i)[b'_H(x_i) - b'_L(x_i)]}{(b'_H(x_i))^2} \le 0.$$

model where $b_L(x_i) = \alpha\, b_H(x_i)$. Now either $\bar{\nu} = 0$ or, because $A(\pi)\, b'_H(x^N_{max}) + B(\pi)\, b'_L(x^N_{max})$ is continuous in $\pi$, the intermediate value theorem guarantees that there exists an interior cutoff $\bar{\nu} \in [\frac{1}{3}, \frac{1}{2}]$ such that $A(\pi)\, b'_H(x^N_{max}) + B(\pi)\, b'_L(x^N_{max}) > 0$ for all $\pi \in (\bar{\nu}, 1)$. So if $b'_L(x^N_{max}) > 0$ and $\bar{\pi}$ is interior, then $\bar{\nu} \in (\frac{1}{3}, \frac{1}{2})$ and there exists another cutoff $\underline{\nu} \in (0, \frac{1}{3})$ such that $A(\pi)\, b'_H(x^N_{max}) + B(\pi)\, b'_L(x^N_{max}) > 0$ for $\pi \in (0, \underline{\nu})$.

### B.4. *Proof of Proposition 3 and further details on its discussion*

Distinguishing $\pi_1$ and $\pi_2$ in the proofs of Lemmas 2 and 3, we obtain the maximal narrow-bracketing implementable goal and an upper bound on the implementable broad-bracketing goal, respectively:

$$\frac{\beta}{1+\eta}\pi_1\{1+\eta\,[1+(\lambda-1)\,\pi_1]\}\, b'(x^N_{max}) = c'(x^N_{max}),$$

$$\frac{\beta}{1+\eta}\pi_1\{1+\eta\,[1+(\lambda-1)\,[\pi_1-\pi_1\,\pi_2\,(1-\pi_2)+(1-\pi_1)\,\pi_2\,(1-\pi_2)]]\}\, b'(x^B_{ub}) = c'(x^B_{ub}).$$

Note that $x^N_{max} > x^B_{ub}$ for $\pi_1 > \frac{1}{2}$, and that for $\pi_1 = 1$, $x^N_{max} > x^B_{ub}$ holds for all $\pi_2 \in (0, 1)$.

The utility of self 0 under narrow bracketing is:

$$U^N_0(x, y|x, y) = \pi_1\,\{1-\eta\,(\lambda-1)\,(1-\pi_1)\}\, b(x) - c(x) + \pi_2\,\{1-\eta\,(\lambda-1)\,(1-\pi_2)\}\, y.$$

The utility of self 0 under broad bracketing is:

$$U^B_0(x, y|x, y) = \begin{cases} \pi_1\,\{1-\eta\,(\lambda-1)\,(1-\pi_1)\}\, b(x) - c(x) \\ \quad + \pi_2\,\{1-\eta\,(\lambda-1)\,(1-\pi_2)\,(1-2\pi_1+2\pi_1^2)\}\, y, & \text{for } b(x) > y, \\ \pi_1\,\{1-\eta\,(\lambda-1)\,(1-\pi_1)\,(1-2\pi_2+2\pi_2^2)\}\, b(x) - c(x) \\ \quad + \pi_2\,\{1-\eta\,(\lambda-1)\,(1-\pi_2)\}\, y, & \text{for } b(x) < y, \\ \pi_1\,\{1-\eta\,(\lambda-1)\,(1-\pi_1)\,(1-\pi_2+\pi_2^2)\}\, b(x) - c(x) \\ \quad + \pi_2\,\{1-\eta\,(\lambda-1)\,(1-\pi_2)\,(1-\pi_1+\pi_1^2)\}\, y, & \text{for } b(x) = y. \end{cases}$$

Note that for $\pi_1 = 1$ and $\pi_2 \in (0, 1)$ the utility functions of self 0 under broad and narrow bracketing coincide, but because $x^N_{max} > x^B_{ub}$, narrow bracketing yields a strictly higher utility: $U^N_0(x^N_{max}, y|x^N_{max}, y)|_{\pi=1} > U^B_0(x^B_{ub}, y|x^B_{ub}, y)|_{\pi=1}$. Since the utility functions and maximal implementable decisions all are continuous in $\pi_1$, there exists some threshold $\check{\pi}_1 < 1$ such that $U^N_0(x^N_{max}, y|x^N_{max}, y) > U^B_0(x^B_{ub}, y|x^B_{ub}, y)$ also holds for $\pi_1 \in (\check{\pi}_1, 1]$.

*Non-CARA preferences for* $\pi_1 = 1$ *and* $\pi_2 \equiv \pi \in (0, 1)$ *(referring to the discussion in the text)* Assume that the consumption utility of the individual is $v(y + b(x))$, where $v(\cdot)$ is strictly increasing and strictly concave and $v(0) = 0$. The utility of self 0 for investment decision $x$ under a broad bracket is:

$$U^B_0(x, y|x, y) = \pi\, v(y + b(x)) + (1-\pi)\, v(b(x)) - \eta\,\lambda\,\pi\,(1-\pi)\,[v(y + b(x)) - v(b(x))].$$

And under narrow bracketing it is:

$$U_0^N(x, y|x, y) = \pi\, v(y + b(x)) + (1-\pi)\, v(b(x)) - \eta\,\lambda\,\pi\,(1-\pi)\,[v(y)].$$

Note that we assume that narrow bracketing only affects the gain–loss utility, but that the individual cannot through bracketing change his perception of the consumption utility. We believe this assumption to be most plausible because our model is about situations where narrow bracketing is by choice, and not by mistake (it is difficult to consciously perceive consumption utility differently, see also the discussion in section 4.1). However, our results would go through also if we assumed that narrow bracketing affected consumption utility as well.

The maximal implementable goal under narrow and broad bracketing, respectively, are defined by (for $x_{ub}^B$ note that for a small deviation $x < \hat{x}$ we have $y + b(x) > b(\hat{x})$):

$$\beta\,\{\pi\, v'(y + b(x_{max}^N)) + (1-\pi)\, v'(b(x_{max}^N)) + \eta\,\lambda\, v'(b(x_{max}^N))\}\, b'(x_{max}^N) = (1+\eta)\, c'(x_{max}^N),$$

$$\beta\,\{\pi\, v'(y + b(x_{max}^N)) + (1-\pi)\, v'(b(x_{max}^N)) + \eta\,[\lambda\,(1-\pi) v'(b(x_{ub}^B)) + \lambda\,\pi^2\, v'(y + b(x_{ub}^B)) + \pi\,(1-\pi)\, v'(y + b(x_{ub}^B))]\}\, b'(x_{ub}^B) = (1+\eta)\, c'(x_{ub}^B).$$

Note that $U_0^N(x, y|x, y) < U_0^B(x, y|x, y)$ except for $\pi \in \{0, 1\}$. In the latter case, the utility functions of self 0 coincide under narrow and broad bracketing. Similarly, note that our above definitions imply that $x_{max}^N > x_{ub}^B$, which can be understood more intuitively by considering the incentive effects of brackets. For degenerate background risk ($\pi = 1$), non-CARA preferences cause broad brackets to have a negative incentive effect: Under narrow brackets, marginal utility is $v'(b(x))$ in the gain–loss evaluation, while under broad brackets marginal utility is $v'(b(x) + y)$. Because of diminishing marginal utility, $v'(b(x) + y) < v'(b(x))$. Intuitively, the individual cares more about marginal losses (or gains) in the investment activity when he focuses narrowly on the investment income than when the investment income just adds to some other income $y$. For non-degenerate background risk ($\pi \in (0, 1)$) there additionally is the usual effect that risk pooling dampens incentives under broad brackets. While an investment $x$ just short of the goal $\hat{x}$ leads to a sure loss under narrow brackets, under broad brackets there is a probability $\pi(1-\pi)$ that $x$ will be evaluated in the gain rather than in the loss domain.

As a consequence of the negative incentive effect, for $\pi = 1$, the maximized utility under narrow bracketing is strictly larger than the one under broad bracketing. Because $x_{max}^N > x_{ub}^B$ for all $\pi \in (0, 1)$ it hence follows, by continuity of the consumption utility function and continuity of the maximal implementable decisions, that narrow brackets are also strictly better for some $\pi < 1$. Broad and narrow brackets yield the same maximized utility for $\pi = 0$. Depending on parameters, broad bracketing may be optimal for intermediate values of $\pi$.

