**Online Supplements – One size fits all? Designing financial incentives tailored to individual economic preferences.**

This online supplementary file contains all additional details related to study. It contains the following Appendices:

1. Experimental flowchart, instructions, choice task, definitions used to elicit economic preferences.
2. Post-hoc power analysis and interpretation
3. Supplementary regression results for additional control variables and interactions

**Online Appendix A: Experimental flowchart, instructions, choice tasks and definitions for economic preferences**

Figure A1 shows the outline of the experiment, and the measures included in this study. The remainder of this Appendix will provide screenshots and instructions for each part of the experiment.

**Figure A1:** Experimental Flowchart



**A.1. General introduction**

Students were welcomed by the experimenter and received a short oral introduction into the goals of the study. They were told that they could ask any question that they had, and that there were no right or wrong answers.

**A.2. Tool for tailored incentives**

A direct link to this tool can be found here: <https://referencepoints.shinyapps.io/Minecentive/>

The following instruction was used: *‘Please imagine the following situation: you have set yourself the goal of losing weight, so you decided to get a gym membership. Now, your employer wants to help you to lose weight. This may decrease your chances of taking up sick leave and increase your overall wellbeing. As such, your employer has offered to pay you a financial reward if you use the gym at least three times each week for a 10 week period. Your employer is quite flexible, and besides the expected pay-out has no preference in how your financial reward is structured. Obviously, you yourself know best what kind of pay-out structure would motivate you to go to the gym and reach your goal of losing weight. Therefore, we ask you to indicate how you would like your pay-out(s) to be structured.’*

Next, they could tailor their incentives in a menu, with a separate interactive heading for each of the incentive dimensions. Figures A2.1 to A2.4 show the choice options. Subject were given feedback of their selected incentive in the same panel (see Figure A2.5)

**Figure A2.1***: Pre-commitment dimension question*



**Figure A2.2**. *Pay-out frequency*



**Figure 2.3**. *Pay-out structure*



**Figure 2.4.** *Chance of winning (risk)*



**Figure A2.5**. *Interactive feedback panel for tailored incentive*



**A.3. Economic preference elicitation**

Economic preferences were elicited in three parts in a chained procedure, i.e. responses from each part carried on to the next. The first two parts, aimed at measuring loss aversion, utility curvature and probability weighting were based on the non-parametric method developed by Abdellaoui et al. (2016). A full justification of their approach can be found in the original paper, including the notational conventions, and theoretical assumptions needed to arrive at these elicitations. Throughout, each elicitation consisted of a bi-sectional approach with 4 choices followed by a slider that allowed respondents to modify and confirm their elicited indifference (see Figure A3.1. for an example). Furthermore, throughout this Appendix, all elicitation start with gains first and losses after, while in reality this was counterbalanced between respondents. Throughout this Appendix, we let $≻,≽,∼$ represent strict preference, weak preference and indifference respectively.

*A.3.1. Loss aversion and utility curvature*

*A.3.1.1. Loss aversion*

Loss aversion is elicited by eliciting three indifferences, that link gains and losses together. First, we elicit an indifference $g\_{p}l\~x\_{0},$ where $x\_{0}$ is the reference-point (set at 0), *p* is a probability that is kept constant throughout this first, and $g$ is a gain of €50. We elicit a loss $l$ (e.g. -25€). The next two indifferences involve certainty equivalence elicitation for outcome $g$ and $l,$ i.e. a certain outcome ($x\_{1}^{+}$ for gains, $x\_{1}^{-}$ for losses) that makes one indifferent between receiving$ g$ or $l$ with probability $p$ or $x\_{0}$ otherwise. These indifferences are denoted $x\_{1}^{+}\~g\_{p}x\_{0}$ and $x\_{1}^{-}\~l\_{p}x\_{0}$. [Abdellaoui et al. (2016](#_ENREF_2)) show that loss aversion (denoted as index $λ)$, as defined by Köbberling and Wakker (2005), can be derived by: $λ= {x\_{1}^{+}}/{-x\_{1}^{-}}$, where respondents with $λ>1, λ=1, λ<1 $are loss averse, loss neutral or gain seeking respectively.

**Figure A3.1.** Example of elicitation procedure for indifference $g\_{p}l\~x\_{0}$, where a respondent is indifferent for $l=25$



 *A.3.1.2. Utility curvature*

Now that we have linked together gains and losses compared to the reference-point $x\_{0},$ we can elicit a series of indifference to estimate utility curvature for both gains and losses separately. In Abdellaoui et al. (2016) this process is based on the trade-off method developed by Wakker and Deneffe (1996), see Figure A3.2. That is, a standard sequence of outcomes is elicited for gains and losses, or in other words a sequence of outcomes spaced equally in terms of utility. This standard sequence elicitation is set-up in the same way for gains and losses. For gains, it starts by fixing a small loss $l$ (in this study fixed at: -10€), and eliciting a larger loss $L$ in the following indifference: $x\_{1}^{+}\_{p}L\~l\_{p}x\_{0}$. These two loss amounts serve as offset losses, in the standard sequence elicitation for gains. Next, the equally-spaced outcomes in the standard sequence are elicited by eliciting $x\_{2}^{+}$ in the following indifference : $x\_{2}^{+}\_{p}L\~l\_{p}x\_{1}^{+}$. This process (i.e. $x\_{j}^{+}\_{p}L\~x\_{j-1}^{+}\_{p}l, j=2,…4)$ is applied 3 times, yielding a standard sequence with 5 data points ($x\_{0},x\_{1}^{+},x\_{2}^{+},$ $x\_{3}^{+},$ and $x\_{4}^{+}$). For losses, a small gain (in this study fixed at: 10€) is fixed, to elicit a larger loss $G$, in the following indifference: $G\_{p}x\_{1}^{-}\~g\_{p}x\_{0}$. Next, again a series of indifferences of the form $G\_{p}x\_{j}^{-}\~g\_{p}x\_{j-1}^{-}$, $j=2,…,4 $is elicited, which yields a standard sequence for losses with 5 data points ($x\_{0},x\_{1}^{-},x\_{2}^{-},$ $x\_{3}^{-},$ and $x\_{4}^{-}$).

