Appendix

A Experimental Task

A.1 Text messages in the Effort Task

A.1.1 Day 0

MCN: Up for a word task?
Earn airtime by writing words backward like this:
baby, parent = ybab, tnerap
Complete the task and get R20 in 7 days time.
The sooner you do it, the easier it is.
You choose when to do it:
Do it tomorrow (1 day) - write 4 words
Do it in 3 days - 6 words
Do it in 5 days - 10 words.
We will send you words tomorrow.
In how many days do you think you will answer?
Reply "1" "3" "5" or "No days".
Receive R5 now for this answer.
For more info, text MORE.

MCN is an acronym which signals to the mothers that the text comes from MomConnect. If the mother Reply "1" "3" "5" or "No days" she gets the following text:

You will now receive R5 for answering.
Tomorrow you will receive words to write backwards

If the mother replies with the word "MORE" she also receives the following information:

In this question, we ask when you think you will write words backwards.
We do this to investigate the importance of timing when doing tedious tasks.
Reply "1" "3" "5" or "No days" and receive R5 today for this answer.
In the following days, we'll send you the words.
If you write them backwards, you will get R20 in a week's time.

A.1.2 Day 1

MCN: Please reply by writing these 4 words backwards:
baby, parent, bottle, mom
Do it today and get R20 airtime in 6 days' time.
If you don´t do it today, you get another chance if
you reply to a text in 2 days with 7 words to write backwards.

If the mother replies with the words written backwards she receives the following information:
Thank you, you will receive R20 in airtime in 6 days.

A.1.3 Day 3

MCN: Please reply by writing these 6 words backwards:
baby, parent, bottle, mom, sleep, child.
Do it today and get R20 airtime in 4 days' time.
If you don't do it today, you get another chance if
you reply to a text in 2 days with 10 words to write backwards

If the mother replies with the words written backwards she receives the following information:
Thank you, you will receive R20 in airtime in 4 days.

A.1.4 Day 5

MCN: Please reply by writing these 10 words backwards:
baby, parent, bottle, mom, sleep, child, play, love, dad, happy.
Last chance to do the task for your R20 airtime in 2 days' time.

If the mother replies with the words written backwards she receives the following information:
Thank you, you will receive R20 in airtime in 2 days.

A.1.5 Day 7

If the mother replies with the words written backwards at any of the days of the experiment, she receives the following information:

MCN: You will now receive your R20 for the word task you did. Have a great day!

B Technical Appendix for the Interpretation of the Time-Preference Task

This appendix provides the technical details related to the interpretation of future and immediate effort choices in the time-preference task. To ease the exposition we assume through-out this appendix that $c_t = E(c_t) = \frac{1}{\phi}$ i.e. there is no uncertainty about effort costs.

B.1 Preferred Behavior at Time $k = 0$ for naives

Assuming that the individual is naive about possible present bias at time $k = 0$ when the plan is made, the table below provides conditions for when the different possible timings are preferred from the point of view of time $k = 0$. That is given the discount factor δ these are the plans the individual would like to implement ex ante if possible.

Table B.1: Pairwise Comparisons and Conditions for Optimality at $k = 0$

	Preferred to $t = 1$	Preferred to $t = 3$	Preferred to $t = 5$	Preferred to never
$t = 1$		$\delta^2 \geq \frac{2}{3}$	$(\delta^4 \geq \frac{2}{5})$	$\delta^6 \geq \frac{1}{5\phi}$
$t = 3$	$\delta^2 \leq \frac{2}{3}$		$\delta^2 \geq \frac{3}{5}$	$\delta^4 \geq \frac{3}{10\phi}$
$t = 5$	$(\delta^4 \leq \frac{2}{5})$	$\delta^2 \leq \frac{3}{5}$		$\delta^2 \geq \frac{1}{2\phi}$
Never	$\delta^6 < \frac{1}{5\phi}$	$\delta^4 < \frac{3}{10\phi}$	$\delta^2 < \frac{1}{2\phi}$	

Notes: conditions in brackets are implied by other conditions in the row and are thus never binding.

B.2 Deriving Actual Behavior at time $k > 0$

This section derives optimal behavior in periods $k = 1$, $k = 3$ and $k = 5$. Consider the behavior of an individual with present-biased preferences. We solve the problem backwards as the sophisticated agent would do. Let $\hat{\beta}$ denote the belief about future present bias. Naivete is then the special case where $\hat{\beta} = 1$ and full sophistication is the special case where $\hat{\beta} = \beta$.

Period $t = 5$

When period 5 arrives $k = 5$ and the agent will do the task if the participation constraint is satisfied i.e. if

$$\beta\delta^2 \geq \frac{1}{2\phi} \tag{3}$$

Consider an individual who planned to to the task in period $t = 5$. Recall that for these individuals $\frac{3}{5} > \delta^2 \geq \frac{1}{2\phi}$ if they are time consistent or naive. For individuals with time-consistent preferences ($\beta = 1$), the participation constraint in Equation (3) is identical to the original participation constraint and it is satisfied for anyone who planned $t = 5$. For people who are present-biased ($\beta < 1$) and naive ($\hat{\beta} = 1$), the participation constraint may not be satisfied if β sufficiently low even if the individual originally planned to do the task at time $t = 5$. Note that when δ is close to the lower bound in the interval $I_5 = [\frac{1}{2\phi}; \frac{3}{5}]$ and the individual thus planned to do the task at time $t = 5$ then a β just below unity is sufficient to violate the participation constraint in Equation (3). If δ is close to the upper bound in I_5 then a smaller value of β is required for violations of Equation (3). For an individual with future bias ($\beta > 1$), the participation constraint may be satisfied now even if it was not satisfied at the planning stage ($k = 0$) and the individual may therefore do the task even if she did not plan to do it (i.e. women with $\delta < \sqrt{\frac{1}{2\phi}}$ may now do the task).