*Correlation between marginal rate of substitution and shocks (referring to the discussion in the text)* Giné et al. (2012) consider deaths of a household member (occurring in 2 percent of the sample) and income shocks (that "tend to be small" in their sample). Tanaka et al. (2010), who examine time preferences in Vietnamese villages, consider whether the household head cannot work at the time of the experiment. According to their descriptive statistics only very few cannot do so. Further, they consider rainfall as a shock. Probabilities are harder to assess here, but according to the Vietnamese embassy in the UK "Rainfall is abundant, with annual rainfall exceeding 1000 mm almost everywhere". In contrast, in Dean and Sautmann (2015), the mean probability of an adverse event is reported to be 33.2 percent. Harrison et al. (2005) consider changes over time in how subjects perceive their own situation and the general macroeconomic situation. The two shocks that correlate with the marginal rate of substitution (changes in the own economic situation, and in some specifications also changes in employment status) are the ones that subjects believe to be most likely to happen (changes in the interest rate are also believed to be likely, but our predictions are not about how the individual uses the external credit market). Meier and Sprenger (2015) consider changes in the employment status (13 percent of subjects have changes), the number of dependents (12 percent show a change) and income (44 percent show a change). That is, only the latter occurs with a high probability. Hence, our model with non-CARA preferences predicts for the latter case that broad bracketing is optimal, i.e., a correlation. This however is not observed in Meier and Sprenger (2015). Yet, such a correlation between income changes and the marginal rate of substitution is observed in Krupka and Stephens (2013), who argue that their data is superior to that of Meier and Sprenger (2015) because they measure income changes during the months of the experiment rather than just in the year of the experiment as in Meier and Sprenger (2015).

### B.5. Proof of Corollary 1

The order of proofs in the appendix follows the one in the text. Yet, the reader is advised to read the proof of Proposition 4 before the proof of Corollary 1. To show that $x^B_{ub} = x^B_{max}$ and $x^B_{lb} = x^B_{min}$ we need to establish that there exists no profitable joint deviation if the individual brackets broadly with goals $\hat{x}_1 = \hat{x}_2 = x^B_0$ and $x^B_0 \in [x^B_{lb}, x^B_{ub}]$. By the same arguments as in the proof of Proposition 4, we can restrict attention to deviations $(x'_1, x'_2)$ of the kind $x'_2 < x^B_0 < x'_1$ and write the utility change from the deviation as a line integral $\int_{x^B_0}^{x'_1} \phi'(x_1)\, dx_1$. The aim is to show that $\phi'(x_1) < 0$ for every point $(x_1, z(x_1))$ on the path along which we are integrating. We abbreviate $b_1 \equiv b(x_1)$, $b_2 \equiv b(z(x_1))$, etc., and use short-hands for the slopes of the path as well as the iso-benefit and iso-cost curves, respectively: $-\frac{dz(x_1)}{dx_1} \equiv r_z$, $\frac{b'_1}{b'_2} \equiv r_b$, and $\frac{c'_1}{c'_2} \equiv r_c$. Taking the derivative of the utility of self 1 yields

$$\phi'(x_1) = \beta\pi \left[(1+\eta\kappa_1)\, b'_1 - (1+\eta\kappa_2)\, b'_2\, r_z\right] - (1+\eta\kappa_3)\left[c'_1 - c'_2\, r_z\right], \tag{33}$$

where the parameters $\kappa_j \in \{1, 2, 3\}$ depend on the kind of deviation (see below).

Case (a): $z(x'_1) < x^{cm}(x'_1)$. We exploit path-independence of the line integral. There exists a path from $(x^B_0, x^B_0)$ to $(x'_1, x'_2)$ for which at every point $0 < r_b < r_c < r_z$. The parameters for the benefit dimension are $\kappa_1 = 1 + (\lambda - 1)\pi^2$ and $\kappa_2 = 1 + (\lambda - 1)\pi\,[1 + (1-\pi)(1-2\pi)]$. And $\kappa_3 = 1$ because self 1 feels a gain in the cost dimension. Hence,

$$\begin{aligned}
\phi'(x_1) &= \beta\pi\, b'_2 \left[(1+\eta\kappa_1)\, r_b - (1+\eta\kappa_2)\, r_z\right] + (1+\eta)\, c'_2 \left[r_z - r_c\right] \\
&< \beta\pi\, b'_2 \left[(1+\eta\kappa_1)\, r_b - (1+\eta\kappa_2)\, r_z\right] + \beta\pi\, b'_2\, (1+\eta\kappa_2) \left[r_z - r_c\right] \\
&\propto (1+\eta\kappa_1)\, r_b - (1+\eta\kappa_2)\, r_c < (1+\eta\kappa_1)\, r_b - (1+\eta\kappa_2)\, r_b \\
&\propto (\kappa_1 - \kappa_2) = -2\pi\,(1-\pi)^2\,(\lambda - 1) < 0.
\end{aligned}$$

The second line exploits that $z(x_1) < x^B_0 \le x^B_{max}$ implies $\beta\pi\left[1+\eta\kappa_2\right] b'_2 > (1+\eta)\, c'_2$ (cf. Lemma 2, setting $b'_L(x) = 0$) and that $r_z > r_c$. The third line uses $r_c > r_b$.

Case (b): $x^{cm}(x'_1) \le z(x'_1) < x^{bm}(x'_1)$. Along the path $0 < r_b < r_z \le r_c$, self 1 feels a loss in the cost-dimension, i.e. $\kappa_3 = \lambda$. Hence,

$$\begin{aligned}
\phi'(x_1) &= \beta\pi\, b'_2 \left[(1+\eta\kappa_1)\, r_b - (1+\eta\kappa_2)\, r_z\right] + (1+\eta\lambda)\, c'_2 \left[r_z - r_c\right] \\
&\le \beta\pi\, b'_2 \left[(1+\eta\kappa_1)\, r_b - (1+\eta\kappa_2)\, r_z\right] < 0,
\end{aligned}$$

because $\kappa_2 > \kappa_1$, where $\kappa_1$ and $\kappa_2$ are the same as in case (a).

Case (c): $x^{bm}(x_1') \leq z(x_1') < x_0^B$. Along path $0 < r_z \leq r_b < r_c$, we have $\kappa_1 = 1 + (\lambda-1)\pi^2(1-\pi)$, $\kappa_2 = 1 + (\lambda-1)\pi(1-\pi)(2-\pi)$, and $\kappa_3 = \lambda$. Hence,

$$\begin{aligned}\phi'(x_1) &= \beta\pi\, b_1' \left[(1+\eta\kappa_1) - (1+\eta\kappa_2)\frac{r_z}{r_b}\right] - (1+\eta\lambda)\,c_1'\left[1-\frac{r_z}{r_c}\right] \\ &< \beta\pi\,(1+\eta\kappa_1)\,b_1'\left[1-\frac{r_z}{r_b}\right] - (1+\beta\eta\lambda)\,c_1'\left[1-\frac{r_z}{r_c}\right] \\ &< \left[\beta\pi\,(1+\eta\kappa_1)\,b_1' - (1+\eta\lambda)\,c_1'\right]\left[1-\frac{r_z}{r_c}\right] < 0.\end{aligned}$$

The second line exploits $\kappa_2 > \kappa_1$; the third uses $1 - r_z/r_c > 1 - r_z/r_b \geq 0$ and that $x_1 > x_0^B \geq x_{min}^B$ implies $\beta\pi\,[1+\eta\kappa_1]\,b_1' < (1+\beta\eta\lambda)\,c_1'$ (cf. Lemma 2, setting $b_L'(x) = 0$).

The last part of the result follows from Proposition 2 because $x_{ub}^B = x_{max}^B$.

### B.6. *Proof of Proposition 4*

With deterministic outcomes, the utility of self 0 is the same for narrow and broad bracketing. Further, the marginal incentives to deviate from the goal in a single task are the same under broad and narrow bracketing, i.e., $x_{max}^N = x_{ub}^B$. Thus, a necessary condition for goals $(x_1, x_2)$ to be broad-bracketing implementable is that $x_1$ and $x_2$ are narrow-bracketing implementable. The proof explores when and whether profitable joint deviations under broad bracketing arise so that $(x_1, x_2)$ are not broad-bracketing implementable.

We repeatedly use the following concepts. For the pair of goals $(\hat{x}_1, \hat{x}_2)$ and a deviation to $x_1 \neq \hat{x}_1$ define the *cost-matching decision* $x^{cm}(x_1)$ by $c(\hat{x}_1) + c(\hat{x}_2) = c(x_1) + c(x^{cm}(x_1))$, and the *benefit-matching decision* $x^{bm}(x_1)$ by $b(\hat{x}_1) + b(\hat{x}_2) = b(x_1) + b(x^{bm}(x_1))$.

*Part 1* Consider first the case where $x_0^N$ is implementable in both activities under narrow bracketing. We show that they are not broad-bracketing implementable. Our proof is for the case where task 1 is a neutral activity and task 2 is a leisure activity. The other cases are analogous.