**Figure A3.1.** Example visual representation of elicitation procedure for utility curvature (indifference $x\_{1}^{+}\_{p}L\~l\_{p}x\_{0}, with x\_{1}^{+}=25$)



To calculate the utility curvature of the utility function for monetary gains $U^{+}(∙)$ or losses$ U^{-}(∙)$, we apply the following scaling: $U^{+}\left(x\_{0}\right)=0, U^{-}\left(x\_{4}^{-}\right)=-1$ and $U^{+}\left(x\_{4}^{-}\right)=1$. Furthermore, as is usual in these type of studies, monetary outcomes are normalized such that each outcome is dived by the highest outcome in its’ respective domain, i.e. $x\_{0},x\_{1}^{-},x\_{2}^{-},$ $x\_{3}^{-},$ $x\_{4}^{-}/ x\_{4}^{-}$ and $x\_{0},x\_{1}^{+},x\_{2}^{+},$ $x\_{3}^{+},$ $x\_{4}^{+}/ x\_{4}^{+}$. Although the non-parametric method allows non-parametric estimation of utility curvature, in this study the most commonly used power utility family is used estimated by non-linear least squares. This allows the estimates to be compared with earlier work. For this family, $U^{+}\left(x\right) $=$ x^{α}$, where represents the utility function over monetary outcomes. For losses, this is estimated by $U^{-}\left(x\right) $=$-(- (x)^{α}$ with $α>0. $For gains [losses], $α>1$ corresponds to convex [concave] utility, $α=1$ corresponds to linear utility, and $α<1$ corresponds to concave [convex] utility.

*A.3.2. Probability weighting functions*

As in Lipman and colleagues (2019), probability weighting is elicited with the method developed by Abdellaoui (2000). This method was employed as follows: to the probability weighting functions $w^{+}(p)$ and $w^{-}(p)$, the certainty equivalents $x\_{p}^{+}$ and $x\_{p}^{-}$ of the prospects $x\_{4\_{ p}}^{+}x\_{0}$ and $x\_{4\_{ p}}^{-}x\_{0}$, for the following probabilities: $p= 0.1, 0.3, 0.5, 0.7, 0.9$.The outcomes $x\_{4}^{+}$ and $x\_{4}^{-}$ are the maximum (minimum) outcome elicited in the standard sequence. Therefore, it follows from the chosen scaling of utility that $U\left(x\_{P}^{+}\right)=w^{+}(p)$ and $-U^{-}\left(x\_{p}^{-}\right)=w^{-}(p)$. The values of $U^{+}\left(x\_{p}^{+}\right)$ and $LU^{-}\left(x\_{p}^{-}\right)$ are interpolated from their respective standard sequences. Figure A3.3 shows an example of a gamble scenario for gains.

**Figure A3.3.** Example visual representation of choice options used for eliciting probability weighting for gains ($x\_{p}^{+}\~x\_{4\_{ p}}^{+}x\_{0}), with x\_{4}^{+}=100, p=0.3$.



To summarize the shape of the weighting functions Tversky and Kahneman’s one-parameter inverse S-shaped probability weighting function is used, i.e. $w^{i}\left(p\right)=p^{γ}/(p^{γ}+(1-p)^{γ})^{1/γ}$ with $i=+,-$. Again, this is estimated by nonlinear least squares*.* The $γ$-parameter controls for the shape of the probability weighting function. If $γ=1$ there is no probability transformation and $w^{i}\left(p\right)=p$. However, if $γ<1$, decision makers underweight large probabilities and overweight small probabilities. This corresponds to the commonly found inverse S-shaped weighting function. If $γ>1$, the opposite pattern holds, corresponding to an S-shaped weighting function.

*A.3.3. Present bias & Discounting*

Present bias and discounting were elicited by means of the approach of Laibson (1997). This model assumes the discounted utility model, i.e. utility for timed outcome $(x,t)$ can be evaluated by $DU\left(x,t\right)=D\left(t\right)U(x)$, where $D\left(⋅\right)$ refers to the discounting function. To reflect the sign-dependent nature of this experiment, we modify this to: $U\left(x,t\right)=D^{i}\left(t\right)U^{i}(x)$, with $i=+$ for gains and $i=-$ for losses. In the quasi-hyperbolic discounting model, $D^{i}\left(t\right)=β^{i}/(1+r^{i})^{t}$, with $i=+,- $for gains and losses respectively, $0<β\leq 1$ for $t>0$ and $D\left(t\right)=$1 otherwise, and $r$ reflects the per-period discount rate. In this part of the elicitation $x\_{T}^{+}$ and $x\_{T}^{-} $ refer to the highest outcome in the standard sequences for gains and losses divided by 2, i.e.$ x\_{4}^{+}/2$ and $x\_{4}^{-}/2$. These outcomes were divided by half to decrease the chances extrapolation beyond the measured standard sequence was necessary. $β$ and $r$ were elicited for gains and losses by means of the following indifferences (see Figure A3.4 for an example of visual stimuli used),

$$\left(x\_{T}^{+},0 weeks\right) \~ \left(y\_{T}^{+},5 weeks\right)$$

$$\left(x\_{T}^{+},5 weeks\right) \~ \left(z\_{T}^{+},10 weeks\right)$$

And for losses by:

$$\left(x\_{T}^{-},0 weeks\right) \~ \left(y\_{T}^{-},5 weeks\right)$$

$$\left(x\_{T}^{-},5 weeks\right) \~ \left(z\_{T}^{-},10 weeks\right)$$

**Figure A3.4.** Visual representation of choice options used for eliciting present bias and discounting for losses.



In both cases, we can evaluate these indifferences as:

$U^{i}(x\_{T}^{i})D^{i}\left(0\right)= U^{i}(y\_{T}^{i})D^{i}\left(5\right) ⟺ U^{i}(x\_{T}^{i})(\frac{1}{\left(1+r^{i}\right)^{1}})=U^{i}(y\_{T}^{i})(\frac{β^{i}}{\left(1+r^{i}\right)^{5}})$

$U^{i}\left(x\_{T}^{i}\right) D^{i}\left(5\right)= U^{i}(z\_{T}^{i})D^{i}\left(10\right) ⟺ U^{i}(x\_{T}^{i})(\frac{β^{i}}{\left(1+r^{i}\right)^{5}})=U^{i}(y\_{T}^{i})(\frac{β^{i}}{\left(1+r^{i}\right)^{10}})$

After rearranging the second indifference we find:

$$r^{i}=\frac{1}{\left(\frac{U^{i}(z\_{T}^{i})}{U^{i}(x\_{T}^{i})}\right)^{1/5}}$$

$ $

And after we have determined $r^{i}$, $β $ is found by:

$$β^{i}=\frac{U^{i}\left(y\_{T}^{i}\right)}{U^{i}(x\_{T}^{i})\frac{1}{\left(1+r^{i}\right)^{1/5}}}$$

**A5. Psychological measures**

After asking respondents to self-report on several health-related characteristics, total of three psychological measures were used, which are reprinted in this Appendix. These questionnaires measured self-control (Tangney et al., 2018), cognitive reflection (Toplak et al., 2011), and personality (Francis et al., 1992).

*A.4.1. Self-reported health behavior*

The following demographics were collected:

*A.4.1. Trait self-control questionnaire*

The questionnaire was adapted from Tangney et al. (2018), and measures self-control as a trait, i.e. the degree to which individuals in general are able to self-regulate. Items marked with \* require reverse coding, and it is reported as a mean in the main text.

|  |
| --- |
| *The following statements may reflect how you perceive yourself. Please indicate below to what extent these statements reflect how you typically are, by circling the answer that applies.* |
|  |  ***Not at all Very much*** ***↓ ↓***  |
| *1* | *I am good at resisting temptation* | *1* | *2* | *3* | *4* | *5* |  |
| *2\** | *I have a hard time breaking bad habits.* | *1* | *2* | *3* | *4* | *5* |  |
| *3\** | *I am lazy* | *1* | *2* | *3* | *4* | *5* |  |
| *4\** | *I say inappropriate things.* | *1* | *2* | *3* | *4* | *5* |  |
| *5\** | *I do certain things that are bad for me, if they are fun.* | *1* | *2* | *3* | *4* | *5* |  |
| *6* | *I refuse things that are bad for me.* | *1* | *2* | *3* | *4* | *5* |  |
| *7\** | *I wish I had more self-discipline.* | *1* | *2* | *3* | *4* | *5* |  |
| *8* | *People would say that I have iron self- discipline.* | *1* | *2* | *3* | *4* | *5* |  |
| *9\** | *Pleasure and fun sometimes keep me from getting work done.* | *1* | *2* | *3* | *4* | *5* |  |
| *10\** | *I have trouble concentrating.* | *1* | *2* | *3* | *4* | *5* |  |
| *11\** | *I am able to work effectively toward long-term goals.* | *1* | *2* | *3* | *4* | *5* |  |
| *12\** | *Sometimes I can’t stop myself from doing something, even if I know it is wrong.* | *1* | *2* | *3* | *4* | *5* |  |
| *13\** | *I often act without thinking through all the alternatives.* | *1* | *2* | *3* | *4* | *5* |  |

*A.4.2. Cognitive reflection task (CRT)*

This three item task developed by Toplak et al. (2011) aims to quantify the degree to which individuals rely on their automatic system by asking questions which seems to have an immediate, simple and right answer, which only after reflecting on it for some time appears to be in fact *wrong.* The CRT is scored as the amount of correct answers. The questions were answered by with a pen by writing down the answer on the open space.

1. *A bat and a ball cost €1.10 in total. The bat costs €1.00 more than the ball. How much does the ball cost? \_\_\_\_\_\_\_\_ cents*
2. *If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? \_\_\_\_\_\_\_ minutes*
3. *In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take*

*A.4.3. Personality questionnaire*

The last questionnaire used is a revised short-form version of the Revised Eysenck Personality Questionnaire, which captures personality on 4 domains. Items 1, 10, 12, 15, 19 and 22 capture Neuroticism. Items 2, 4, 14, 16, 21 and 24 capture Extraversion. Items 3, 6, 9, 13, 17 and 23 capture Psychoticism, and finally, items 5, 7, 8, 11, 18, and 20 capture Social desirability. Items marked with \* are recoded, meaning that code 1 means has characteristic related to personality dimension and code 0 means does not relate to that dimension. Means are reported in main text.

*Please answer the following questions by answering “Yes” or “No” (circle which applies). There are no right or wrong answers. It is not necessary to think very long about these questions.*