Period $t = 3$

When period 3 arrives, i.e. $k = 3$, the individual does the task in period $t = 3$ if

$$\beta\delta^4 \geq \frac{3}{10\phi} \text{ and } \beta\delta^2 \geq \frac{3}{5} \tag{4}$$

or if

$$\beta\delta^4 \geq \frac{3}{10\phi} \text{ and } \beta\delta^2 < \frac{3}{5} \text{ and } \hat{\beta}\delta^2 < \frac{1}{2\phi}. \tag{5}$$

In Equation (4), the first inequality ensures that the individual prefers that the task is done now to not at all and the second inequality ensures that the individual prefers doing the task now to delaying to time $t = 5$. These are the conditions that are required to ensure that the individual prefers doing the task at time $t = 3$. We can compare this to the conditions for planning at time $k = 0$ to the the task at $t = 3$. For women who are time-consistent or naive, such a plan implies that $\delta^2 \in \left[\frac{3}{5}; \frac{2}{3}\right]$ and $\delta^4 \geq \frac{3}{10\phi}$ (see in Table 1). A time-consistent individual thus also carries out the task as planned if she planned to do it at time $t = 3$. However, the conditions in Equation (4) may not be satisfied for an individual with present-biased preferences ($\beta < 1$) even if she originally (naively) planned $t = 3$ and as a result she may delay the task. This may either happen because the individual no longer prefers that the task is done (the first inequality is not satisfied) or because the individual prefers that the task is done in the next period. Realizing this the woman may also do the task in period $t = 3$ if she knows that by postponing she risks that she will never actually do the task even if she prefers doing the task now at time $t = k = 1$ to never doing it. This is captured by the conditions in Equation (5). The first inequality in Equation (5) ensures

that the individual prefers that the task is done now to not at all and the second inequality implies that she would prefer to delay to the next period while the third inequality states that the individual predicts that the task will not be carried out if it is delayed. Clearly a smaller $\hat{\beta}$ makes is more likely that the second inequality is satisfied. Intuitively the stronger the predicted future present bias the less likely it is that the women will predict that the task will be done in period $t = 5$ if it is delayed. For an individual who is naive about her future present bias the second inequality reduces to the $k = 0$ non-participation constraint for doing the task in period $t = 5$. For an individual who is fully sophisticated about her future present bias at time $k = 3$ the condition is equivalent to the actual non-participation constraint when $k = 5$ (see above).

Period $t = 1$

When period 1 arrives, i.e. $k = 1$, the individual does the task in period $t = 1$ if

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 \geq \frac{2}{3} \text{ and } \beta\delta^4 \geq \frac{2}{5} \tag{6}$$

or if

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 < \frac{2}{3} \text{ and } \beta\delta^4 \geq \frac{2}{5} \text{ and } \left(\hat{\beta}\delta^2 < \frac{3}{5} \text{ or } \hat{\beta}\delta^4 < \frac{3}{10\phi}\right) \tag{7}$$

or if

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 < \frac{2}{3} \text{ and } \beta\delta^4 < \frac{2}{5} \text{ and } \left(\hat{\beta}\delta^2 < \frac{3}{5} \text{ or } \hat{\beta}\delta^4 < \frac{3}{10\phi}\right) \text{ and } \hat{\beta}\delta^2 < \frac{1}{2\phi} \tag{8}$$

or if

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 \geq \frac{2}{3} \text{ and } \beta\delta^4 < \frac{2}{5} \text{ and } \left(\hat{\beta}\delta^2 \geq \frac{3}{5} \text{ and } \hat{\beta}\delta^4 \geq \frac{3}{10\phi}\right) \tag{9}$$

The inequalities in Equation (6) are the conditions that ensure that the individual at time $k = 1$ prefers to do the task now (i.e. at time $t = 1$) rather than to delay to a future period. If this is the case, she does the task immediately.

The inequalities in Equation (7) are the conditions that describe a situation where the individual at time $k = 1$ prefers that the task is done in period $t = 3$ rather than at time $t = 1$ ($\beta\delta^2 < \frac{2}{3}$) but also prefers doing it in period $t = 1$ to period $t = 5$ ($\beta\delta^4 \geq \frac{2}{5}$). Note that these two conditions also imply that $\delta^2 > \frac{3}{5}$, i.e. that the individual prefers at time $k = 1$ that the task is done at time $t = 3$ to at time $t = 5$. However, when one of the inequalities in the bracket holds she predicts that if she delays to period $k = 3$, she will delay again to time $k = 5$ (and possibly never do the task). Therefore it is better to do the task immediately to avoid unnecessary delay.

The conditions in Equation (8) ensure that the individual would actually like the task to be done but would prefer that it is done in a later time period ($\beta\delta^2 < \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$) but she predicts that the task will not be done if it is delayed. Therefore it is again optimal to do the task now to avoid delaying and not getting the task done.