Consider a deviation along the iso-cost curve to $x_1' < x_0^N$ and $x^{cm}(x_1) > x_0^N$. Such a deviation causes a loss in the benefit dimension relative to sticking with the goals because the cost-matching decision is lower than the benefit-matching one: $x^{bm}(x_1) > x^{cm}(x_1) > \hat{x}_2$. To see this note that the iso-cost and iso-benefit curves are tangent at $\hat{x}_1 = \hat{x}_2$. For joint deviations from the goals $x_1 < \hat{x}_1$ and $x_2 > \hat{x}_2$, we have $c'(x_1) \leq c'(\hat{x}_1)$ and $c'(x_2) \geq c'(\hat{x}_2)$ as well as $b'(x_1) \geq b'(\hat{x}_1)$ and $b'(x_2) \leq b'(\hat{x}_2)$. Hence

$$\begin{aligned}\left|\left.\frac{d\,x_2}{d\,x_1}\right|_{\sum_i c(x_i)=\sum_i c(\hat{x}_i)}\right| &= \frac{c'(x_1)}{c'(x_2)} \leq \frac{c'(\hat{x}_1)}{c'(\hat{x}_2)} = \frac{b'(\hat{x}_1)}{b'(\hat{x}_2)} = 1 \leq \left|\left.\frac{d\,x_2}{d\,x_1}\right|_{\sum_i b(x_i)=\sum_i b(\hat{x}_i)}\right| \\ &= \frac{b'(x_1)}{b'(x_2)}.\end{aligned}$$

At least one of the inequalities above is strict. Hence, the utility of self 1 from such a deviation is:

$$\begin{aligned}&b(x_1') + b(x^{cm}(x_1)) - \left[c(x_1') + \beta\,c(x^{cm}(x_1))\right] \\ &+\eta\lambda\left(b(x_1') + b(x^{cm}(x_1)) - b(x_0^N) - b(x_0^N)\right).\end{aligned}$$

Evaluating the impact on the utility of self 1 of a marginal deviation along the iso-cost curve at point $(x_0^N, x_0^N)$, we obtain (omitting arguments: $b_i' \equiv b'(x_i)$ etc.):

$$(1+\eta\,\lambda)\left[b_1' + b_2'\,\frac{d\,x_2}{d\,x_1}\bigg|_{\sum c_i=\sum \hat{c}_i}\right] - \left[c_1' + \beta\,c_2'\,\frac{d\,x_2}{d\,x_1}\bigg|_{\sum c_i=\sum \hat{c}_i}\right]$$
$$= (1+\eta\,\lambda)\left[b_1' - b_2'\,\frac{c_1'}{c_2'}\right] - \left[c_1' - \beta\,c_1'\right] = -\left[1-\beta\right]c_1',$$

where the last line uses $c'(x_0^N) = b'(x_0^N)$. Hence, decreasing $x_1$ (and increasing $x_2$) starting from $(x_0^N, x_0^N)$ increases the utility of self 1. Thus, $x_0^N$ is not broad-bracketing implementable.

Consider next the case where $x_0^N$ (in at least one task) is not implementable under narrow bracketing. Consider the case where task 1 is a leisure activity and where task 2 is an investment activity (the other cases are analogous). The maximal and minimal implementable goals differ for investment and leisure activities. We denote them by $x_{max,L}/x_{min,L}$ for the leisure activity and by $x_{max,I}/x_{min,I}$ for the investment activity.[37] Suppose now $x_0^N > x_{max,I}^N$ for the investment activity, i.e., $\frac{\beta\,(1+\eta\,\lambda)}{1+\eta} < 1$. This implies that $x_0^N < x_{min,L}^N$ for the leisure activity. Thus, self 0 sets goals $(\hat{x}_1, \hat{x}_2) = (x_{min,L}^N, x_{max,I}^N)$ under narrow brackets. Using the definitions of $x_{min,L}$ and $x_{max,I}$, and that $x_0^N > x_{max,I}^N$,

$$\frac{(1+\eta)^2}{\beta^2\,(1+\eta\,\lambda)^2}\,\frac{b'(x_{min,L}^N)}{b'(x_{max,I}^N)} = \frac{c'(x_{min,L}^N)}{c'(x_{max,I}^N)} \Longrightarrow \frac{b'(x_{min,L}^N)}{b'(x_{max,I}^N)} > \frac{c'(x_{min,L}^N)}{c'(x_{max,I}^N)}.$$

Consider now a move along the iso-benefit curve to $x_1' > x_{min,L}^N$ and $x^{bm}(x_1) < x_{max,I}^N$. Such a move causes a gain in the cost dimension relative to sticking with the goals. The utility of self 1 from such a deviation is:

$$b(x_1') + \beta\,b(x^{cm}(x_1)) - \left[\beta\,c(x_1') + c(x^{cm}(x_1))\right]$$
$$+\beta\,\eta\left(c(x_1') + c(x^{cm}(x_1)) - c(x_{min,L}^N) - c(x_{max,I}^N)\right).$$

Evaluating the impact on the utility of self 1 of a marginal deviation along the iso-cost curve at point $(x_{min,L}^N, x_{max,I}^N)$, we obtain (omitting arguments):

$$b_1' + \beta\,b_2'\,\frac{d\,x_2}{d\,x_1}\bigg|_{\sum b_i=\sum \hat{b}_i} - \left[\beta\,(1+\eta)\,c_1' + (1+\beta\,\eta)\,c_2'\,\frac{d\,x_2}{d\,x_1}\bigg|_{\sum b_i=\sum \hat{b}_i}\right]$$
$$= (1-\beta)\,b_1' + c_2'\left[(1+\beta\,\eta)\,\frac{b_1'}{b_2'} - \beta\,(1+\eta)\,\frac{c_1'}{c_2'}\right] > 0,$$

where the strict inequality follows because $1-\beta > 0$, $(1+\beta\,\eta) > \beta\,(1+\eta)$ and $\frac{b'(x_{min,L}^N)}{b'(x_{max,I}^N)} > \frac{c'(x_{min,L}^N)}{c'(x_{max,I}^N)}$. Hence, increasing $x_1$ (and decreasing $x_2$) starting from $(x_{min,L}^N, x_{max,I}^N)$ increases the utility of self 1. Thus, $(x_{min,L}^N, x_{max,I}^N)$ is not broad-bracketing implementable.

[37] For the investment activity $x_{min,I}^N$ is defined by: $\frac{\beta\,(1+\eta)}{1+\eta\,\lambda}\,b'(x_{min,I}^N) = c'(x_{min,I}^N)$, $x_{max,I}^N$ is defined by: $\frac{\beta\,(1+\eta\,\lambda)}{1+\eta}\,b'(x_{max,I}^N) = c'(x_{max,I}^N)$. For the leisure activity $x_{min,L}^N$ is defined by: $\frac{(1+\eta)}{\beta\,(1+\eta\,\lambda)}\,b'(x_{min,L}^N) = c'(x_{min,L}^N)$, $x_{max,L}^N$ is defined by: $\frac{(1+\eta\,\lambda)}{\beta\,(1+\eta)}\,b'(x_{max,L}^N) = c'(x_{max,L}^N)$.

*Part 2* In the following, we consider two leisure activities – the other cases are analogous. As activities are symmetric, self 0 sets goals $\hat{x}_1 = \hat{x}_2$ under narrow bracketing and thus the iso-cost and the iso-benefit curve are tangent (with slope 1) at $(\hat{x}_1, \hat{x}_2)$. Note that the goals being narrow-bracketing implementable rules out unilateral deviations, and it rules out joint deviations in which self 1 either increases both decisions or decreases both decisions (changing both decisions in the same way results in the same gains and losses as a unilateral deviation). So we only need to consider joint deviations in which self 1 increases one decision and decreases the other, which requires first that $\hat{x}_1 = \hat{x}_2 \in (x^N_{min}, x^N_{max})$. Without loss of generality, suppose that self 1 deviates to $x'_1 < \hat{x}_1$ and $x'_2 > \hat{x}_2$. We are interested in the utility difference $U_1(x'_1, x'_2|\hat{x}_1, \hat{x}_2) - U_1(\hat{x}_1, \hat{x}_2|\hat{x}_1, \hat{x}_2)$. By the gradient theorem, we can express this change in utility as a line integral along a curve with endpoints $(\hat{x}_1, \hat{x}_2)$ and $(x'_1, x'_2)$. That is, abbreviating $U_1(\cdot) \equiv U_1(\cdot|\hat{x}_1, \hat{x}_2)$,

$$U_1(x'_1, z(x'_1)) - U_1(x^N_0, z(x^N_0)) = \\ -\int_{x'_1}^{x^N_0} \underbrace{\left[\frac{\partial}{\partial x_1} U_1(x_1, z(x_1)) + \frac{\partial}{\partial x_2} U_1(x_1, z(x_1)) \frac{d\,z(x_1)}{d\,x_1}\right]}_{\equiv \phi'(x_1)} d\,x_1, \tag{34}$$

where $z : [x'_1, \hat{x}_2] \to [\hat{x}_1, x'_2]$ is an arbitrary bijective function with $z(x'_1) = x'_2$ and $z(\hat{x}_1) = \hat{x}_2$. Further, the gradient theorem implies that the line integral does not depend on the path $z(x_1)$ between the two endpoints, which will turn out to be useful for our purposes. We show that the integrand $\phi'(x_1) > 0$ for any $x_1 \in [x'_1, \hat{x}_1]$, which implies that the utility difference in equation (34) is negative. That is, any deviation $(x'_1, x'_2)$ causes a drop in overall utility. There are three possible types of deviations $(x'_1, x'_2)$ with $x'_1 < \hat{x}_1$ and $x'_2 > \hat{x}_2$: (a) $\hat{x}_2 < z(x'_1) < x^{cm}(x_1)$, (b) $x^{cm}(x_1) \le z(x'_1) < x^{bm}(x_1)$, and (c) $x^{bm}(x_1) \le z(x'_1)$.