|  |  |  |  |
| --- | --- | --- | --- |
| *1* | *Does your mood often go up and down?* | *Yes* | *No* |
| *2* | *Are you a talkative person?* | *Yes* | *No* |
| *3\** | *Would being in debt worry you?* | *Yes* | *No* |
| *4* | *Are you rather lively?* | *Yes* | *No* |
| *5\** | *Were you ever greedy by helping yourself to more than your share of anything?* | *Yes* | *No* |
| *6* | *Would you take drugs which may have strange or dangerous effects?* | *Yes* | *No* |
| *7\** | *Have you ever blamed someone for doing something you knew was really your fault?* | *Yes* | *No* |
| *8* | *Do you always practice what you preach?* | *Yes* | *No* |
| *9* | *Do you prefer to go your own way rather than act by the rules?* | *Yes* | *No* |
| *10* | *Do you often feel ‘fed-up’?* | *Yes* | *No* |
| *11\** | *Have you ever taken anything (even a pin or button) that belonged to someone else?* | *Yes* | *No* |
| *12* | *Would you call yourself a nervous person?* | *Yes* | *No* |
| *13* | *Do you think marriage is old-fashioned and should be done away with?* | *Yes* | *No* |
| *14* | *Can you easily get some life into a rather dull party?* | *Yes* | *No* |
| *15* | *Are you a worrier?* | *Yes* | *No* |
| *16\** | *Do you tend to keep in the background on social occasions?* | *Yes* | *No* |
| *17\** | *Does it worry you if you know there are mistakes in your work?* | *Yes* | *No* |
| *18\** | *Have you ever cheated at a game?* | *Yes* | *No* |
| *19* | *Do you suffer from ‘nerves’?* | *Yes* | *No* |
| *20\** | *Have you ever taken advantage of someone?* | *Yes* | *No* |
| *21\** | *Are you mostly quiet when you are with other people?* | *Yes* | *No* |
| *22* | *Do you often feel lonely?* | *Yes* | *No* |
| *23\** | *Is it better to follow society’s rules than go your own way?* | *Yes* | *No* |
| *24* | *Do other people think of you as being very lively* | *Yes* | *No* |

**Appendix B. Post-hoc power analysis and interpretation of effect sizes**

This Appendix will reflect on the statistical power of the experiment reported. Seeing as this is the first study that aims to study the association between tailored incentives and behavioral insights (obtained as stated preferences), no effect sizes were available for a-priori power analysis. Furthermore, a-priori power analyses are often performed for studies with random assignment, which ensures that samples are independent and gives the researcher control over the number of respondents in each group. In this study, however, it is questionable if samples are independent, and barely any work existed to predict how individuals self-select incentives. Hence, in my view, only post-hoc power analyses were available, which I report in this appendix.

For simplicity, I will base power analysis on independent samples *T* tests with 2 samples, as this will easily allow me to demonstrate the lowest mean differences this study was powered to detect. This means that for each incentive dimension I divide the self-selected incentives into 2 samples, that I treat as being independent. The lack of significant results in my study indicates that for most measures my mean differences were lower. If this study is underpowered, it runs the risk of unjustly categorizing these small differences as error rather resulting from distributions with a true mean difference. Any study has such a risk of such a Type II error, which is captured by (1-$β$), where $β$ reflects the desired test power, which I set at 0.8, as is suggested by Cohen (1988). However, it is impossible to determine if such an error actually occurred. Hence, whether or not this study is sufficiently powered is a value judgment to be made by reflecting on the economic significance of the lowest mean differences this study is able to detect. If one believes that mean differences smaller than this study was able to detect are economically significant (i.e. are meaningful in terms of stated preferences or real decisions), this study is underpowered. If on the other hand, one believes mean differences of the order of magnitude reported in Table B1 to Table B4 to be negligible, this study was adequately powered.

First, the sample and its’ empirical distribution over the tailored incentive dimensions is used to determine the lowest effect size this study is powered to observe (Table B1-B4, *Cohen’s d*). For Pre-commitment and Weekly, which had two answer categories, this will be based on an independent samples *t* test. The minimal effect size is a function of the size of each group (Table B1-B4, *n1* and *n2*), significance level ($∝=0.05$), and the preferred statistical power ($β$=0.8). This approach can also be applied to Structure and Risk, but this first requires dichotomizing these dimensions. Structure is dichotomized as those selecting those who chose constant (32%) or one of the non-constant payment structures (68%). Risk will be dichotomized as those who chose certain pay-outs (48%) and those who introduced some risk (52%).

Cohen’s d is calculated as follows: $d=M2-M1/(SD\_{pooled})$., i.e. the mean difference divided by the pooled standard deviation. Hence, for each observation, I calculated the pooled standard deviation, and multiply by Cohen’s *d* to obtain the smallest parameter difference this study was powered to observe. Table B1 to B4 report the outcomes of this analysis (and the inputs used). It shows that Cohen’s d ranged from 0.42 to 0.51, which are between small and medium effect sizes using Cohen’s (1988) interpretation. Whether or not the lowest mean difference this study was powered to observe includes economically significant differences is ultimately a value judgement. My conclusion is that due to the high heterogeneity observed for the economic preferences elicited, the study runs some risk of being underpowered. For example, for effects of discounting for gains and losses the lowest mean differences this study was powered to observe was 0.05 and 0.03 respectively. Given that these are *weekly* discount rates, differences much smaller than these could very well be economically significant. However, given the exploratory nature of this study, many measures were collected, and none appeared to significantly and systematically predict tailored incentives. Hence, it is straightforward to see that it is unlikely that all of these null-results (i.e. the small actual differences reported in Table B1-B4) are Type II errors. In fact, without correcting for multiple testing it would have been far more likely to have found significant but non-existing effects (i.e. Type I errors).