Finally, Equation (9) describes a situation where the individual prefers that the task is done to not at all but prefers at time $k = 1$ that it is done in period $t = 5$ to that it is done in period $t = 1$ ($\beta\delta^4 < \frac{2}{5}$) which again is better than if it is done in period $t = 3$ ($\beta\delta^2 \geq \frac{2}{3}$). Note that these two conditions together also imply $\delta^2 \leq \frac{3}{5}$, i.e. that the individual at time $k = 1$ prefers if the task is delayed from period $t = 3$ to $t = 5$. However, by delaying from period $t = 1$ she predicts that the task will be done in period $t = 3$ rather than delayed to period $t = 5$. Then it is better to do the task now.

For a time-consistent individual with $\beta = \hat{\beta} = 1$ this reduces to

$$\delta^6 \geq \frac{1}{5\phi} \text{ and } \delta^2 \geq \frac{2}{3}$$

We note that the situations captured by Equation (7)[23], (8)[24] and (9)[25], cannot arise (i.e. all inequalities cannot be satisfied simultaneously), as the time-consistent individual will have the same preferences when making decisions at time $k = 3$ and $k = 5$ as she does at time $k = 1$. Therefore she will continue to rank the utility associated with doing the task in the different periods in the same way. As a result, the individual does the task at time $t = 1$ only if the conditions in the equation above satisfied.

For a naive $\beta \neq \hat{\beta} = 1$ the conditions reduce to

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 \geq \frac{2}{3} \text{ and } \beta\delta^4 \geq \frac{2}{5}$$

Again the situations captured by Equation (7)[26], (8)[27] and (9)[28] cannot arise (i.e. all in-

[23] In particular $\delta^2 < \frac{2}{3}$ and $\delta^4 \geq \frac{2}{5}$ imply $\delta^2 \geq \frac{3}{5}$ which contradicts $\delta^2 < \frac{3}{5}$. In addition, $\delta^4 < \frac{3}{10\phi}$ and $\delta^6 \geq \frac{1}{5\phi}$ imply $\delta^2 \geq \frac{2}{3}$ which contradicts $\delta^2 < \frac{2}{3}$. Hence, the situation cannot arise.

[24] One can show that $\delta^4 < \frac{3}{10\phi}$ and $\delta^6 \geq \frac{1}{5\phi}$ imply $\delta^2 \geq \frac{2}{3}$ which contradicts $\delta^2 < \frac{2}{3}$. At the same time, $\delta^2 < \frac{1}{2\phi}$ and $\delta^6 \geq \frac{1}{5\phi}$ imply $\delta^4 \geq \frac{2}{5}$ which contradicts $\delta^4 < \frac{2}{5}$. Hence, the situation described by Equation (8) cannot arise.

[25] In particular $\beta\delta^2 \geq \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$ imply $\delta^2 < \frac{3}{5}$ which contradicts $\delta^2 \geq \frac{3}{5}$.

[26] In particular $\beta\delta^2 < \frac{2}{3}$ and $\beta\delta^4 \geq \frac{2}{5}$ imply $\delta^2 \geq \frac{3}{5}$ which contradicts $\delta^2 < \frac{3}{5}$. In addition, $\delta^4 < \frac{3}{10\phi}$ and $\beta\delta^6 \geq \frac{1}{5\phi}$ imply $\beta\delta^2 \geq \frac{2}{3}$ which contradicts $\beta\delta^2 < \frac{2}{3}$. Hence, the situation cannot arise.

[27] One can show that $\delta^4 < \frac{3}{10\phi}$ and $\beta\delta^6 \geq \frac{1}{5\phi}$ imply $\beta\delta^2 \geq \frac{2}{3}$ which contradicts $\beta\delta^2 < \frac{2}{3}$. At the same time, $\delta^2 < \frac{1}{2\phi}$ and $\beta\delta^6 \geq \frac{1}{5\phi}$ imply $\beta\delta^4 \geq \frac{2}{5}$ which contradicts $\beta\delta^4 < \frac{2}{5}$. Hence, the situation described by Equation (8) cannot arise.

[28] In particular $\beta\delta^2 \geq \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$ imply $\delta^2 < \frac{3}{5}$ which contradicts $\delta^2 \geq \frac{3}{5}$.

equalities cannot be satisfied simultaneously), as the naive individual (wrongly) *predicts* that he will have the same preferences when making decisions at time $k = 3$ and at time $k = 5$ as a time-consistent individual. Therefore, the individual does the task at time $t = 1$ only if the conditions in the first equation are satisfied because she thinks that she will choose to act as a time-consistent individual in future periods. Note that the conditions in the first equation differ from the comparable conditions for the time-consistent individual because of the present-bias parameter. When $\beta < 1$ the conditions are less likely to be satisfied i.e. it is less likely that a naive individual will do the task in period $t = 1$ compared to a time-consistent individual. In contrast, it is more likely that the conditions will be satisfied if $\beta > 1$ i.e. in the case of future bias.

For a sophisticated individual with $\beta = \hat{\beta} \neq 1$ the conditions reduce to

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 \geq \frac{2}{3} \text{ and } \beta\delta^4 \geq \frac{2}{5}$$

or if

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 < \frac{2}{3} \text{ and } \beta\delta^4 \geq \frac{2}{5} \text{ and } \left(\beta\delta^2 < \frac{3}{5} \text{ or } \beta\delta^4 < \frac{3}{10\phi}\right)$$

or if

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 < \frac{2}{3} \text{ and } \beta\delta^4 < \frac{2}{5} \text{ and } \left(\beta\delta^2 < \frac{3}{5} \text{ or } \beta\delta^4 < \frac{3}{10\phi}\right) \text{ and } \beta\delta^2 < \frac{1}{2\phi}$$

or if

$$\beta\delta^6 \geq \frac{1}{5\phi} \text{ and } \beta\delta^2 \geq \frac{2}{3} \text{ and } \beta\delta^4 < \frac{2}{5} \text{ and } \beta\delta^4 \geq \frac{3}{10\phi}$$

The first condition here is identical to that for naively present-biased individuals. But the remaining conditions do not become irrelevant because the sophisticated individual in contrast to the naive individual realizes that preferences are different in future periods meaning that the ranking of the utility associated with the different timings as well as the participation constraints may not be satisfied in future periods.