Case (a): Because $(x'_1, x'_2)$ lies below the iso-cost curve that goes through $(\hat{x}_1, \hat{x}_2)$, there exists a path connecting the two points for which at every point the slope is less steep than the slope of the iso-cost curve, which in turn is less steep than the slope of the iso-benefit curve. That is,

$$0 < -\frac{d\,z(x_1)}{d\,x_1} < \frac{c'(x_1)}{c'(x_2)} < \frac{b'(x_1)}{b'(x_2)}. \tag{35}$$

Hence, at every point along this path the individual would experience a gain in the cost dimension and a loss in the benefit dimension. Exploiting path-independence, we can fix a feasible path by imposing conditions (35), and write the integrand for a given $x_1$ as

$$\begin{aligned}\phi'(x_1) &= \kappa\left[b'(x_1) + b'(z(x_1))\frac{d\,z(x_1)}{d\,x_1}\right] - \omega\left[c'(x_1) + c'(z(x_1))\frac{d\,z(x_1)}{d\,x_1}\right] \\ &= \kappa\, b'(x_1)\left[1 - \frac{b'(z(x_1))}{b'(x_1)}\left(-\frac{d\,z(x_1)}{d\,x_1}\right)\right] \\ &\quad - \omega\, c'(x_1)\left[1 - \frac{c'(z(x_1))}{c'(x_1)}\left(-\frac{d\,z(x_1)}{d\,x_1}\right)\right],\end{aligned} \tag{36}$$

where $\kappa = 1 + \eta\,\lambda$, and $\omega = \beta\,(1+\eta)$. Conditions (35) give

$$0 < \frac{c'(z(x_1))}{c'(x_1)}\left(-\frac{d\,z(x_1)}{d\,x_1}\right) < 1 \quad \text{and} \quad \frac{c'(z(x_1))}{c'(x_1)} > \frac{b'(z(x_1))}{b'(x_1)},$$

which implies that

$$1 - \frac{b'(z(x_1))}{b'(x_1)}\left(-\frac{d\,z(x_1)}{d\,x_1}\right) > 1 - \frac{c'(z(x_1))}{c'(x_1)}\left(-\frac{d\,z(x_1)}{d\,x_1}\right) > 0. \tag{37}$$

Hence,

$$\phi'(x_1) > \underbrace{\left[\kappa\, b'(x_1) - \omega\, c'(x_1)\right]}_{>0 \text{ for } x_1 < x^N_{max}} \underbrace{\left[1 - \frac{c'(z(x_1))}{c'(x_1)}\left(-\frac{d\,z(x_1)}{d\,x_1}\right)\right]}_{>0;\text{ see eq. (37)}} > 0.$$

Case (b): Now $(x_1', x_2')$ lies between iso-cost and -benefit curves, so there exists a path with

$$0 < \frac{c'(x_1)}{c'(z(x_1))} \leq -\frac{d\,z(x_1)}{d\,x_1} < \frac{b'(x_1)}{b'(z(x_1))}. \tag{38}$$

At every point along the path the individual would experience a loss in the cost and benefit dimensions. So we replace in equation (36) $\kappa = 1 + \eta\,\lambda$, and $\omega = \beta\,(1 + \eta\,\lambda)$. The first term in equation (36) remains positive, but conditions (38) now imply $1 - \frac{c'(z(x_1))}{c'(x_1)}\left(-\frac{d\,z(x_1)}{d\,x_1}\right) \leq 0$. Hence, $\phi'(x_1) > 0$.

Case (c): Now $(x_1', x_2')$ lies above iso-cost and -benefit curves, so there exists a path with

$$0 < \frac{c'(x_1)}{c'(z(x_1))} < \frac{b'(x_1)}{b'(z(x_1))} \leq -\frac{d\,z(x_1)}{d\,x_1}. \tag{39}$$

At every point along the path the individual would experience a gain in the benefit dimension and a loss in the cost dimension. So we replace in equation (36) $\kappa = 1 + \eta$, and $\omega = \beta\,(1 + \eta\,\lambda)$. Rearranging terms gives

$$\begin{aligned}\phi'(x_1) &= \omega\, c'(z(x_1))\left[\left(-\frac{d\,z(x_1)}{d\,x_1}\right) - \frac{c'(x_1)}{c'(z(x_1))}\right] - \kappa\, b'(z(x_1))\left[\left(-\frac{d\,z(x_1)}{d\,x_1}\right) - \frac{b'(x_1)}{b'(z(x_1))}\right]\\ &> \underbrace{\left[\omega\, c'(z(x_1)) - \kappa\, b'(z(x_1))\right]}_{>0 \text{ for } x_2 > x^N_{min}} \underbrace{\left[\left(-\frac{d\,z(x_1)}{d\,x_1}\right) - \frac{b'(x_1)}{b'(z(x_1))}\right]}_{\geq 0;\text{ see eq. (39)}} \geq 0.\end{aligned}$$

The second-to-last inequality exploits conditions (39).

### B.7. Proof of Proposition 5

*Utility of self 0 under narrow bracketing* Following Kőszegi and Rabin (2009), the individual compares the worst percentile of outcomes under the new distribution with the worst percentile of outcomes under the old distribution, then the second worst percentile of outcomes, and so on. Self 1 (3) inherits for each activity 1 (2) from self 0 the reference distribution $[\pi \circ b(\hat{x}_i);\ (1-\pi) \circ 0]$ for the delayed benefit and $[\pi \circ c(\hat{x}_i);\ (1-\pi) \circ 0]$ for the cost. He then observes the state for the activity to be conducted at $t = 2$ ($t = 4$) and updates the reference distributions. If he observes that the state is productive, he knows that he will incur benefit $b(\hat{x}_i)$ and cost $c(\hat{x}_i)$ with probability 1. However, he previously expected that with probability $1 - \pi$ the unproductive state occurs, i.e., costs and benefits of zero. The comparison of the new reference distribution with past expectations triggers anticipatory utility $\eta\,(1-\pi)\,[\gamma\, b(\hat{x}_i) - \gamma\,\lambda\, c(\hat{x}_i)]$. Similarly, if self $t \in \{1, 3\}$ observes that the state at $t + 1$ is unproductive, this triggers anticipatory utility

$-\eta\,\pi\,[\gamma\,\lambda\,b(\hat{x}_i) - \gamma\,c(\hat{x}_i)]$. Taking expectations over the anticipatory utility in the possible states for the two tasks and adding consumption utility, the utility of self 0 is

$$U_0^N(\hat{x}_1, \hat{x}_2 | \hat{x}_1, \hat{x}_2) = \beta\,\pi \sum_{i=1}^{2} \left\{ b(\hat{x}_i) - c(\hat{x}_i) - \eta\,(1-\pi)\,(\lambda - 1)\,\gamma\,(b(\hat{x}_i) + c(\hat{x}_i)) \right\}. \quad (40)$$

*Incentive constraints under narrow bracketing* We derive the bounds that determine the maximal implementable goals under narrow bracketing (arguments for the minimal implementable goals are analogous). In the unproductive state, the individual chooses zero and no self has an incentive to deviate from this. Denote the equilibrium decision in task $i \in \{1, 2\}$ if it is in the productive state by $\hat{x}_i$. For these to indeed be equilibrium decisions, they need to be implementable, i.e., self $t$, $t \in \{1, 2, 3, 4\}$ should have no incentive to deviate from them, or revise them. Suppose that at $t = 1$ the individual observes the productive state for task 1 (the case where self 3 observes the productive state at $t = 2$ is parallel). If he revises the goal for activity 1 downwards to $x_1 < \hat{x}_1$ his utility from activity 1 is (as utility is additive separable under narrow bracketing across activities, we suppress the utility from activity 2):

$$\begin{aligned} U_1^N(x_1|\hat{x}_1) = \beta\,b(x_1) - c(x_1) + \eta\,\gamma\,\left[\pi\,\lambda\,(b(x_1) - b(\hat{x}_1)) + (1-\pi)\,(b(x_1) - 0)\right] \\ -\eta\,\gamma\,\left[\pi\,(c(x_1) - c(\hat{x}_1)) + (1-\pi)\,\lambda\,(c(x_1) - 0)\right]. \end{aligned}$$

