Pre-analysis Plan Outline for Sleepless in Chennai: The Consequences of Sleep Deprivation Among the Urban Poor

Pedro Bessone Tepedino — MIT
Gautam Rao — Harvard University
Frank Schilbach — MIT
Heather Schofield — University of Pennsylvania
Mattie Toma — Harvard University

December 8, 2017

Contents

1 Introduction **4**

2 Benchmark Models **4**

3 First-Stage Outcomes **8**

3.1 Sleep 8

3.1.1 Overview 8

3.1.2 Empirical Framework 8

3.2 Sleep deprivation/Sleepiness 8

3.2.1 Overview 8

3.2.2 Regressions 9

3.3 Self-Reported Sleep and Sleepiness 9

4 Earnings, Productivity, and Labor Supply **11**

4.1 Earnings and Labor Supply 13

4.1.1 Overview 13

4.1.2 Regressions 13

4.2 Productivity 13

4.2.1 Overview 13

4.2.2 Regressions 14

4.3 Decomposing earnings and productivity effects 14

4.3.1 Overview 14

4.3.2 Regressions 15

4.3.3 Heterogeneous Treatment Effects 15

5 Decision Making Outcomes **16**

5.1 Savings 16

5.1.1 Overview 16

5.1.2 Regressions 16

5.2 Defaults 17

5.2.1 Overview 17

5.2.2 Regressions 17

5.3 Present Bias 19

5.3.1 Overview. 19

5.3.2 Empirical Framework. 20

5.4 Limited Attention 23

5.4.1 Overview 23

5.5 Risk and Social Preferences 24

5.5.1 Overview 24

5.5.2 Regressions 25

6 Cognitive Outcomes **28**

6.1 Inhibitory Control 28

6.2 Memory 29

6.2.1 Regressions. 29

7 Happiness and Subjective Well-Being **30**

7.1 Overview. 30

7.2 Regressions. 30

1. Introduction

This study seeks to improve our understanding of the role of sleep deprivation in the lives of the poor in developing countries. We will estimate the impact of improved sleep on economic outcomes, decision-making, cognitive function, health outcomes and happiness.

The primary outcomes that we are interested in include: (i) work-related outcomes (labor supply, earnings, productivity) and (ii) decision-making (time preferences, default effects, inattention, risk preferences, social preferences).

Other secondary outcomes include: (i) tests of cognitive function, which are intended primarily as mechanisms for effects on the primary outcomes; (ii) health-related outcomes, pre-registered at clinicaltrials.gov, and not discussed here; and (iii) happiness.

We anticipate that we may write multiple papers on different outcomes (for example, a paper on the effect of sleep on labor market outcomes as well as a paper on the effect on decision-making). However, it is possible that all results will be combined into a single paper.

We will adjust p-values for multiple hypothesis testing as appropriate within classes of primary outcomes.

2. Benchmark Models

The experimental design involves multiple baseline and post-treatment observations per participant for the majority of outcomes. The main regression specifications will be ANCOVA regressions, which control for the average value of the outcome variable in the baseline period, as well as other individual-specific baseline control variables (such as average baseline sleep, age, or gender). The ANCOVA specification was selected based on Monte Carlo simulations using pilot data, in which the ANCOVA specification achieved higher power than models with individual fixed effects. For all regressions, we will cluster standard errors at the participant level.

The following variables are used in multiple models we estimate:

- $Treat\_S_i$ is the treatment group we call the "sleep group". Individuals in this group receive sleep devices, information, and encouragement to sleep. The randomization is between individuals.
- $Treat\_S\&I_i$ is the treatment group we call the "sleep and incentives group". Individuals in this group receive sleep devices, information, encouragement to sleep, *and* monetary incentives based on increases in their sleep level. The randomization is between individuals.