Optimal Behavior under Different Assumptions

Now let us characterize optimal behavior under different assumptions regarding the individual's time preferences.

Time-Consistent Preferences

First consider the time-consistent individual ($\beta = \hat{\beta} = 1$). She will plan and do the task in accordance with the plan as detailed in the table below.

Table B.2: Planned and Actual Behavior Predicted for Individuals with Time-Consistent Preferences

	Timing optimal		Participation constraint
Plan and do $t=1$ if	$\delta^2 > \frac{2}{3}$	and	$\delta^6 \geq \frac{1}{5\phi}$
Plan and do $t=3$ if	$\delta^2 \in \left[\frac{3}{5}; \frac{2}{3}\right]$	and	$\delta^4 \geq \frac{3}{10\phi}$
Plan and do $t=5$ if	$\delta^2 < \frac{3}{5}$	and	$\delta^2 \geq \frac{1}{2\phi}$
Plan not to do it			if none of the above hold

We note that the participation constraint is what ensure that the utility from doing it in the stated period is positive and the conditions on δ stated under "Timing optimal" are the conditions that ensure that the utility is maximized when the task is done at the specified point in time. Note that it is therefore not possible that for given timing t the conditions for optimal timing are met but the the participation constraint is not met in period t but it is met in a future period. Therefore the conditions above fully describe the optimal choice and behavior of the time-consistent individual.

Present-Bias and Naivete

Now, let's consider the behavior of a naive individual with $\beta \neq \hat{\beta} = 1$. This individual will ex-ante plan to behave exactly as the time-consistent individual but will behave as described by the table below.

Table B.3: Actual Behavior Predicted for Individuals with Naive Present-Biased Preferences

	Timing optimal		Participation constraint
Do $t=1$ if	$\beta\delta^2 > \frac{2}{3}$ and $\beta\delta^4 > \frac{2}{5}$	and	$\beta\delta^6 \geq \frac{1}{5\phi}$
Do $t=3$ if	$\beta\delta^2 \in \left[\frac{3}{5}; \frac{2}{3}\right]$	and	$\beta\delta^4 \geq \frac{3}{10\phi}$
Do $t=5$ if	$\beta\delta^2 < \frac{3}{5}$	and	$\beta\delta^2 \geq \frac{1}{2\phi}$
Don't do it			if none of the above hold

That is the individual may delay doing the task or fail to do the task if $\beta < 1$ in which case she is naively present biased. The individual may do the task early if $\beta > 1$ in which case she is naively future biased. Note, however, that doing the task early is not possible if the individual planned to do the task at time $t = 1$.

Time-Inconsistent Preferences and Sophistication

Now, let's consider the behavior of an individual who is fully sophisticated with $\beta = \hat{\beta} \neq 1$. When asked when she expects to do the task, she realizes that she may procrastinate and she

will predict her own behavior. Her plan will therefore coincide perfectly with the behavior detailed in the table below. That is in this case planned behavior depends not only on her realization of δ but also on her realization of β.

Table B.4: Planned and Actual Behavior Predicted for Individuals who have Present-Biased Preferences and are Fully Sophisticated

	Timing optimal
Do $t=1$ if	$\beta\delta^2 > \frac{2}{3}$ and $\beta\delta^4 > \frac{2}{5}$ and $\beta\delta^6 \geq \frac{1}{5\phi}$
or if	$\beta\delta^6 \geq \frac{1}{5\phi}$ and $\beta\delta^2 < \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$ and $\left(\beta\delta^2 < \frac{3}{5} \text{ or } \beta\delta^4 < \frac{3}{10\phi}\right)$
or if	$\beta\delta^6 \geq \frac{1}{5\phi}$ and $\beta\delta^2 < \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$ and $\left(\beta\delta^2 < \frac{3}{5} \text{ or } \beta\delta^4 < \frac{3}{10\phi}\right)$ and $\beta\delta^2 < \frac{1}{2\phi}$
or if	$\beta\delta^6 \geq \frac{1}{5\phi}$ and $\beta\delta^2 \geq \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$ and $\beta\delta^4 \geq \frac{3}{10\phi}$
Do $t=3$ if	$\beta\delta^2 \geq \frac{3}{5}$ and $\beta\delta^4 \geq \frac{3}{10\phi}$
or if	$\beta\delta^4 \geq \frac{3}{10\phi}$ and $\beta\delta^2 < \frac{3}{5}$ and $\beta\delta^2 < \frac{1}{2\phi}$
Do $t=5$ if	$\beta\delta^2 < \frac{3}{5}$ and $\beta\delta^2 \geq \frac{1}{2\phi}$
Don't do it	if none of the above hold

Finally, consider an individual who is partially sophisticated $\beta < \hat{\beta} \neq 1$. This individual underestimates her present bias and plans as a fully sophisticated individual where β is replaced by $\hat{\beta}$ throughout as in the table below.