Self 1 experiences anticipatory utility from comparing the new (degenerate) reference distributions with the inherited reference distribution. That is, for the benefits he compares $b(x_1)$ with $(\pi \circ b(\hat{x}_1); (1-\pi) \circ 0)$. Similarly for the costs. He has no incentive to revise the goal for activity 1 whenever $U_1^N(\hat{x}_1|\hat{x}_1) \geq U_1^N(x_1|\hat{x}_1)$, which is the case for $\hat{x}_1 \leq \tilde{x}_{max}^N$ defined by:

$$(\beta + \eta\,\gamma\,[1 + (\lambda - 1)\,\pi])\;b'(\tilde{x}_{max}^N) = (1 + \eta\,\gamma\,[\lambda - (\lambda - 1)\,\pi])\;c'(\tilde{x}_{max}^N). \quad (41)$$

If at $t = 1$ the individual revises the goal for the productive state of task 2 to $x_2 < \hat{x}_2$, his utility is given by (suppressing the utility from activity 1):

$$\begin{aligned} U_1^N(x_2|\hat{x}_2) = \beta\,\pi\,b(x_2) - \beta\,\pi\,c(x_2) + \eta\,\gamma\,\left[\pi\,\lambda\,(b(x_2) - b(\hat{x}_2)) - \pi\,(c(x_2) - c(\hat{x}_2))\right] \\ -\eta\,\beta\,\gamma\,\pi\,(1-\pi)\,(\lambda - 1)\;(b(x_2) + c(x_2)). \end{aligned}$$

For the benefits, self 1 compares $(\pi \circ b(x_2); (1-\pi) \circ 0)$ with $(\pi \circ b(\hat{x}_2); (1-\pi) \circ 0)$. Similarly for the costs. Moreover, he anticipates that the goal revision will affect the anticipatory utility that self 2 feels after receiving pre-task information. The expected anticipatory gain–loss utility is $-\eta\,\gamma\,\pi\,(1-\pi)\,(\lambda - 1)\;(b(x_2) + c(x_2))$. It results from comparisons of benefits and costs in each state with the then inherited reference distributions $(\pi \circ b(x_2); (1-\pi) \circ 0)$ and $(\pi \circ c(x_2); (1-\pi) \circ 0)$, respectively. Self 1 has no incentive to revise the goal for activity 2 whenever $U_1^N(\hat{x}_2|\hat{x}_2) \geq U_1^N(x_2|\hat{x}_2)$, which is the case for $\hat{x}_2 \leq \check{x}_{max}$:

$$\begin{aligned} &(\beta + \eta\,\gamma\,\lambda - \eta\,\beta\,\gamma\,(1-\pi)\,(\lambda - 1))\;b'(\check{x}_{max}^N) \\ &\quad = (\beta + \eta\,\gamma + \eta\,\beta\,\gamma\,(1-\pi)\,(\lambda - 1))\;c'(\check{x}_{max}^N). \end{aligned} \quad (42)$$

Comparing equations (41) and (42) reveals that $\check{x}_{max}^N \geq \tilde{x}_{max}^N$. At $t = 2$ ($t = 4$, respectively) the individual faces no uncertainty. We now ask when he has no incentive to deviate from his goal in activity 1 (2, respectively). The incentive constraints are the same as in inequality (1) for $\pi = 1$, with the only difference that gain–loss utility in the benefit dimension receives weight $\gamma$ as a result of goal revision. That is, the individual has no incentive to deviate from the (possibly already revised) goal if it does not exceed $x^N_{max}$ defined by $(\beta + \eta\,\gamma\,\lambda)\,b'(x^N_{max}) = (1+\eta)\,c'(x^N_{max})$. Depending on whether $\gamma$ is smaller or larger than $\beta$, $x^N_{max}$ may be smaller or larger than the threshold defined in equation (4) for $\pi = 1$. Overall, the maximal implementable goal for task 1 is $\min\{x^N_{max}, \tilde{x}^N_{max}\}$, and for task 2 it is $\min\{x^N_{max}, \tilde{x}^N_{max}, \check{x}^N_{max}\}$.

*Part 1* We need to verify that the utility of self 0 under broad bracketing reduces to equation (40), i.e., that $U^N_0(\hat{x}_1, \hat{x}_2|\hat{x}_1, \hat{x}_2) = U^B_0(\hat{x}_1, \hat{x}_2|\hat{x}_1, \hat{x}_2)$. Self 1 inherits from self 0 the reference distribution $[\pi^2 \circ (b(\hat{x}_1) + b(\hat{x}_2));\, \pi\,(1-\pi) \circ b(\hat{x}_1);\, \pi\,(1-\pi) \circ b(\hat{x}_2);\, (1-\pi)^2 \circ 0]$. He learns the productivity for task 1 and updates the reference distribution to $[\pi \circ (b(x_1) + b(\hat{x}_2));\, (1-\pi) \circ (b(x_1) + 0)]$, where $x_1 \in \{0, \hat{x}_1\}$, depending on the realized state. Self 3 learns the productivity for task 2 and updates the reference distribution to $b(x_1) + b(x_2)$, $x_2 \in \{0, \hat{x}_2\}$. Similarly, for the costs. Each revision of expectations triggers anticipatory utility. For example, if self 1 learns that the state is unproductive, this triggers anticipatory utility $-\eta\,\pi\,[\gamma\,\lambda\,b(\hat{x}_i) - \gamma\,c(\hat{x}_i)]$. Proceeding in this fashion for the other states and dates, one obtains exactly the same utility of self 0 as under narrow bracketing, given in equation (40). As a result, the preferred decisions of self 0 under broad and narrow bracketing are the same. For the unproductive state it is zero. For the productive state in task $i \in \{1, 2\}$, the preferred decision $x^N_0 = x^B_0$ maximizes the utility in equation (40) and is defined by

$$\frac{1 - \eta\,\gamma\,(1-\pi)\,(\lambda-1)}{1 + \eta\,\gamma\,(1-\pi)\,(\lambda-1)}\,b'(x^N_0) = c'(x^N_0). \tag{43}$$

Assumption 1 is sufficient to ensure an interior solution.

*Part 2* It suffices to show that narrow-bracketing implementable goals $(x^N_0, x^N_0)$ are not broad-bracketing implementable. Then – because $(x^N_0, x^N_0)$ maximize the utility of self 0 under both broad and narrow brackets – $U^N_0(x^N_0, x^N_0|x^N_0, x^N_0) > U^B_0(\hat{x}_1, \hat{x}_2|\hat{x}_1, \hat{x}_2)$ for any goals $(\hat{x}_1, \hat{x}_2)$ that are broad-bracketing implementable.

Consider a small deviation along the iso-cost curve for which $c(x_1) + c(x_2) = c(x^N_0) + c(x^N_0)$ and $b(x^N_0) < b(x_1) + b(x_2) < b(x^N_0) + b(x^N_0)$. What are the marginal incentives for such a deviation from the goals $(x^N_0, x^N_0)$? Upon observing the productive state and revising goals in this way, the individual at $t = 1$ has utility (omitting arguments: $b_i \equiv b(x_i)$ etc.)

$$\begin{aligned} &\beta\,b_1 - c_1 + \beta\,\pi\,(b_2 - c_2) \\ &+ \eta\,\gamma\, \times \text{[anticipatory utility: learn state in task 1 \& goal revision]} \\ &+ \beta\,\eta\,\gamma\, \times \text{[future anticipatory utility: learn state in task 2].} \end{aligned} \tag{44}$$