- $Treat\_Nap_i$ is a treatment arm where individuals are allowed to take naps of up to 30 minutes in the afternoon. The randomization is between individuals and is cross-randomized with the previous two interventions to improve sleep (i.e. with $Treat\_S_i$ and $Treat\_S\&I_i$).
- Age_i is the age bin that participant i is in, collected at baseline.
- $Female_i$ is an indicator equal to one when participant i is female.
- $SleepNap_{it}$ is the actual time asleep during participant i's nap at date t, as measured by the actigraph (more on this below).
- $SleepNight_{it}$ is the duration of participant i's sleep during the previous night (that is, the night from day $t-1$ to t) as measured by the actigraph (more on this below).
- $SleepTime_{it} = SleepNap_{it} + SleepNight_{it}$

Empirical Models

The following empirical models are the backbone of our analysis. We will use them repeatedly for different outcomes and will refer to them by their abbreviation.

(1) Basic Reduced-Form Equation

$$\begin{aligned} y_{it} = {} & \alpha_0 + \alpha_1 \text{Treat\_S}_i + \alpha_2 \text{Treat\_S\&I}_i + \alpha_3 \text{Treat\_Nap}_i \\ & + \alpha_4 \text{Treat\_S}_i * \text{Treat\_Nap}_i + \alpha_5 \text{Treat\_S\&I}_i * \text{Treat\_Nap}_i \\ & + \boldsymbol{\theta_1}' \textbf{Baseline}_i + \theta_2 \text{Female}_i + \theta_3 \text{Age}_i + \theta_4 \text{Other\_vars}_{it} \\ & + \boldsymbol{\gamma_1}' \textbf{Date FE}_t + \boldsymbol{\gamma_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \end{aligned} \tag{RF}$$

where

- y_{it} is an outcome to be specified for each context.
- **Baseline** is a vector of variables measured at the baseline (we will specify which variables we use for each outcome of interest).
- Date FE_t is a vector of dummies capturing the calendar date.
- Day in Study FE_{it} is a vector of dummies capturing the day in study of participant i at date t (i.e. for how many days participant i has been enrolled in the study on date t).
- Every regression will have the controls Baseline, Female, and Age. Other_vars captures other variables (specific to particular outcomes) that might be added for specific regressions. For each regression, we will specify which other variables are considered, if any.

Remark. We will run one specification with the demographic control variables, and one without them.

(2) IV Specification

We will estimate the impact of sleep on different outcomes using IV methods. The mediating variable is Sleep Time_{it}, which includes the time spent sleeping the previous night and the time spent napping on the day in question. This variable might be endogenous, so we will instrument it with the treatment variables.

The benchmark specification is

$$\begin{aligned} y_{it} = {} & \beta_0 + \beta_1 \text{Sleep Time}_{it} \\ & + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i + \gamma_3 \text{Age}_i + \gamma_4 \text{Other\_vars}_{it} \\ & + \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \end{aligned} \tag{IV}$$

The instruments for Sleep Time_{it} are Treat_S_i, $\text{Treat\_S\&I}_i$, Treat_Nap_i, and the interactions $\text{Treat\_S}_i * \text{Treat\_Nap}_i$ and $\text{Treat\_S\&I}_i * \text{Treat\_Nap}_i$.

(3) Heterogeneous Effects Regressions

Finally, we will estimate additional specifications for both the reduced-form and instrumental variable regressions to explore heterogeneous treatment effects. We will run the specifications below for every outcome variable unless stated otherwise. In some cases, we include additional interactions, and we state that when relevant. The specific interactions included in these extended regressions are outlined below:

Heterogeneous Effects Reduced-Form Equation

$$\begin{aligned} y_{it} = {} & \alpha_0 + \alpha_1 \text{Treat}_i + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i \\ & + \gamma_3 \text{Age}_i + \gamma_4 \text{Female}_i * \text{Treat}_i \\ & + \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \\ y_{it} = {} & \alpha_0 + \alpha_1 \text{Treat}_i + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i \\ & + \gamma_3 \text{Age}_i + \gamma_4 \text{Second Half}_{it} + \gamma_5 \text{Second Half}_{it} * \text{Treat}_i \\ & + \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \\ y_{it} = {} & \alpha_0 + \alpha_1 \text{Treat}_i + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i \\ & + \gamma_3 \text{Age}_i + \gamma_4 \text{Above Med Sleep}_{it} + \gamma_5 \text{Above Med Sleep}_{it} * \text{Treat}_i \\ & + \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \\ y_{it} = {} & \alpha_0 + \alpha_1 \text{Treat}_i + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i \\ & + \gamma_3 \text{Age}_i + \gamma_4 \text{Above Med Base}_{it} + \gamma_5 \text{Above Med Base}_{it} * \text{Treat}_i \\ & + \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \end{aligned} \tag{HRF}$$