Table B.5: Planned Behavior Predicted for Individuals with Present-Biased Preferences who are Partially Sophisticated

	Timing optimal
Do $t=1$ if	$\hat{\beta}\delta^2 > \frac{2}{3}$ and $\hat{\beta}\delta^4 > \frac{2}{5}$ and $\hat{\beta}\delta^6 \geq \frac{1}{5\phi}$
or if	$\hat{\beta}\delta^6 \geq \frac{1}{5\phi}$ and $\hat{\beta}\delta^2 < \frac{2}{3}$ and $\hat{\beta}\delta^4 < \frac{2}{5}$ and $\left(\hat{\beta}\delta^2 < \frac{3}{5} \text{ or } \hat{\beta}\delta^4 < \frac{3}{10\phi}\right)$
or if	$\hat{\beta}\delta^6 \geq \frac{1}{5\phi}$ and $\hat{\beta}\delta^2 < \frac{2}{3}$ and $\hat{\beta}\delta^4 < \frac{2}{5}$ and $\left(\hat{\beta}\delta^2 < \frac{3}{5} \text{ or } \hat{\beta}\delta^4 < \frac{3}{10\phi}\right)$ and $\hat{\beta}\delta^2 < \frac{1}{2\phi}$
or if	$\hat{\beta}\delta^6 \geq \frac{1}{5\phi}$ and $\hat{\beta}\delta^2 \geq \frac{2}{3}$ and $\hat{\beta}\delta^4 < \frac{2}{5}$ and $\hat{\beta}\delta^4 \geq \frac{3}{10\phi}$
Do $t=3$ if	$\hat{\beta}\delta^2 \geq \frac{3}{5}$ and $\hat{\beta}\delta^4 \geq \frac{3}{10\phi}$
or if	$\hat{\beta}\delta^4 \geq \frac{3}{10\phi}$ and $\hat{\beta}\delta^2 < \frac{3}{5}$ and $\hat{\beta}\delta^2 < \frac{1}{2\phi}$
Do $t=5$ if	$\hat{\beta}\delta^2 < \frac{3}{5}$ and $\hat{\beta}\delta^2 \geq \frac{1}{2\phi}$
Don't do it	if none of the above hold

That is the partially sophisticated individual realizes that she is present biased but she discounts future periods at a lower rate than she should and therefore may mispredict behavior. Her actual behavior is determined by the conditions in the table below. Compared

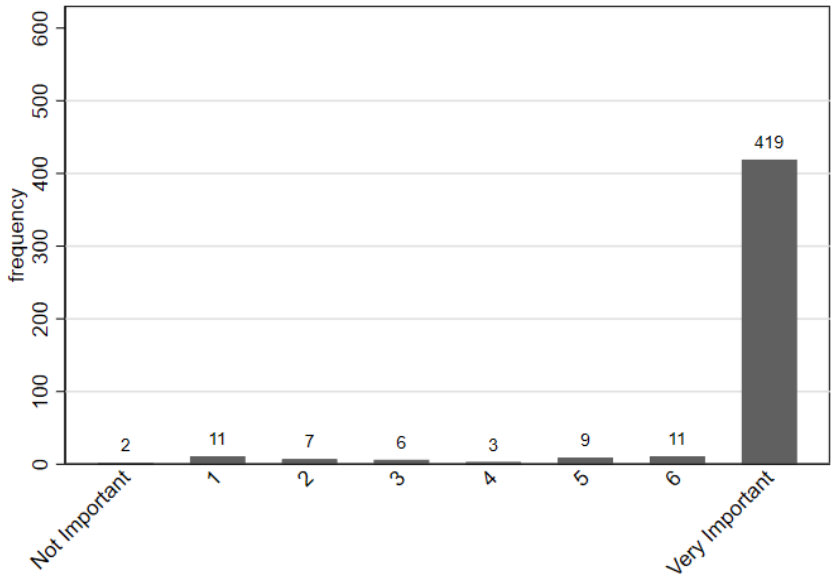
to the table detailing planned behavior, $\hat{\beta}$ is replaced by $\beta < \hat{\beta}$ in some conditions because thr present bias experience between the now and the future is β whereas the present bias associated with plans for future periods is still $\hat{\beta}$. The fact that β replaces $\hat{\beta}$ means that some conditions are no longer satisfied, when the time arrives and as a result the individual may chose to delay the task more than planned or choose not to do a task that she had planned to do.

Table B.6: Actual Behavior for Individuals with Present-Biased Preferences who are Partially Sophisticated

	Timing optimal
Do $t=1$ if	$\beta\delta^2 > \frac{2}{3}$ and $\beta\delta^4 > \frac{2}{5}$ and $\beta\delta^6 \geq \frac{1}{5\phi}$
or if	$\beta\delta^6 \geq \frac{1}{5\phi}$ and $\beta\delta^2 < \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$ and $\left(\hat{\beta}\delta^2 < \frac{3}{5} \text{ or } \hat{\beta}\delta^4 < \frac{3}{10\phi}\right)$
or if	$\beta\delta^6 \geq \frac{1}{5\phi}$ and $\beta\delta^2 < \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$ and $\left(\hat{\beta}\delta^2 < \frac{3}{5} \text{ or } \hat{\beta}\delta^4 < \frac{3}{10\phi}\right)$ and $\hat{\beta}\delta^2 < \frac{1}{2\phi}$
or if	$\beta\delta^6 \geq \frac{1}{5\phi}$ and $\beta\delta^2 \geq \frac{2}{3}$ and $\beta\delta^4 < \frac{2}{5}$ and $\hat{\beta}\delta^4 \geq \frac{3}{10\phi}$
Do $t=3$ if	$\beta\delta^2 \geq \frac{3}{5}$ and $\beta\delta^4 \geq \frac{3}{10\phi}$
or if	$\beta\delta^4 \geq \frac{3}{10\phi}$ and $\beta\delta^2 < \frac{3}{5}$ and $\hat{\beta}\delta^2 < \frac{1}{2\phi}$
Do $t=5$ if	$\beta\delta^2 < \frac{3}{5}$ and $\beta\delta^2 \geq \frac{1}{2\phi}$
Don't do it	if none of the above hold