Taking the derivative of the utility of self 1 in equation (44) with respect to $x_1$, dividing by $c'_1$, and using the fact that cost matching implies $\frac{d\,x_2}{d\,x_1} = -\frac{c'_1}{c'_2}$, we obtain

$$\begin{aligned} &\beta\,\frac{b'_1}{c'_1} - 1 + \beta\,\pi\left(\frac{b'_2}{c'_2} - 1\right)\frac{c'_2}{c'_1}\frac{d\,x_2}{d\,x_1} - \beta\,\eta\,\gamma\,\pi\,(1-\pi)\,(\lambda-1)\left(\frac{b'_2}{c'_2} + 1\right)\frac{c'_2}{c'_1}\frac{d\,x_2}{d\,x_1} \\ &+ \eta\,\gamma\left[[1 + (\lambda-1)\,\pi]\left(\frac{b'_1}{c'_1} + \pi\,\frac{b'_2}{c'_2}\frac{c'_2}{c'_1}\frac{d\,x_2}{d\,x_1}\right) - (1-\pi)\,[1 + (\lambda-1)\,(1-\pi)]\right] \\ \propto\;&\beta\,\frac{b'_1}{c'_1} - 1 - \beta\,\pi\left(\frac{b'_2}{c'_2} - 1\right) + \beta\,\eta\,\gamma\,\pi\,(1-\pi)\,(\lambda-1)\left(\frac{b'_2}{c'_2} + 1\right) \end{aligned}$$

$$+\,\eta\,\gamma\left[[1+(\lambda-1)\,\pi]\left(\frac{b_1'}{c_1'}-\pi\,\frac{b_2'}{c_2'}\right)-(1-\pi)\,[1+(\lambda-1)\,(1-\pi)]\right]$$

$$\le\,\beta\,(1-\pi)\,\frac{b_2'}{c_2'}-(1-\beta\,\pi)+\eta\,\gamma\,(1-\pi)\left\{[1+(\lambda-1)\,\pi]\,\frac{b_2'}{c_2'}-[1+(\lambda-1)\,(1-\pi)]\right.$$

$$\left.+\,\beta\,\pi\,(\lambda-1)\left(\frac{b_2'}{c_2'}+1\right)\right\}\equiv\Psi(\pi). \qquad (45)$$

The last step uses $\frac{b_2'}{c_2'}\ge\frac{b_1'}{c_1'}>1$ at $(x_1=x_0^N, x_2=x_0^N)$. Note that $\Psi(1)=-(1-\beta)<0$. Further, for $\pi=0$, the expression in equation (45) becomes $\frac{b_1'}{c_1'}\,(\beta+\eta\,\gamma)-(1+\eta\,\lambda\,\gamma)>0$, where the inequality follows from applying the definition of $\tilde{x}_{max}^N$ in equation (41) and the fact that $x_0^N<\tilde{x}_{max}^N$. Hence, by the intermediate value theorem, there exists a $\hat{\pi}\in(0,1)$, such that for all $\pi\ge\hat{\pi}$ decision substitution pays off. Here the preferred decisions of self 0 are not broad-bracketing implementable and narrow bracketing is optimal.

## Appendix C. Taxi driver application (Proof of Proposition 6)

We apply the pre-task information model from section 4.2 of the paper, assuming that self 1 (3) observes the daily wage in the morning and self 2 (4) makes the decision how many hours to drive. It is convenient to normalize $w_H=1$ and work with $w_L=\alpha\,w_H$, $\alpha<1$, $c(x)=(c\,x_s)^2/2$ and $c=2$. One can think of $x_s\in[0,1]$ as the fraction of the day spent working. Scaling $x_s$ up by 24 gives the hours in Fig. 2 in the paper.

### *C.1. Narrow bracketing*

#### *C.1.1. Negative wage elasticity of labor supply ($x_L>x_H$)*

*The preferred decisions of self 0* The utility of self 0 from decisions $x_L>x_H$ is given by

$$\beta\left\{\pi\,(x_H-c(x_H))+(1-\pi)\,(\alpha\,x_L-c(x_L))\right.$$
$$\left.-\,\eta\,\pi\,(1-\pi)\,\gamma\,(\lambda-1)\,(x_H-\alpha\,x_L+c(x_L)-c(x_H))\right\}$$

and takes into account future anticipatory utility after learning the state. Taking the derivative and using the quadratic cost function gives that the utility is increasing in the respective decisions up to $x_{0,H}^*=1/c$ and $x_{0,L}^*=\alpha/c$, respectively.

*After learning that it is a high-wage day* Suppose self 1 (3) revises his goal slightly downward to $x<x_H$, such that still $x>\alpha\,x_L$ and $c(x_H)>c(x)$. This triggers an adjustment from the original reference distribution $(\pi\circ x_H;\,1-\pi\circ x_L)$ to the realized state and revised goal. Hence, the utility of self 1 (3) from a small deviation from the goal $x_H$ for the high-wage day is

$$\beta\,x-c(x)+\eta\,\gamma\,\{\pi\,[-\lambda\,(x_H-x)+c(x_H)-c(x)]+(1-\pi)\,[x-\alpha\,x_L+c(x_L)-c(x)\}\,.$$

Taking the derivative and using the quadratic cost function gives:

$$x_{max,H}^N=\frac{\beta+\eta\,\gamma\,(1+\pi\,(\lambda-1))}{(1+\eta\,\gamma)\,c}.$$

*After learning that it is a low-wage day* If self 1 (3) revises the goal slightly downward to $x < x_L$, such that still $x > \alpha x_L$ and $c(x_L) > c(x) > c(x_H)$, his utility is:

$$\beta \alpha x - c(x) + \eta \gamma \left\{\pi \left[-\lambda (x_H - x) + c(x_H) - c(x)\right] + (1-\pi)\left[\alpha x - \alpha x_L + c(x_L) - c(x)\right\}.\right.$$

Taking the derivative and using the quadratic cost function gives:

$$x^N_{max,L} = \frac{(\beta + \eta \gamma \lambda) \alpha}{(1 + \eta \gamma (1 + \pi (\lambda - 1)) c}.$$

It is easy to show that $x^N_{max,L}$ and $x^N_{max,H}$ are the relevant constraints on the goals of self 0.[38]

*C.1.2. Alternative case: $x_L \leq x_H$*

*The preferred decisions of self 0* The utility of self 0 from decisions $x_L \leq x_H$ is:

$$\beta \left\{\pi (x_H - c(x_H)) + (1-\pi)(\alpha x_L - c(x_L)) - \eta \pi (1-\pi) \gamma (\lambda - 1)(x_H - \alpha x_L + c(x_H) - c(x_L))\right\}.$$

Taking the derivative and using the quadratic cost function gives that the utility is increasing in the respective decisions up to

$$\tilde{x}^*_{0,H} = \frac{1 - \eta \gamma (1-\pi)(\lambda - 1)}{(1 + \eta \gamma (1-\pi)(\lambda - 1)) c} \quad \text{and} \quad \tilde{x}^*_{0,L} = \frac{(1 + \eta \gamma \pi (\lambda - 1)) \alpha}{(1 - \eta \gamma \pi (\lambda - 1)) c}.$$

*The maximal implementable goals in states $s = L, H$ are*

$$\tilde{x}^N_{max,H} = \frac{\beta + \eta \gamma (1 + \pi (\lambda - 1))}{(1 + \eta \gamma(\lambda - \pi (\lambda - 1)) c} \quad \text{and} \quad \tilde{x}^N_{max,L} = \frac{(\beta + \eta \gamma \lambda) \alpha}{(1 + \eta \gamma) c}.$$

*Comments* The parameter values $\alpha = 0.75$, $\eta = 1$, $\lambda = 2.5$, $\beta = \gamma = 0.5$, and $c = 2$, for example, ensure that narrow-bracketing implementable goals have negative elasticities of labor supply ($x_L > x_H$). Specifically, with these parameter values, the alternative case cannot arise because the chosen decisions satisfy $x_L > x_H$, which contradicts the assumption that $x_L \leq x_H$. Equilibria for the case with negative wage elasticity however do exist for $\pi$ up to around 0.16. Scaling parameters produces Fig. 2 in the paper: $hours_s = 24 x_s$, $s = L, H$, $earnings_H = 30\, hours_H$, $earnings_L = 22.5\, hours_L$ (i.e. $\alpha = 0.75$).

## C.2. Broad bracketing

In our parameterized example, the optimal narrow-bracketing implementable goals are $x_L = x^*_{0,L}$ and $x_H = x^N_{max,H}$. With broad brackets we now show that on the morning of the first day the

[38] The corresponding upper bounds derived by looking at deviations by self 2 (4) are:

$$x^N_{max,H2} = \frac{\beta + \eta \gamma \lambda}{(1 + \eta \gamma) c} > x^N_{max,H} \quad \text{and} \quad x^N_{max,L2} = \alpha\, x^N_{max,H2} > x^N_{max,L}.$$

The extra constraints that self 1 does not want to change the effort goal for the second day are:

$$\check{x}^N_{max,H} = \frac{\beta + \eta \gamma \lambda - \eta \gamma (1-\pi)}{(1 + \eta \gamma - \eta \gamma (1-\pi)) c} > x^N_{max,H} \quad \text{and} \quad \check{x}^N_{max,L} = \frac{\beta + \eta \gamma \lambda + \eta \gamma \pi}{(1 + \eta \gamma + \eta \gamma \pi) c} > x^N_{max,L}.$$

individual (self 1) has an incentive to engage in decision substitution by changing the goal for the first and second days. Given goal $x_H = x^N_{max,H}$, self 4 is committed not to increase effort on a high-wage day. But since $x_L = x^*_{0,L} < x^N_{max,L}$ there is slack on a low-wage day. We find that it pays off to work a bit less if the first day is a high-wage day and plan to make up on the next day, in case this turns out to be a low-wage day.