Heterogeneous Effects IV Equation

We will estimate the following heterogeneous treatment effects regressions:

$$\begin{aligned}
y_{it} =& \alpha_0 + \alpha_1 \text{Sleep Time}_{it} + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i \\
&+ \gamma_3 \text{Age}_i + \gamma_4 \text{Female}_i * \text{Sleep Time}_i \\
&+ \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \\
y_{it} =& \alpha_0 + \alpha_1 \text{Sleep Time}_{it} + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i \\
&+ \gamma_3 \text{Age}_i + \gamma_4 \text{Second Half}_{it} + \gamma_5 \text{Second Half}_{it} * \text{Sleep Time}_{it} \\
&+ \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \\
y_{it} =& \alpha_0 + \alpha_1 \text{Sleep Time}_{it} + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i \\
&+ \gamma_3 \text{Age}_i + \gamma_4 \text{Above Med Sleep}_{it} + \gamma_5 \text{Above Med Sleep}_{it} * \text{Sleep Time}_{it} \\
&+ \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it} \\
y_{it} =& \alpha_0 + \alpha_1 \text{Sleep Time}_{it} + \boldsymbol{\gamma_1}' \textbf{Baseline}_i + \gamma_2 \text{Female}_i \\
&+ \gamma_3 \text{Age}_i + \gamma_4 \text{Above Med Base}_{it} + \gamma_5 \text{Above Med Base}_{it} * \text{Sleep Time}_{it} \\
&+ \boldsymbol{\delta_1}' \textbf{Date FE}_t + \boldsymbol{\delta_2}' \textbf{Day in Study FE}_{it} + \varepsilon_{it}
\end{aligned} \tag{HEIV}$$

We will use the sleep groups treatment assignments interacted with the interaction variables above as instruments for sleep time.

Standard interaction variables:

- *Female*$_i$ is an indicator for whether the participant is female.
- *Second Half*$_{it}$ is an indicator for whether the participant is in the first half or second half of their days in the treatment period (to determine whether there are cumulative effects of sleep such that the treatment effects increase over time).
- *Above Med Sleep*$_i$ is an indicator for whether the participant's baseline average sleep is above or below the median baseline average sleep in our study population (to determine whether our treatment has a stronger impact on, say, those who sleep less during the baseline period).
- *Above Med Base*$_i$ is an indicator for whether the participant's baseline average outcome variable y is above or below the median baseline average value for that outcome in our study population. (This specification is excluded when the outcome variable is sleep.)

3. First-Stage Outcomes

The first aim of our study is to establish a significant first stage, i.e. a meaningful impact of the randomized interventions on sleep. Below we explain in detail the variables capturing sleep time.

3.1. Sleep

3.1.1 Overview

Our primary first-stage measure will be sleep time over the last 24 hours (*Sleep Time*$_{it}$), including both sleep during the previous night and time spent napping on the day in question, as measured by actigraphs. Actigraphs are watch-like devices that participants wear on their wrists and that measure movement throughout the day; they are considered the gold standard for measuring sleep outside of sleep labs in the sleep literature (Jean-Louis et al. (1997); Blackwell et al. (2008); Schutte-Rodin et al. (2008)). Using the data collected from the actigraphs, we can measure when participants fall asleep, wake up in the middle of the night or in the morning, how long they sleep, how often they wake up, etc.

3.1.2 Empirical Framework

We will estimate the reduced-form model (RF, see Section 2) for the first stage where

- $y_{it} = \text{Sleep Time}_{it},\ \text{Sleep Night}_{it},\ \text{Sleep Nap}_{it}$
- **Baseline**$_i$ = $\frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$, i.e. the average sleep for participant i during the τ baseline days (before the sleep-related treatments are administered).