**Table B1.** Post-hoc power analysis for Pre-commit dimension of tailored incentives

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Dimension: Pre-commitment |  |  |  |  |  |
|  | ***N1*** | ***N2*** | **Cohen’s *d*** | ***SD1*** | ***SD2*** | $$SD\_{pooled}$$ | **Lowest difference** | **Actual difference** |
| Loss aversion ($λ$) | 59 | 123 | 0.45 | 3.15 | 2.22 | 2.73 | 1.22 | 0.65 |
| Utility curvature ($α$) - gains | 59 | 123 | 0.45 | 1.53 | 3.07 | 2.43 | 1.08 | -0.24 |
| Utility curvature ($α$) - losses | 59 | 123 | 0.45 | 4.72 | 3.93 | 4.34 | 1.94 | 0.29 |
| Probability weighting ($γ$) - gains | 59 | 123 | 0.45 | 2.67 | 2.05 | 2.38 | 1.06 | 0.05 |
| Probability weighting ($γ$) - losses | 59 | 123 | 0.45 | 2.89 | 2.32 | 2.62 | 1.17 | 0.51 |
| Present Bias ($β$) - gains | 58 | 123 | 0.45 | 0.18 | 0.33 | 0.27 | 0.12 | -0.05 |
| Present Bias ($β$) - losses | 59 | 123 | 0.45 | 0.18 | 0.21 | 0.2 | 0.09 | 0 |
| Discounting ($δ$) - gains | 59 | 123 | 0.45 | 0.05 | 0.14 | 0.1 | 0.05 | -0.02 |
| Discounting ($δ$) - losses | 59 | 123 | 0.45 | 0.04 | 0.08 | 0.07 | 0.03 | -0.02 |
| Age | 59 | 123 | 0.45 | 1.45 | 1.49 | 1.47 | 0.66 | 0 |
| Cigarettes (per week) | 59 | 123 | 0.45 | 2.82 | 2.57 | 2.7 | 1.2 | 0.04 |
| BMI | 59 | 123 | 0.45 | 5.98 | 3.13 | 4.78 | 2.13 | -0.35 |
| Alcohol (glasses/week) | 59 | 123 | 0.45 | 8.34 | 9.93 | 9.17 | 4.09 | -0.35 |
| Exercise (days/week) | 59 | 123 | 0.45 | 1.74 | 1.64 | 1.69 | 0.75 | -0.19 |
| Trait self-control  | 55 | 107 | 0.47 | 0.59 | 0.57 | 0.58 | 0.27 | 0.03 |
| Cognitive reflection  | 50 | 96 | 0.49 | 1.16 | 1.16 | 1.16 | 0.57 | 0.13 |
| EPQ - Neuroticism  | 45 | 90 | 0.52 | 0.22 | 0.18 | 0.2 | 0.11 | -0.06 |
| EPQ - Extraversion  | 45 | 90 | 0.52 | 0.19 | 0.17 | 0.18 | 0.09 | -0.01 |
| EPQ - Psychoticism  | 45 | 90 | 0.52 | 0.2 | 0.18 | 0.19 | 0.1 | -0.04 |
| EPQ – Social desirability  | 45 | 90 | 0.52 | 0.23 | 0.22 | 0.23 | 0.12 | 0.05 |

**Table B2.** Post-hoc power analysis for Weekly dimension of tailored incentives

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dimension: Weekly** |  |  |  |  |  |
|  | ***N1*** | ***N2*** | **Cohen’s *d*** | ***SD1*** | ***SD2*** | $$SD\_{pooled}$$ | **Lowest difference** | **Actual difference** |
| **Loss aversion (**$λ$**)** | 86 | 96 | 0.42 | 2.67 | 2.49 | 2.58 | 1.08 | 0.08 |
| **Utility curvature (**$α$**) - gains** | 86 | 96 | 0.42 | 3.79 | 0.69 | 2.73 | 1.14 | 0.7 |
| **Utility curvature (**$α$**) - losses** | 86 | 96 | 0.42 | 0.81 | 5.7 | 4.07 | 1.7 | -0.88 |
| **Probability weighting (**$γ$**) - gains** | 86 | 96 | 0.42 | 2.86 | 1.52 | 2.29 | 0.96 | 0.47 |
| **Probability weighting (**$γ$**) - losses** | 86 | 96 | 0.42 | 1.61 | 3.09 | 2.47 | 1.03 | -0.67 |
| **Present Bias (**$β$**) - gains** | 86 | 95 | 0.42 | 0.34 | 0.24 | 0.3 | 0.12 | 0.04 |
| **Present Bias (**$β$**) - losses** | 86 | 96 | 0.42 | 0.23 | 0.18 | 0.2 | 0.08 | -0.06 |
| **Discounting (**$δ$**) - gains** | 86 | 96 | 0.42 | 0.13 | 0.1 | 0.12 | 0.05 | 0.02 |
| **Discounting (**$δ$**) - losses** | 86 | 96 | 0.42 | 0.06 | 0.09 | 0.07 | 0.03 | 0 |
| **Age** | 86 | 96 | 0.42 | 1.38 | 1.56 | 1.47 | 0.62 | -0.04 |
| **Cigarettes (per week)** | 86 | 96 | 0.42 | 2.73 | 2.57 | 2.65 | 1.11 | 0.39 |
| **BMI** | 86 | 96 | 0.42 | 5.3 | 2.98 | 4.3 | 1.8 | 1.08 |
| **Alcohol (glasses/week)** | 86 | 96 | 0.42 | 10.03 | 8.89 | 9.48 | 3.96 | 0.58 |
| **Exercise (days/week)** | 86 | 96 | 0.42 | 1.64 | 1.69 | 1.67 | 0.7 | -0.19 |
| **Trait self-control**  | 81 | 81 | 0.44 | 0.56 | 0.59 | 0.57 | 0.25 | -0.13 |
| **Cognitive reflection**  | 74 | 72 | 0.47 | 1.12 | 1.17 | 1.15 | 0.54 | 0.33 |
| **EPQ - Neuroticism**  | 70 | 65 | 0.49 | 0.21 | 0.18 | 0.2 | 0.1 | -0.05 |
| **EPQ - Extraversion**  | 70 | 65 | 0.49 | 0.18 | 0.18 | 0.18 | 0.09 | 0 |
| **EPQ - Psychoticism**  | 70 | 65 | 0.49 | 0.18 | 0.19 | 0.19 | 0.09 | -0.03 |
| **EPQ – Social desirability**  | 70 | 65 | 0.49 | 0.22 | 0.24 | 0.23 | 0.11 | 0.01 |