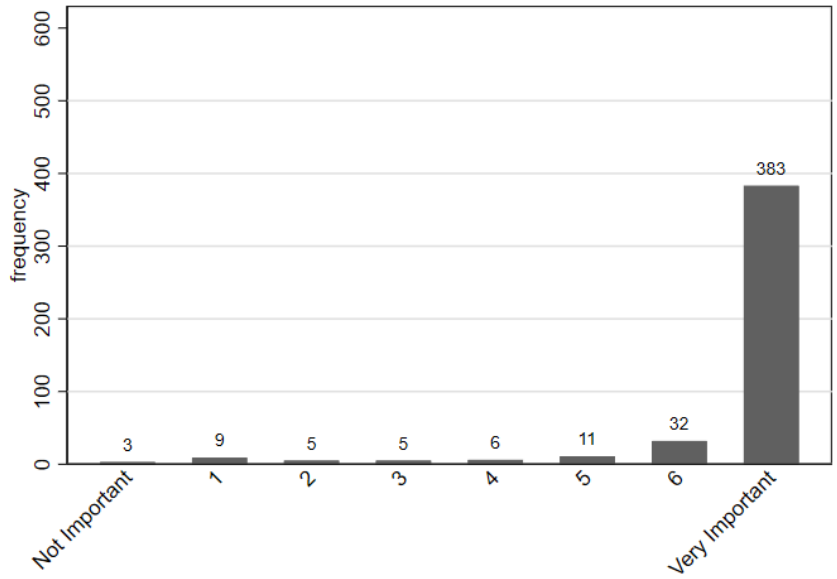
C Additional Results

Figure C.1: Iron Intake Important for Healthy Mothers



Notes: To be a healthy mom, is it important to take iron pills? How important do you think it is? Reply with a number from 0 (not important) to 7 (very important). In total 468 women in our sample answered this question.

Figure C.2: Iron Intake Important for Healthy Babies



Notes: To have a health baby, are iron pills important? How important do you think it is? Reply with a number from 0 (not important) to 7 (very important). In total, 454 women in our sample answered this question.

Table C.7: Testing for Selection into the Sample

	Out-of-sample	In-sample	Treatment Comparison
Birthyear	1991	1992	0.63
	(5.84)	(5.34)	
English	0.67	0.65	-0.01
	(0.47)	(0.48)	
Urban	0.39	0.29	-0.10**
	(0.49)	(0.46)	
Number of Previous Children	1.05	0.98	-0.07
	(1.01)	(1.13)	
N	214	480	

Notes (i) Standard deviations are reported in parentheses, (ii) The Treatment Comparison column reports the difference between the means of the In-sample and Out-of-sample groups, with a Wald test used to test for a statistically significant difference, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.8: Testing for Selection on Time Preferences into the Sample

	Out-of-sample	In-sample	TreatComparison
Late Planners	0.36	0.31	-0.06
	(0.48)	(0.46)	
Late Doers	0.15	0.09	-0.06
	(0.36)	(0.29)	
Early Doers	0.26	0.28	0.03
	(0.44)	(0.45)	
On Timers	0.59	0.63	0.04
	(0.50)	(0.48)	
N	66	480	

Notes Late Planners is a dummy variable where planning to do the task on day 1 is denoted 0 and planning to do the task on day 3 or 5, or not planning to do the task is denoted with a 1. Late doers are people who planned to do the word task sooner, did it later or did not do it. Early doers planned to do the task late but did it sooner OR did not plan to do it but did it in the end. On timers are people who do the task when they planned.
Standard deviations are reported in parentheses. The Treatment Comparison column reports the difference between the means of the In-sample and Out-of-sample groups, with a Wald test used to test for a statistically significant difference, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.9: Descriptive Statistics

	Means (std)
Time Preference Variables	
Time Preference Plan	1.78
	(1.45)
Attempt on Day 1	0.90
	(0.30)
Success fraction day 1 (conditional)	0.81
	(0.40)
Attempt on Day 3	0.04
	(0.19)
Success fraction day 3 (conditional)	0.67
	(0.49)
Attempt on Day 5	0.02
	(0.14)
Success fraction day 5 (conditional)	0.50
	(0.53)
Future Bias	0.28
	(0.45)
Present Bias	0.09
	(0.29)
Consistent	0.63
	(0.48)
Questionnaire	
Q1.Weekly Iron Intake	6.59
	(1.17)
Q2.Difficulty Remembering [=1]	0.10
	(0.30)
Q3.Failure to Carry out plans	4.64
	(3.74)
Q4.Willingness to take risks	4.26
	(3.84)
Q5.Iron pills – Stomach feels bad	0.12
	(0.33)
Q6.Weekly days tired / dizzy	2.35
	(1.81)
Q7.Low iron levels [=1]	0.10
	(0.31)
Q8.Years of Schooling	12.04
	(2.54)
N	480

Notes: (i) The table reports means and standard deviations for variables that we use in our main analysis, e.g. *Time Preference Plan* is the average number of days respondents plan to take to complete the task, *Attempt on day 1* is a binary indicator variable for an attempt on the first day, whereas *Success fraction on day x* reports the rate of success, conditional making an attempt on day x, *Future Bias, Present Bias* and *Consistent* are defined in the main text.