3.2. Sleep deprivation/Sleepiness

3.2.1 Overview

There is individual variation in the amount of sleep needed. Hence, to objectively measure participants' sleep *deprivation* (as opposed to sleep itself), we follow the literature in sleep medicine by asking participants to complete the Psychomotor Vigilance Task (PVT). The PVT is a test of reaction time and attention that has been used extensively in the sleep literature to measure sleep deprivation (Lim and Dinges (2008); Basner and Dinges (2011); Basner et al. (2011)).

The task consists of a stimulus appearing on the screen at random. Participants must then click on their mouse as soon as they spot the stimulus to score points in the task.

PVT performance is measured in terms of participant payments for the task. Based on piloting evidence, we decided to base payments on a weighted average of standard measures used in the sleep literature (Basner and Dinges, 2011). This weighted average includes (i) inverse reaction time (the average of 1/reaction times–determines 50% of the payment), (ii) false starts (instances where a participant presses a button before the appearance of a stimulus–determines 25% of the payment), and (iii) minor lapses (instances in which the participant does not respond to the stimulus within 500ms–determines 25% of the payment).

We also introduce experimental variation in the payment structure of the task. On 50% of the days, participants' given payments from the task are doubled. We denote the earnings from the PVT task when under low incentives rather than high incentives by $\textit{PVT Performance}_{it}$.

3.2.2 Regressions

The following reports first-stage regressions looking at the impact of our sleep treatment on PVT performance.

We will estimate the reduced-form model (RF) where

- $y_{it} = \textit{PVT Performance}_{it}$.
- $\textbf{Baseline}_i = \frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$, i.e., the average baseline performance for participant i during the τ baseline PVT measures.
- $\textit{Other\_vars}_{it} = \text{High Task Pay}_{it}$, an indicator variable that equals one if participant i faces high incentives on day t.

An additional specification will include the interaction of the sleep treatment with High Task Pay$_{it}$.

3.3. Self-Reported Sleep and Sleepiness

We will also explore the impact of the sleep treatments on self-reported sleep and sleepiness.

Self-reported sleep is measured by the number of hours a participant reports sleeping the previous night. In addition, we will measure self-reported sleep quality using the Pittsburgh Sleep Quality Index. Self-reported sleepiness is measured by standard measures in the sleep literature, e.g. the Karolinska Sleepiness Scale.

These measures will not however serve as primary first-stage outcome measures comparable to the actigraph and PVT data. Rather, they will be used to shed light on how participants perceive the changes in sleep due to the treatments, as compared to the objective actigraph and PVT measures.

4. Earnings, Productivity, and Labor Supply

Participants are presented with the task of entering (artificially created) data into a computer. We design the work task to closely mimic an application form filled out by hand in English, as it is common in the data-entry industry.

The participants complete the task on a computer with a split screen. On the left, the participants see the image of the text form they are meant to transcribe. On the right, they see the text they are inputting through the computer's keyboard. The text displayed on the screen is partitioned into fields of text. Participants cannot move to the next field until they finish the previous one. They are paid by the number of correct entries, and are penalized for each incorrect entry. The number of mistakes is computed using the Levenshtein distance (loosely speaking, the number of additions and deletions that would need to be made to perfectly match the text) within field of text.

While typing, participants face time-varying incentives. The piece rates are randomized to (potentially) vary every 30 minutes to be either high or low. Under high piece rates, participants earn 2 paise (Rs. 0.02) for each correct character typed, and under low piece rates, participants earn .5 paise (Rs. 0.005) for each correct character typed. The penalty rate is fixed at 10 paise (Rs. 0.1) per mistake.

Participants have the option to type for between six and ten hours a day, depending on the day in question. Specifically, 6 of the 28 days are randomly assigned to be "short" days in which participants only have the option to type for about three hours and no longer; 12 of the 28 days are randomly assigned to be "long" days when participants have the option to type up to ten hours; and the remaining 10 days are fixed (rather than randomly assigned) "special" days when the day has a varying length of typing time available to accommodate certain special tasks that a participant must complete, e.g. the present bias task described below. The variation in short versus long days was introduced to facilitate the direct analysis of the impact of the sleep treatment on labor supply. On short days, we try to guarantee only minimal variations in labor supply, allowing us to estimate the effect of sleep on productivity controlling for labor supply. On long days, both productivity and labor supply (length of work time) might be affected.