**Table B3.** Post-hoc power analysis for Weekly dimension of tailored incentives

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dimension: Sequence** |  |  |  |  |  |  |  |  |
|  | ***N1*** | ***N2*** | **Cohen’s *d*** | ***SD1*** | ***SD2*** | $$SD\_{pooled}$$ | **Lowest difference** | **Actual difference** |
| **Loss aversion (**$λ$**)** | 117 | 65 | 0.44 | 3.05 | 1.3 | 2.35 | 1.02 | 0.47 |
| **Utility curvature (**$α$**) - gains** | 117 | 65 | 0.44 | 3.28 | 0.74 | 2.37 | 1.03 | 0.51 |
| **Utility curvature (**$α$**) - losses** | 117 | 65 | 0.44 | 5.16 | 1.06 | 3.73 | 1.62 | 0.5 |
| **Probability weighting (**$γ$**) - gains** | 117 | 65 | 0.44 | 2.53 | 1.69 | 2.15 | 0.94 | 0.24 |
| **Probability weighting (**$γ$**) - losses** | 117 | 65 | 0.44 | 2.03 | 3.2 | 2.68 | 1.17 | -0.64 |
| **Present Bias (**$β$**) - gains** | 117 | 64 | 0.44 | 0.32 | 0.25 | 0.28 | 0.12 | 0 |
| **Present Bias (**$β$**) - losses** | 117 | 65 | 0.44 | 0.2 | 0.2 | 0.2 | 0.09 | -0.04 |
| **Discounting (**$δ$**) - gains** | 117 | 65 | 0.44 | 0.14 | 0.06 | 0.11 | 0.05 | 0.03 |
| **Discounting (**$δ$**) - losses** | 117 | 65 | 0.44 | 0.05 | 0.1 | 0.08 | 0.03 | -0.01 |
| **Age** | 117 | 65 | 0.44 | 1.41 | 1.6 | 1.51 | 0.66 | 0.1 |
| **Cigarettes (per week)** | 117 | 65 | 0.44 | 2.8 | 2.37 | 2.59 | 1.13 | 0.26 |
| **BMI** | 117 | 65 | 0.44 | 4.65 | 3.37 | 4.06 | 1.77 | 1.04 |
| **Alcohol (glasses/week)** | 117 | 65 | 0.44 | 9.91 | 8.55 | 9.26 | 4.03 | 0.28 |
| **Exercise (days/week)** | 117 | 65 | 0.44 | 1.63 | 1.75 | 1.69 | 0.74 | -0.2 |
| **Trait self-control**  | 108 | 54 | 0.47 | 0.57 | 0.59 | 0.58 | 0.27 | -0.03 |
| **Cognitive reflection**  | 97 | 49 | 0.49 | 1.14 | 1.17 | 1.16 | 0.57 | 0.28 |
| **EPQ - Neuroticism**  | 91 | 44 | 0.52 | 0.21 | 0.17 | 0.19 | 0.1 | -0.05 |
| **EPQ - Extraversion**  | 91 | 44 | 0.52 | 0.18 | 0.18 | 0.18 | 0.09 | -0.03 |
| **EPQ - Psychoticism**  | 91 | 44 | 0.52 | 0.18 | 0.19 | 0.19 | 0.1 | -0.06 |
| **EPQ – Social desirability**  | 91 | 44 | 0.52 | 0.22 | 0.24 | 0.23 | 0.12 | -0.02 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dimension: Risk** |  |  |  |  |  |  |  |  |
|  | ***N1*** | ***N2*** | **Cohen’s *d*** | ***SD1*** | ***SD2*** | $$SD\_{pooled}$$ | **Lowest difference** | **Actual difference** |
| **Loss aversion (**$λ$**)** | 87 | 95 | 0.42 | 2.6 | 2.56 | 2.58 | 1.08 | 0.09 |
| **Utility curvature (**$α$**) - gains** | 87 | 95 | 0.42 | 2.67 | 2.68 | 2.68 | 1.12 | 0.03 |
| **Utility curvature (**$α$**) - losses** | 87 | 95 | 0.42 | 3.93 | 4.44 | 4.19 | 1.75 | 0.01 |
| **Probability weighting (**$γ$**) - gains** | 87 | 95 | 0.42 | 2.47 | 2.05 | 2.27 | 0.95 | 0.35 |
| **Probability weighting (**$γ$**) - losses** | 87 | 95 | 0.42 | 3.11 | 1.81 | 2.54 | 1.06 | 0.54 |
| **Present Bias (**$β$**) - gains** | 87 | 94 | 0.42 | 0.21 | 0.35 | 0.29 | 0.12 | -0.04 |
| **Present Bias (**$β$**) - losses** | 87 | 95 | 0.42 | 0.17 | 0.23 | 0.2 | 0.08 | 0.05 |
| **Discounting (**$δ$**) - gains** | 87 | 95 | 0.42 | 0.12 | 0.11 | 0.12 | 0.05 | 0 |
| **Discounting (**$δ$**) - losses** | 87 | 95 | 0.42 | 0.08 | 0.06 | 0.07 | 0.03 | -0.01 |
| **Age** | 87 | 95 | 0.42 | 1.57 | 1.38 | 1.48 | 0.62 | 0.25 |
| **Cigarettes (per week)** | 87 | 95 | 0.42 | 2.41 | 2.85 | 2.64 | 1.1 | -0.31 |
| **BMI** | 87 | 95 | 0.42 | 5.07 | 3.38 | 4.3 | 1.8 | 0.09 |
| **Alcohol (glasses/week)** | 87 | 95 | 0.42 | 9.58 | 9.33 | 9.46 | 3.95 | -0.35 |
| **Exercise (days/week)** | 87 | 95 | 0.42 | 1.59 | 1.74 | 1.67 | 0.7 | 0.08 |
| **Trait self-control**  | 79 | 83 | 0.44 | 0.57 | 0.59 | 0.58 | 0.26 | -0.07 |
| **Cognitive reflection**  | 72 | 74 | 0.47 | 1.15 | 1.17 | 1.16 | 0.54 | -0.05 |
| **EPQ - Neuroticism**  | 69 | 66 | 0.49 | 0.18 | 0.21 | 0.2 | 0.1 | 0.01 |
| **EPQ - Extraversion**  | 69 | 66 | 0.49 | 0.17 | 0.19 | 0.18 | 0.09 | 0 |
| **EPQ - Psychoticism**  | 69 | 66 | 0.49 | 0.2 | 0.18 | 0.19 | 0.09 | 0.02 |
| **EPQ – Social desirability**  | 69 | 66 | 0.49 | 0.21 | 0.24 | 0.23 | 0.11 | -0.03 |