Table C.10: Variable Definitions

Variable	Description
Birthyear	Expectant mother's year of birth
Language	Indicator [=1] if respondent's mother tongue language is English.
Urban	Indicator [=1] if respondent resides in an Urban area.
Numer of Previous Children	Number of children born to respondent prior to current pregnancy.
Q1.Weekly Iron Intake	How many days did you take your iron pills last week? (Please reply a number from 0 to 7)
Q2.Difficulty Remembering [=1]	Do you find it difficult to remember to take the iron pills? (Reply with yes or no.)
Q3.Failure to Carry out plans	How often do you plan to do something and then don't do it? (Reply 0 (never), 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 (always))
Q4.Willingness to take risks	In general, how willing are to take risks? (Please reply on a scale from 0 (completely unwilling) to 10 (very willing)
Q5.Iron pills – Stomach feels bad	Keep going! Does your stomach feel bad if you take iron pills? (Reply yes or no.)
Q6.Weekly days tired / dizzy	How many days did you feel very tired or dizzy last week? (Please reply a number from 0 to 7.)
Q7.Low iron levels [=1]	Did your doctor or nurse tell you that you have low iron levels? (Reply 1) Yes 2) No 3) Not anymore 4) Don't know ... recoded to "yes" =1, else =0)
Q8.Years of Schooling	How many years did you go to school? Reply with number of years.

Notes: (i) The table describes variables used in the main analysis.

Table C.11: Adherence and Discounting. Robustness, Risk, Language, and Education

	OLS Nº of Pills (1)	OLS If Missed Pills (2)	OLS Nº of Pills (3)	OLS Missed Pills =1 (4)	OLS Nº of Pills (5)	OLS Missed Pills =1 (6)
Late Planners	-0.271**	0.087**	-0.272**	0.085**	-0.248*	0.074*
	(0.125)	(0.039)	(0.125)	(0.039)	(0.127)	(0.039)
Risk Taking	-0.004	0.004				
	(0.013)	(0.004)				
English Speaking			0.154	-0.078**		
			(0.114)	(0.037)		
Years of School					0.031	-0.014*
					(0.026)	(0.008)
Constant	6.798***	0.090***	6.686***	0.158***	6.404***	0.286***
	(0.085)	(0.031)	(0.095)	(0.034)	(0.335)	(0.097)
Failure Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Observations	480	480	480	480	480	480
R^2	0.016	0.016	0.020	0.024	0.021	0.024

Results are shown for the 480 participants who have done the full survey and use the full information on all attempts to do the task at any of the possible days. Risk-taking is a measure from 0 to 10, where the participant answers the flowing question: In general, how willing are to take risks? Please reply on a scale from 0 (completely unwilling) to 10 (very willing). English Speaking is a dummy variable where 1 indicate that English is their preferred language and, 0 indicate that they prefer another language. The variable Years of School is the participants reported number of years of education. High Discounters is a dummy variable where planning to do the task on day 1 is denoted 0 and not planning to do the task or planning to do the task on day 3 or 5 is denoted with a 1. Failure Dummy is a dummy for technical problems leading to payment delay. OLS Regressions. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table C.12: Adherence and Discounting - Restricted sample

	OLS Nº of Pills (1)	OLS Missed Pills =1 (2)	Logit Missed Pills =1 (3)
Late Planners	-0.238*	0.084**	0.582**
	(0.126)	(0.041)	(0.264)
Constant	6.760***	0.113***	-2.030***
	(0.069)	(0.025)	(0.222)
Failure Dummy	Yes	Yes	Yes
Observations	465	465	465
R^2	0.013	0.014	

From the 480 participants that answered all eight questions of the final survey, the regressions in this table exclude the 15 particpants answering that they do not plan to do the task or writing words backwards. Missed Pills =1 is a dummy variable which is 0 if the participant reported having taken all 7 pills last week, and which is 1 if she missed one pill or more. High Discounters is a dummy variable where planning to do the task on day 1 is denoted 0 and planning to do the task on day 3 or 5, or not planning to do the task is denoted with a 1. Failure Dummy is a dummy for technical problems leading to payment delay. Regressions include the 480 participants that answered all eight questions of the final survey. Robust standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1.