Participants are free to take breaks at any point during their allotted typing time at the office. Similarly, there are no minimum work requirements.

Main Variables of Interest

- *Time at Office*$_{it}$: Time spent at the office, which is defined as the difference between the time the participants leave and the time they arrive at the office.
- *Time Typing*$_{it}$: Hours (continuous variable) participant i worked on the typing task at date t. We include fraction of hours worked on the measure. *Time Typing*$_{it}$ is defined as the difference between *Time at Office*$_{it}$ and the sum of the following two variables:

  1. Mandatory Pauses - Any time of the day when the participant is doing other scheduled tasks (e.g. cognitive tests, lunch, etc). The participants are not allowed to type during mandatory pauses.
  2. Voluntary Pauses - Any time the participant is scheduled to type but has not interacted with the computer for at least 2 minutes is counted as a voluntary pause. The rationale behind this variable is to capture labor supply responses at the intensive margin. For example, sleep deprivation might cause participants to take voluntary breaks more often.

- *Performance Payment*$_{it}$: The total payment participants receive per day based on their performance while typing. Participants get 2 paise/correct entry under high incentives, 0.5 paise/correct entry under low incentives and a penalty rate of Rs. 0.1/incorrect entry (i.e. 10 paise/correct entry).
- *Attendance Payment*$_{it}$: The payment participants receive per day for the time spent typing. This measure is independent of the output produced. The formula for the attendance payment is Attendance Payment$_{it} = 21.60 *$ Time Typing$_{it}$.
- *Typing Earnings*$_{it}$ = Performance Payment$_{it}$ + Attendance Payment$_{it}$
- *Number of Correct Entries*$_{it}$: The number of characters a participant enters correctly per day.
- *Fraction High Pay*$_{it}$: The fraction of sessions with high piece rate *assigned* to participant i at t. Note that we will not use the the actual time spent typing under each session, but the randomly determined assignment.

  Remark: We set all variables above to zero when participant i is absent on day t .
- *Long Day*$_{it}$: is a dummy variable equal to 1 if the worker i is assigned to a long day at date t.
- *Special Day*$_{it}$: is a dummy variable equal to 1 if the worker i is assigned to a special day at t. Special days are day 1 and day 2 (training period), day 8 (treatment assignment), days 7 and 26, when there are risk and social tasks, days 6, 14, 19 and 25, when workers might do overtime work from the present bias task, and day 28, the last day in the study.

4.1. Earnings and Labor Supply

4.1.1 Overview

The first two primary outcomes related to the typing task are (i) daily earnings; (ii) labor supply, as measured by time typing and time at the office.

4.1.2 Regressions

We will estimate the impact of sleep on the primary outcomes outlined above with the models RF and IV, with the following variables:

- y_{it} = IHS Transformed Typing Earnings$_{it}$[1], Time Typing$_{it}$, Time at Office$_{it}$.
- **Baseline**$_i$ = $\frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$, i.e. participant i's average value for the corresponding outcome variable during his or her τ baseline days.
- $Other\_vars_{it}$ = Fraction High Pay$_{it}$, Long Day$_{it}$, Special Day$_{it}$.

As described in section 4.3.3, we will also run this specification with interactions with the dummy for Long Day$_{it}$ and Special Day$_{it}$ to separate out the productivity effects on short days (where labor supply responses are likely to be more limited) from those on longer days.