**Table B4.** Post-hoc power analysis for Weekly dimension of tailored incentives

**Appendix C: Regression results including additional control variables and interactions**

A full overview of all models ran can be found below (Table C1, which confirm that selected incentives could not reliably be predicted from any of the measures collected (except BMI for the timing dimension). For the Timing dimension, after many exploratory regression analyses, a model with some significant predictors could be developed, which is reported in Table C2. Due to the exploratory process through which these results were obtained, no conclusions are based on it in the main text. Although more model specifications were possible, any correction for multiple hypothesis testing (which would be advised given the plethora of tests applied here) would quickly lead to null results.

**Table C1:** All models ran, including significant ($p<0.05)$ predictors (**boldfaced**), adjusted R-squared, Akaike’s Information Criterions (AIC) and Bayesian Information Criterion (BIC).

**Note:** All models are specified as R model formulas, where x~y indicates predicting x by y. \* signifies that these economic preferences were also not a significant predictor of incentive choice after controlling for all demographics and/or psychological measures,

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Outcome** | **Model ran**  | $$R^{2}$$ | **AIC** | **BIC** |
| Pre-commit (PRC) | Logistic regression |  |  |  |
|  | *Economic preferences* |  |  |  |
|  | PRC ~ Loss aversion\* | 0.012 | 231.15 | 233.13 |
|  | PRC ~ Utility curvature (losses)\* | <0.001 | 233.13 | 239.54 |
|  | PRC ~ Probability weighting (losses)\* | 0.009 | 231.75 | 238.2 |
|  | PRC ~ Present bias (losses)\* | 0.004 | 233.24 | 239.65 |
|  | PRC ~ Discounting (losses)\* | 0.009 | 231.37 | 237.78 |
|  | PRC ~ Loss aversion + Utility curvature (losses) + Probability weighting (losses) + Present bias (losses) + Discounting (losses) \* | 0.026 | 236.23 | 255.46 |
|  | *Demographics*  |  |  |  |
|  | PRC ~ BMI | 0.002 | 233.02 | 239.43 |
|  | PRC ~ Age | <0.001 | 233.31 | 239.72 |
|  | PRC ~ Gender | 0.010 | 231.02 | 237.43 |
|  | PRC ~ Exercise + Smoking + Alcohol | <0.001 | 236.69 | 249.51 |
|  | *Psychological measures* |  |  |  |
|  | PRC ~ Cognitive reflection task (CRT) | <0.001 | 192.18 | 198.16 |
|  | PRC ~ Trait self-control (TSC) | <0.001 | 212.33 | 218.52 |
|  | PRC ~ Eysenck Personality Questionnaire (EPQ) – Extraversion (E) + Neuroticism (N) + Psychoticism (P) + Social Desirability (SD) | 0.04 | 177.35 | 191.92 |
|  | PRC ~ BMI + Age + Gender + CRT + TSC + EPQ-E + EPQ-N + EPQ-P + EPQ-SD | 0.05 | 188.10 | 220.14 |
|  |  |  |  |  |
| Timing | Logistic regression |  |  |  |
|  | *Economic preferences* |  |  |  |
|  | TIMING ~ Loss aversion\*\* | <0.001 | 255.48 | 261.89 |
|  | TIMING ~ Utility curvature (gains)\* | 0.017 | 251.71 | 258.12 |
|  | TIMING ~ Utility curvature (losses)\* | 0.011 | 252.80 | 259.21 |
|  | TIMING ~ Present bias (gains)\* | 0.007 | 252.92 | 259.31 |
|  | TIMING ~ **Present bias (losses)***Note: Present Bias (losses) was only significant after controlling for demographics* | 0.018 | 252.17 | 285.58 |
|  | TIMING ~ Loss aversion + Utility curvature (gains) + Utility curvature (losses) + Present bias (gains) + **Present bias (losses)** *Note: Present Bias (losses) was significant after controlling for demographics* | 0.059 | 250.36 | 269.55 |
|  | *Demographics*  |  |  |  |
|  | TIMING ~ BMI | 0.017 | 252.23 | 258.64 |
|  | TIMING ~ Age | <0.001 | 255.49 | 261.90 |
|  | TIMING ~ Gender | <0.001 | 255.48 | 261.89 |
|  | TIMING ~ Exercise + Smoking + Alcohol | 0.010 | 257.80 | 270.61 |
|  | *Psychological measures* |  |  |  |
|  | TIMING ~ Cognitive reflection task (CRT) | 0.020 | 205.21 | 211.19 |
|  | TIMING ~ Trait self-control (TSC) | 0.010 | 227.67 | 233.85 |
|  | TIMING ~ Eysenck Personality Questionnaire (EPQ) – Extraversion (E) + Neuroticism (N) + Psychoticism (P) + Social Desirability (SD) | 0.020 | 195.48 | 210.04 |
|  | TIMING ~ BMI + Age + Gender + CRT +TSC+EPQ-E + EPQ-N + EPQ-P + EPQ-SD | 0.070 | 199.64 | 231.68 |
|  |  |  |  |  |
| Sequence | Linear regression |  |  |  |
|  | *Economic preferences* |  |  |  |
|  | SEQUENCE ~ Loss aversion\* | 0.001 | 427.88 | 437.49 |
|  | SEQUENCE ~ Utility curvature (gains)\* | 0.007 | 426.87 | 436.48 |
|  | SEQUENCE ~ Utility curvature (losses)\* | 0.001 | 427.98 | 437.59 |
|  | SEQUENCE ~ Discounting (gains)\* | 0.011 | 426.14 | 435.75 |
|  | SEQUENCE ~ Discounting (losses)\* | 0.001 | 427.93 | 437.54 |
|  | SEQUENCE ~ Loss aversion + Utility curvature (gains) + Utility curvature (losses) + Discounting (gains)+ Discounting (losses)\* | 0.022 | 432.09 | 454.52 |
|  | *Demographics*  |  |  |  |
|  | SEQUENCE ~ BMI | 0.009 | 426.51 | 436.12 |
|  | SEQUENCE ~ Age | 0.010 | 426.22 | 435.83 |
|  | SEQUENCE ~ Gender | <0.001 | 428.09 | 437.71 |
|  | SEQUENCE ~ Exercise + Smoking + Alcohol | 0.009 | 430.48 | 446.50 |
|  | *Psychological measures* |  |  |  |
|  | SEQUENCE ~ Cognitive reflection task (CRT) | <0.001 | 340.59 | 349.56 |
|  | SEQUENCE ~ Cognitive reflection task (TSC) | <0.001 | 371.62 | 380.91 |
|  | SEQUENCE ~ Eysenck Personality Questionnaire (EPQ) – Extraversion (E) + Neuroticism (N) + Psychoticism (P) + Social Desirability (SD) | 0.030 | 315.68 | 333.16 |
|  | SEQUENCE ~ BMI + **Age +** Gender + CRT +TSC+ EPQ-E + EPQ-N + EPQ-P + EPQ-SD | 0.110 | 315.78 | 350.73 |
|  |  |  |  |  |
| Risk | Linear regression |  |  |  |
|  | *Economic preferences* |  |  |  |
|  | RISK ~ Loss aversion\* | 0.001 | 1685.40 | 1695.01 |
|  | RISK ~ Probability weighting (gains)\* | 0.008 | 1684.09 | 1693.70 |
|  | RISK ~ Probability weighting (losses)\* | 0.011 | 1683.45 | 1693.06 |
|  | RISK ~ Loss aversion + Utility curvature (losses) + Probability weighting (losses) + Present bias (losses) + Discounting (losses)\* | 0.020 | 1685.78 | 1701.80 |
|  | *Demographics*  |  |  |  |
|  | RISK ~ BMI | 0.003 | 1684.95 | 1694.56 |
|  | RISK ~ Age | 0.001 | 1685.39 | 1695.00 |
|  | RISK ~ Gender | <0.001 | 1685.44 | 1695.05 |
|  | RISK ~ Exercise + Smoking + Alcohol | 0.010 | 1687.60 | 1703.62 |
|  | *Psychological measures* |  |  |  |
|  | RISK ~ Cognitive reflection task (CRT) | 0.010 | 1372.68 | 1381.65 |
|  | RISK ~ Trait self-control (TSC) | <0.001 | 1517.89 | 1287.15 |
|  | RISK ~ Eysenck Personality Questionnaire (EPQ) – Extraversion (E) + Neuroticism (N) + Psychoticism (P) + Social Desirability (SD) | 0.010 | 1269.67 | 1287.15 |
|  | RISK ~ BMI + Age + Gender + CRT + TSC + EPQ-E + EPQ-N + EPQ-P + EPQ-SD | 0.030 | 1277.66 | 1312.61 |