Consider a small deviation where, if the first day is a high-wage day, the individual works less that day and works more on the second day if it turns out to be a low-wage day, to match across-state joint costs: $c(x_1) + c(x_2) = c(x_H) + c(x_L)$: $x_1 < x_H < x_L < x_2$, but still $\alpha\, x_L < x_1$ and $\alpha\, x_2 < x_H$. Then the utility of self 1 is given by:

$$\begin{aligned}
&\beta\, x_1 - c_1 + \beta\,[\pi\,(x_H - c(x_H)) + (1-\pi)\,(\alpha\, x_2 - c(x_2))]\\
&\quad + \eta\,\gamma\,[\text{anticipatory utility from learning the wage for the first day \& goal revision}]\\
&\quad + \beta\,\eta\,\gamma\,[\text{future anticipatory utility when learning the wage for the second day}].
\end{aligned}$$

The anticipatory utility from learning the wage for the first day and goal revision compares the old reference distribution for joint earnings $(\pi^2 \circ 2\, x_H; \pi(1-\pi) \circ \alpha\, x_L + x_H; \pi(1-\pi) \circ x_H + \alpha\, x_L; (1-\pi)^2 \circ 2\,\alpha\, x_L)$ with the revised one $(\pi^2 \circ x_1 + x_H; \pi(1-\pi) \circ x_1 + x_H; \pi(1-\pi) \circ x_1 + \alpha\, x_2; (1-\pi)^2 \circ x_1 + \alpha\, x_2)$. Similarly, for the costs, yielding anticipatory utility

$$\begin{aligned}
&\pi^2\,[-\lambda\,(x_H - x_1) + (c(x_H) - c(x_1))] + \pi\,(1-\pi)\,[(x_1 - \alpha\, x_L) + (c(x_L) - c(x_1))]\\
&\quad + \pi\,(1-\pi)\,[-\lambda\,(x_H + \alpha\, x_L - x_1 - \alpha\, x_2) + \underbrace{(c(x_1) + c(x_2) - c(x_L) - c(x_H))}_{=0}]\\
&\quad + (1-\pi)^2\,[(x_1 + \alpha\, x_2 - 2\,\alpha\, x_L) + (2\, c(x_L) - c(x_1) - c(x_2))]\\
&= (1 + \pi\,(\lambda - 1))\,(x_1 + (1-\pi)\,\alpha\, x_2) - (1 - \pi\,(1-\pi))\,c(x_1) - (1-\pi)^2\, c(x_2)\\
&\quad + \text{constant terms.}
\end{aligned}$$

The future anticipatory utility when learning the wage for the second day compares the old reference distribution for earnings $(\pi \circ x_H; 1-\pi \circ \alpha\, x_2)$ with the certain wage $w_s$, $s = L, H$. Similarly, for the costs, yielding expected anticipatory utility $-\pi\,(1-\pi)\,(\lambda-1)\,(x_H - \alpha\, x_2 + c(x_2) - c(x_H))$. Collecting everything and taking the derivative w.r.t. $x_1$ gives the marginal utility from a joint deviation:

$$\begin{aligned}
&\beta + \eta\,\gamma\,(1 + \pi\,(\lambda-1)) + (1-\pi)\,[\beta + \eta\,\gamma\,(1 + \pi\,(\lambda-1)) + \beta\,\eta\,\gamma\,\pi\,(\lambda-1)]\,\alpha\,\frac{d\,x_2}{d\,x_1}\\
&- [1 + \eta\,\gamma\,(1 - \pi\,(1-\pi))]\,c'(x_1) - (1-\pi)\,[\beta + \eta\,\gamma\,(1-\pi) + \beta\,\eta\,\gamma\,\pi\,(\lambda-1)]\,c'(x_2)\,\frac{d\,x_2}{d\,x_1}.
\end{aligned}$$

At the cost-matching point we have that $\frac{d\,x_2}{d\,x_1} = -\frac{c'(x_1)}{c'(x_2)} = -\frac{x_1}{x_2}$. Evaluating at $x_1 = x^N_{max,H}$ and $x_2 = x^*_{0,L} = \alpha/c$, we get the marginal utility

$$-\eta\,\gamma\,\pi\,(1-\pi)\,(\lambda-1)\,\frac{\beta + \eta\,\gamma\,(1 + \pi\,(\lambda-1))}{1 + \eta\,\gamma} < 0.$$

That is, it pays off to deviate downwards from the high-wage goal for the first day along the cost-matching curve.

## Appendix D. Bundling strategies of competing firms (Proof of Proposition 7)

We provide here the remainder of the derivations for Proposition 7 that are not given in the text. The monopolist, of course, can always induce the individual to buy the bundle by setting the individual product prices $p_1$ and $p_2$ high enough. Does such a strategy survive competitive pressure? Consider $n \geq 2$ firms in Bertrand competition. We want to check whether bundling is robust to competition (zero profits), i.e., whether it is an equilibrium that only the bundle of both products is sold at price $p_B = k_1 + k_2$ and that the price of the single products is $p_i > k_i$. Suppose $k_1 + k_2 < \min\{b_1 + b_2 - c_1, \bar{p}_B\}$. As we solely want to check whether purchasing at $p_B = k_1 + k_2$ can be an equilibrium, we fix the expectations of the consumer (he expects to buy the bundle at price $p_B$) and ask whether a firm can make a profit by offering individual products at a sufficiently low price, assuming that self 1 can sample firms' prices at no cost before making his purchase decision.[39]

Clearly, no firm on its own has an incentive to in- or decrease the price of the bundle and cannot attract self 1 to a single product by charging some other price $p_i \geq p_B$. To get the consumer to act against his expectation of buying the bundle, the price of a single product, $p_i$ must be low enough. Specifically, self 1 buys product 1 instead of the bundle if $p_1 < k_1 + k_2 - \beta\, b_2 \frac{1+\lambda\,\eta}{1+\eta} = \bar{p}_{B1}$. It only pays off for a firm to deviate and offer product 1 at this price if $p_1 > k_1$. Further, self 1 buys product 2 instead of the bundle if $p_2 < k_1 + k_2 - \left(\frac{1+\beta\,\eta\,\lambda}{1+\eta}\, b_1 - \beta\, c_1\right) = \bar{p}_{B2}$. Setting such a low price however is loss-making if

$$k_1 < b_1 \frac{1+\beta\,\eta\,\lambda}{1+\eta} - \beta\, c_1 \quad \text{and} \quad k_2 < \beta\, b_2 \frac{1+\lambda\,\eta}{1+\eta}. \tag{46}$$

## References


Abeler, Johannes, Marklein, Felix, 2010. Fungibility, Labels, and Consumption. CeDEx Discussion Paper 2010-13.

Atella, Vincenzo, Brunetti, Marianna, Maestas, Nicole, 2012. Household portfolio choices, health status and health care systems: a cross-country analysis based on SHARE. J. Bank. Finance 36, 1320–1335.

Bénabou, Roland, Tirole, Jean, 2004. Willpower and personal rules. J. Polit. Econ. 112 (4), 848–886.

Benartzi, Shlomo, Thaler, Richard H., 1995. Myopic loss aversion and the equity premium puzzle. Q. J. Econ. 110 (1), 73–92.

Bertrand, Marianne, Mullainathan, Sendhil, Shafir, Eldar, 2006. Behavioral economics and marketing in aid of decision making among the poor. J. Public Policy Mark. 25 (1), 8–23.

Brocas, Isabelle, Carrillo, Juan D., Dewatripont, Mathias, 2004. Commitment devices under self-control problems: an overview. In: Brocas, Isabelle, Carrillo, Juan D. (Eds.), Reasons and Choices. In: The Psychology of Economic Decisions, vol. 2. Oxford University Press, Oxford, pp. 49–66. Chapter 4.

Camerer, Colin, Babcock, Linda, Loewenstein, George, Thaler, Richard H., 1997. Labor supply of New York city cabdrivers: one day at a time. Q. J. Econ. 112 (2), 407–441.

Carrillo, Juan D., Dewatripont, Mathias, 2008. Promises, promises, .... Econ. J. 118, 1453–1473.

Chandon, Pierre, 2013. How package design and packaged-based marketing claims lead to overeating. Appl. Econ. Perspect. Policy 35 (1), 7–31.

Chandon, Pierre, Wansink, Brian, 2011. Is Food Marketing Making Us Fat? A Multi-disciplinary Review. Foundations and Trends in Marketing, vol. 1. Now Publishers, Delft.

Cheema, Amar, Soman, Dilip, 2008. The effect of partitions on controlling consumption. J. Mark. Res. 45 (6), 665–675.