4.2. Productivity

4.2.1 Overview

The third primary outcome related to the data-entry task is productivity. Our preferred productivity measure is quality-adjusted output divided by time typing (excluding all breaks). Output produced daily is given by the formula

$$\begin{aligned} \text{Output} &= \text{Correct Entries} - \frac{\text{Penalty Rate}}{(\text{High Piece Rate} + \text{Low Piece Rate})/2} \times \text{Mistakes} \\ &= \text{Correct Entries} - 8 \times \text{Mistakes} \end{aligned} \tag{1}$$

where the second line uses the values of piece rates and penalty implemented on the study. The rationale behind this formula is that it captures the trade-off between quantity (speed) and

[1] An inverse hyperbolic sine transformation is defined as $f(y) = ln(y + \sqrt{1+y^2})$. It has the benefit that, except for values of y close to zero, $f(y) \approx ln(2) + ln(y)$. We employ the IHS transformation here, because $Typing\ Earnings_{it} = 0$ if participant i is absent on day t. The coefficients can be interpreted as the standard log transformation away form zero.

quality (accuracy) of output, using as relative weight the compensation ratio of correct entries and mistakes as established by the employer (in our study, us).

4.2.2 Regressions

We will estimate the impact of sleep on productivity with the models RF and IV, with the following variables:

- y_{it} = Productivity$_{it}$
- **Baseline**$_i$ = $\frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$ of participant i, i.e. the average productivity for participant i during his or her τ baseline days.
- *Other_vars*$_{it}$ = Fraction High Pay$_{it}$, Long Day$_{it}$, Special Day$_{it}$.

As described in section 4.3.3, we will also run this specification with interactions with the dummy for Long Day$_{it}$ and Special Day$_{it}$ to separate out the productivity effects on short days (where labor supply responses are likely to be more limited) from those on longer days.

4.3. Decomposing earnings and productivity effects

4.3.1 Overview

In secondary analyses, we will decompose the effect on earnings and productivity further into three potential channels: (i) speed, (ii) accuracy and (iii) sensitivity to incentives. These variables are defined as follows:

$$\text{Speed}_{it} = \log\left(\frac{\text{Total Entries}_{it}}{\text{Time Typing}_{it}}\right)$$
$$\text{Accuracy}_{it} = \frac{\text{Incorrect Entries}_{it}}{\text{Number of Entries}_{it}}$$

i.e. speed is measured by the log of the total number of entries submitted per hour worked, while accuracy is measured by the percentage of mistakes per submitted character.

Finally, the sensitivity to incentives is measured by considering how productivity and labor supply responds to changes in wages for treatment and control groups.

4.3.2 Regressions

We will use RF and IV to estimate the reduced-form and IV regressions for speed and accuracy, where:

- y_{it} = Speed$_{it}$, Accuracy$_{it}$
- **Baseline**$_i$ = $\frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$, i.e. participant i's performance during his or her τ baseline days.
- *Other_vars*$_{it}$ = Fraction High Pay$_{it}$, Long Day$_{it}$, Special Day$_{it}$

For the analysis of sensitivity to incentives, we will estimate the following RF specification

$$\begin{aligned} y_{it} =& \alpha_0 + \alpha_1 \text{Treat}_i + \alpha_2 \text{Treat\_Nap}_i + \alpha_3 \text{Default}_{it} \\ &+ \alpha_4 \text{Treat}_i * \text{High Piece Rate}_{it} + \alpha_5 \text{Treat\_Nap}_i * \text{High Piece Rate}_{it} \\ &+ \boldsymbol{\theta_1}' \textbf{Baseline}_i + \theta_2 \text{Female}_i + \theta_3 \text{Age}_i + \theta_4 \text{Other\_vars}_{it} \\ &+ \text{Date FE}_t + \text{Day in Study FE}_{it} + \varepsilon_{it}, \end{aligned} \tag{2}$$

where

- y_{it} = Time Typing$_{it}$, Productivity$_{it}$, as outlined above;
- **Baseline**$_i$ = $\frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$ of participant i.
- Treat$_i$ pools both night sleep treatment groups.

To directly estimate the effect of sleep on sensitivity, we will estimate a similar regression, substituting the treatment assignment variable for the total sleep time variable described above. The regression will be estimated by IV, using treatment assignment variables and treatment assignment variables interacted with high piece-rate to instrument for sleep time and sleep time interacted with high piece-rate.

4.3.3 Heterogeneous Treatment Effects

We will estimate models (HRF) and (HEIV) for every outcome variable outlined in 4.1.2, 4.2.2 and 4.3.2, using the standard interaction variables outlined in the Benchmark Models section.