For the Timing dimension, after exploring many different model specifications, one of the better fitting models included: present bias for losses, Eysenck Personality Questionnaire dimensions: Neuroticism and Psychoticism, Cognitive Reflection and BMI. Logistic regression results are reported in Table C2.

**Table C2.** Results for exploratory logistic regression for timing dimension.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Predictor** | **Estimate** | **SE** | **Z-value** | $p$ **value** |
| (Intercept) | -0.46 | 1.42 | -0.33 | 0.74 |
| Present bias (losses) | 1.65 | 0.94 | 1.76 | 0.08 |
| Eysenck Personality Questionnaire (Neuroticism) | 1.71 | 0.97 | 1.76 | 0.08 |
| Eysenck Personality Questionnaire (Psychoticism) | 0.98 | 1.02 | 0.97 | 0.34 |
| Cognitive reflection | -0.38 | 0.16 | -2.33 | 0.02 |
| BMI | -0.09 | 0.04 | -1.80 | 0.07 |

These results indicate that: those with weaker present bias for losses (marginally significant), those are more prone to neuroticism (marginally significant), those who more on their automatic system, and those with a lower BMI (marginally significant) are more likely to choose a weekly pay-out structure.

**References**

ABDELLAOUI, M. 2000. Parameter-free elicitation of utility and probability weighting functions. *Management science,* 46**,** 1497-1512.

ABDELLAOUI, M., BLEICHRODT, H., L’HARIDON, O. & VAN DOLDER, D. 2016. Measuring Loss Aversion under Ambiguity: A Method to Make Prospect Theory Completely Observable. *Journal of Risk and Uncertainty,* 52**,** 1-20.

FRANCIS, L. J., BROWN, L. B. & PHILIPCHALK, R. 1992. The development of an abbreviated form of the Revised Eysenck Personality Questionnaire (EPQR-A): Its use among students in England, Canada, the USA and Australia. *Personality and individual differences,* 13**,** 443-449.

KÖBBERLING, V. & WAKKER, P. P. 2005. An index of loss aversion. *Journal of Economic Theory,* 122**,** 119-131.

LAIBSON, D. 1997. Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics,* 112**,** 443-478.

TANGNEY, J. P., BOONE, A. L. & BAUMEISTER, R. F. 2018. High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Self-Regulation and Self-Control.* Routledge.

TOPLAK, M. E., WEST, R. F. & STANOVICH, K. E. 2011. The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition,* 39**,** 1275.

WAKKER, P. & DENEFFE, D. 1996. Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management science,* 42**,** 1131-1150.