Crawford, Vincent, Meng, Juanjuan, 2011. New York city cabdrivers' labor supply revisited: reference-dependence preferences with rational-expectations targets for hours and income. Am. Econ. Rev. 101 (5), 1912–1932.


[39] This follows Heidhues and Kőszegi (2008), who assume that (cf. their footnote 9) "at the stage when firms' prices are chosen, these prices do not influence a consumer's reference point".

Dean, Mark, Sautmann, Anja, 2015. Credit constraints and the measurement of time preferences. Working Paper. Brown University.
DellaVigna, Stefano, 2009. Psychology and economics: evidence from the field. J. Econ. Lit. 47 (2), 315–372.
Farber, Henry, 2005. Is tomorrow another day? The labor supply of New York city cabdrivers. J. Polit. Econ. 113 (1), 46–82.
Fehr, Ernst, Götte, Lorenz, 2007. Do workers work more if wages are high? Evidence from a randomized field experiment. Am. Econ. Rev. 97 (1), 298–317.
Fishbach, Ayelet, Zhang, Ying, 2008. Together or apart: when goals and temptations complement versus compete. J. Pers. Soc. Psychol. 94 (4), 547–559.
Fudenberg, Drew, Levine, David K., 2006. A dual-self model of impulse control. Am. Econ. Rev. 96 (5), 1449–1476.
Fudenberg, Drew, Levine, David K., 2011. Risk, delay, and convex self-control costs. Am. Econ. J. Microecon. 3 (3), 34–68.
Giné, Xavier, Goldberg, Jessica, Silverman, Dan, Yang, Dean, 2012. Revising commitments: field evidence on the adjustment of prior choices. Working Paper. National Bureau of Economic Research.
Gourville, John, Soman, Dilip, 2002. Pricing and the psychology of consumption. Harv. Bus. Rev. 80 (9), 90–96.
Guiso, Luigi, Paiella, Monica, 2008. Risk aversion, wealth, and background risk. J. Eur. Econ. Assoc. 6 (6), 1109–1150.
Harrison, Glenn W., Lau, Morten I., Rutström, Elisabet, 2005. Dynamic consistency in Denmark: a longitudinal field experiment. Working Paper 2005-09. Durham University.
Heath, Chip, Soll, Jack B., 1996. Mental budgeting and consumer decisions. J. Consum. Res. 23 (1), 40–52.
Heath, Chip, Soll, Jack B., Larrick, Richard P., Wu, George, 1999. Goals as reference points. Cogn. Psychol. 38, 79–109.
Heidhues, Paul, Kőszegi, Botond, 2008. Competition and price variation when consumers are loss averse. Am. Econ. Rev. 98 (4), 1245–1268.
Herweg, Fabian, Müller, Daniel, Weinschenk, Philipp, 2010. Binary payment schemes: moral hazard and loss aversion. Am. Econ. Rev. 100 (5), 2451–2477.
Hollenbeck, John R., Williams, Charles R., Klein, Howard J., 1989. An empirical examination of the antecedents of commitment to difficult goals. J. Appl. Psychol. 74 (1), 18–23.
Hsiaw, Alice, 2015. Goal bracketing and self-control. Working Paper. Brandeis University.
Hsiaw, Alice, 2013. Goal-setting and self-control. J. Econ. Theory 148 (2), 601–626.
Kahneman, Daniel, Tversky, Amos, 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–291.
Kahneman, Daniel, Tversky, Amos, 1984. Choices, values, and frames. Am. Psychol. 39 (4), 341–350.
Kahneman, Daniel, Tversky, Amos, Lovallo, Dan, 1993. Timid choices and bold forecasts: a cognitive perspective on risk taking. Manag. Sci. 39 (1), 17–31.
Khan, Uzma, Dhar, Ravi, 2007. Where there is a way, is there a will? The effect of future choices on self-control. J. Exp. Psychol. 136 (2), 277–288.
Kőszegi, Botond, Rabin, Matthew, 2006. A model of reference-dependent preferences. Q. J. Econ. 121 (4), 1133–1165.
Kőszegi, Botond, Rabin, Matthew, 2007. Reference-dependent risk attitudes. Am. Econ. Rev. 97 (4), 1047–1073.
Kőszegi, Botond, Rabin, Matthew, 2009. Reference-dependent consumption plans. Am. Econ. Rev. 99 (3), 909–936.
Koch, Alexander K., Nafziger, Julia, 2009. Motivational goal bracketing. IZA Discussion Paper No. 4471.
Koch, Alexander K., Nafziger, Julia, 2011a. Goals and psychological accounting. IZA Discussion Paper No. 5802.
Koch, Alexander K., Nafziger, Julia, 2011b. Self-regulation through goal setting. Scand. J. Econ. 113 (1), 212–227.
Koch, Alexander K., Nafziger, Julia, 2016. Motivational effects of goal bracketing: an experiment. Working Paper. Aarhus University.
Krupka, Erin L., Stephens, Melvin, 2013. The stability of measured time preferences. J. Econ. Behav. Organ. 85, 11–19.
Laibson, David, 1997. Golden eggs and hyperbolic discounting. Q. J. Econ. 112 (2), 443–477.
Locke, Edwin A., Latham, Gary P., 2002. Building a practically useful theory of goal setting and task motivation. Am. Psychol. 57 (9), 705–717.
Loewenstein, George, 1999. Because it is there: the challenge of mountaineering for utility theory. Kyklos 52 (3), 315–343.
Love, David A., Smith, Paul A., 2010. Does health affect portfolio choice? Health Econ. 19 (2), 1441–1460.
Meier, Stephan, Sprenger, Charles, 2015. Temporal stability of time preferences. Rev. Econ. Stat. 97 (2), 273–286.
O'Donoghue, Ted, Rabin, Matthew, 1999. Doing it now or later. Am. Econ. Rev. 89 (1), 103–124.
Pope, Devin G., Schweitzer, Maurice E., 2011. Is tiger woods loss averse? Persistent bias in the face of experience, competition, and high stakes. Am. Econ. Rev. 101 (1), 129–157.
Prelec, Drazen, Loewenstein, George, 1998. The red and the black: mental accounting of savings and debt. Mark. Sci. 17 (1), 4–28.

Rabin, Matthew, Weizsäcker, Georg, 2009. Narrow bracketing and dominated choices. Am. Econ. Rev. 99 (4), 1508–1543.

Read, Daniel, Loewenstein, George, Rabin, Matthew, 1999. Choice bracketing. J. Risk Uncertain. 19 (1–3), 171–197.

Shefrin, Hersh M., Thaler, Richard H., 1988. The behavioral life-cycle hypothesis. Econ. Inq. 26 (4), 609–643.

Soman, Dilip, 2004. Framing, loss aversion, and mental accounting. In: Koehler, Derek J., Harvey, Nigel (Eds.), Blackwell Handbook of Judgment and Decision Making. Blackwell Publishing, Oxford, UK, pp. 379–398. Chapter 19.

Soman, Dilip, Cheema, Amar, 2011. Earmarking and partitioning: increasing saving by low-income households. J. Mark. Res. 48, S14–S22.

Soman, Dilip, Gourville, John T., 2001. Transaction decoupling: how price bundling affects the decision to consume. J. Mark. Res. 38 (1), 30–44.

Suvorov, Anton, van de Ven, Jeroen, 2008. Goal setting as a self-regulation mechanism. Working Paper w0122. Center for Economic and Financial Research.

Tanaka, Tomomi, Camerer, Colin F., Nguyen, Quang, 2010. Risk and time preferences: linking experimental and household survey data from Vietnam. Am. Econ. Rev. 100 (1), 557–571.

Thaler, Richard, 1999. Mental accounting matters. J. Behav. Decis. Mak. 12, 183–206.

Thaler, Richard, 1985. Mental accounting and consumer choice. Mark. Sci. 4 (3), 199–214.

Thaler, Richard, Tversky, Amos, Kahneman, Daniel, Schwartz, Alan, 1997. The effect of myopia and loss aversion on risk taking: an experimental test. Q. J. Econ. 112 (2), 647–661.

Thaler, Richard, Shefrin, Hersh, 1981. An economic theory of self-control. J. Polit. Econ. 89 (2), 392–406.

Tversky, Amos, Kahneman, Daniel, 1981. The framing of decisions and the psychology of choice. Science 211 (4481), 453–458.

UK National Health Service, 2012. Health service, "Losing weight—getting started". http://www.nhs.uk/Tools/Documents/WEIGHT-LOSS-PACK/all-weeks.pdf.

Zhang, Ying, Fishbach, Ayelet, Dhar, Ravi, 2007. When thinking beats doing: the role of optimistic expectations in goal-based choice. J. Consum. Res. 34 (4), 567–578.