Table C.13: Prediction Analysis: Missed Pills =1

	Full sample Missed Pills=1 (1)	Full sample Missed Pills=1 (2)	Full sample Missed Pills=1 (3)	Est. sample Missed Pills=1 (4)	Est. sample Missed Pills=1 (5)	Est. sample Missed Pills=1 (6)
Age	-0.010***	-0.009**	-0.009***	-0.012**	-0.011**	-0.011**
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
English Speaking	-0.052	-0.056	-0.058	-0.100	-0.111*	-0.111*
	(0.039)	(0.039)	(0.038)	(0.061)	(0.061)	(0.061)
Years of Schooling	-0.014*	-0.012*	-0.012	-0.021*	-0.020*	-0.020*
	(0.008)	(0.007)	(0.007)	(0.011)	(0.011)	(0.011)
Number of Kids	0.002	-0.001	-0.001	0.007	0.006	0.006
	(0.019)	(0.019)	(0.019)	(0.029)	(0.029)	(0.029)
Urban	-0.057	-0.062*	-0.067*	-0.027	-0.032	-0.032
	(0.036)	(0.036)	(0.036)	(0.054)	(0.054)	(0.054)
Risk Taking	0.004	0.005	0.005	0.007	0.009	0.009
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
High Discounters		0.080**	0.085**		0.090	0.091
		(0.039)	(0.040)		(0.059)	(0.059)
Late Doers			0.104			0.028
			(0.065)			(0.117)
Constant	0.596***	0.524***	0.519***	0.904***	0.808***	0.773**
	(0.128)	(0.129)	(0.129)	(0.290)	(0.279)	(0.306)
Failure Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Observations	473	473	473	237	237	237
R^2	0.047	0.056	0.062	0.064	0.074	0.075
RMSE	0.365	0.364	0.363	0.379	0.378	0.379
RMSPE Out of Sample				0.393	0.386	0.378

Notes: The full sample include the 473 participants that answered all eight questions of the final survey and for whom all controls are available. Note 7 observations are dropped because we lack information on the number of kids. The estimation sample excludes a hold-out sample consisting of the 50% of women who are due last. RMSE and RMSPE are the Root Mean Square Error corrected for degrees of freedom of the sample used for estimation and hold-out sample respectively. Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

Table C.14: How Quickly They Answer Other Messages

	High Discounters (1)	Actually do task 1,3,5 (2)	Actually Delay (3)
Answer First Hour	-0.061	0.003	-0.002
	(0.048)	(0.071)	(0.025)
No Answer	-0.023	0.003	0.006
	(0.072)	(0.106)	(0.039)
Constant	0.394***	1.185***	0.064***
	(0.046)	(0.078)	(0.025)
Failure Dummy	Yes	Yes	Yes
Observations	480	460	460
R^2	0.014	0.001	0.000

Column 1 in this table include all 480 participants in our main sample. In column 2 and 3, the sample is restricted further, as the main explanatory variable here is when the participants did the word task. These regressions therefore only include participants who actually made an attempt to reply in this task, excluding 20 individuals. Variable definitions: No Answer is a binary variable for the 61 participants that did not reply to an additional question asking them too reply with their first 4 digits in their social security number. Answer First Hour is a dummy variable where 1 denotes a person that answered the 4 digits social security number question within the first hour, which 270 women did. Plan to Delay is a dummy variable where planning to do the task on day 1 is denoted 0 and planning not to do the task or planning to do the task on day 3 or 5 is denoted with a 1. The variable Actually Delay is a binary variable where 0 denotes if a person performs the task on day 1, and 1 denotes if a person delays to do the task untill to day 3 or 5. Failure Dummy is a dummy for technical problems leading to payment delay. OLS regressions. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table C.15: Day of the Week

	Late Planners (1)	Actually do task 1,3,5 (2)	Actually Delay (3)
Monday	-0.106	0.050	0.039
	(0.078)	(0.098)	(0.039)
Tuesday	-0.062	0.116	0.043
	(0.082)	(0.123)	(0.040)
Wednesday	-0.105	0.058	0.006
	(0.078)	(0.115)	(0.032)
Thursday	-0.082	0.163	0.063
	(0.082)	(0.127)	(0.042)
Friday	-0.061	0.106	0.051
	(0.085)	(0.119)	(0.042)
Saturday	0.038	0.011	0.004
	(0.090)	(0.109)	(0.034)
Constant	0.410***	1.122***	0.038
	(0.064)	(0.072)	(0.025)
Failure Dummy	Yes	Yes	Yes
Observations	480	460	460
R^2	0.021	0.007	0.010

Column 1 in this table include all 480 participants in our main sample. In column 2 and 3, the sample is restricted further, as the main explanatory variable here is when the participants did the word task. These regressions therefore only include participants who actually made an attempt to reply in this task, excluding 20 individuals. Late Planner is a dummy variable where planning to do the task on day 1 is denoted 0 and planning not to do the task or to do the task on day 3 or 5 is denoted with a 1. The variable Actually Delay is a binary variable where 0 denotes if a person performs the task on day 1, and 1 denotes if a person procrastinates to day 3 or 5. Failure Dummy is a dummy for technical problems leading to payment delay. OLS regressions. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table C.16: Vitamin Adherence and Discounting. Robustness, Nº Kids.

	OLS Nº of Pills (1)	OLS If Missed Pills (2)	Logit If Missed Pills (3)
Late Planners	-0.291**	0.089**	0.610**
	(0.126)	(0.040)	(0.258)
Number of Kids	0.015	-0.011	-0.080
	(0.052)	(0.020)	(0.161)
Constant	6.767***	0.121***	-1.968***
	(0.092)	(0.034)	(0.281)
Failure Dummy	Yes	Yes	Yes
Observations	473	473	473
R^2	0.018	0.016	

From the 480 participants that answered all eight questions of the final survey, the regressions in this table include only the 473 individuals from which we have information on how many kids they already have. The variable Number of Kids contains this information on how many kids the participant already has. Missed Pills =1 is a dummy variable which is 0 if the mother reported having taken all 7 pills last week, and which is 1 if she missed one pill or more. Late Planners is a dummy variable where planning to do the task on day 1 is denoted 0 and not planning to do the task or planning to do the task on day 3 or 5 is denoted with a 1. Failure Dummy is a dummy for technical problems leading to payment delay. OLS and Logit Regressions. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.