In addition, we will also explore the effects in the sub-sample of long and short days to capture the impact of sleep on labor supply (long days) and productivity controlling for labor supply (short days).

5. Decision Making Outcomes

5.1. Savings

5.1.1 Overview

At the end of each work day, participants can choose to deposit up to Rs. 400 using either some or all of their study earnings or any other money that they may have available. Participants can withdraw any amounts from their existing savings on any study day (but the overall savings amount cannot be negative, i.e. participants cannot borrow any money from the study). The daily savings will be counted as zero if the participant does not show up on a given day. During the baseline period (i.e. before the sleep treatments are assigned), all participants face a 1% daily interest rate. On the day of the sleep treatment assignment (day 8), half of the sample (half of each treatment group) will be randomly assigned to receive a 2% daily interest rate for the remaining days of their participation in the study. The remaining individuals will continue to face a 1% daily interest rate.

5.1.2 Regressions

The main savings regressions will look at daily net savings as an outcome measure, i.e. the difference between daily deposit and withdrawal amounts. We will use RF and IV to estimate the impact of sleep time and treatment for daily net savings, where:

- y_{it} = *Daily Net Savings*$_{it}$
- ***Baseline***$_i = \frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$, i.e. participant i's average daily net savings during his or her τ baseline days.
- *Other_vars*$_{it}$ = *Default Amount*$_{it}$, *Fraction High Pay*$_{it}$, *Max Pay Cognitive Tasks*$_{it}$, *Present Bias*$_{it}$ $*$ *Piece Rate*$_{it}$, *Risk and social day*$_{it}$, *Interest Rate*$_{it}$, ***Surveyor FE***$_{it}$.
  where:
  1. *Default Amount*$_{it}$: The amount that was defaulted into savings on a given day (either Rs. 0 or Rs. 40; see the below defaults section);
  2. *Fraction High Pay*$_{it}$: as defined above (Section 4).
  3. *Max Pay Cognitive Tasks*$_{it}$: The maximum amount that a participant could have earned during cognitive tasks during the day. This variable is not based on the participant's actual (endogenous) performance. Rather, it is the sum of the maximum payments possible from all the cognitive tasks the participant takes in a given day, which is randomly determined (as described below).

4. *Present Bias*$_{it}$, *Piece Rate*$_{it}$:, A dummy variable equal to one if the participant receives the payment for the present bias task on that day. "Piece Rate" is the randomized piece rate that the participant is paid for the present bias task on that day.
5. *Risk and social day*$_{it}$: A dummy variable equal to one if the participant completed the risk and social tasks that day.
6. *Interest Rate*$_i$: The interest rate that the participant faces after the baseline period: 1% or 2%.
7. ***Surveyor FE***$_{it}$: A vector of dummies capturing which surveyor conducted the savings survey for participant i on day t.

Heterogeneous treatment effects:

Apart from the standard analysis outlined in the benchmark section, we will also estimate how the treatment effect varies with the interest rate level, by estimating the models (HERF) and HEIV using *Interest Rate*$_i$ as an additional interaction variable.

5.2. Defaults

5.2.1 Overview

Starting on the second day of the study, payments for completing the daily survey (Rs. 40) are either paid out in cash at the end of the day or put into the lockbox (i.e. defaulted into savings) in the morning, right after the daily survey. That is, default savings are either Rs. 0 or Rs. 40. Once participants reach the end of the day, they make their savings decisions where they can choose to save any additional amount (up to the daily maximum of Rs. 400, including any amount defaulted into savings) or to withdraw any amount that is in the lockbox, including any payment deposited earlier in the day (default savings) or in previous days. Accordingly, default savings are not binding in any way. Participants learn whether their survey payment will be defaulted into their savings account after completing their daily survey but will not be reminded again of their default type over the course of the day.

Participants are randomly assigned to receive the default either once every three days or twice every three days. We randomize defaults in that manner to investigate whether timing or expectation of the defaults are driving participants' responses.

5.2.2 Regressions

There are two main regression specifications.

Specification 1

The first default specification follows the savings specification, but includes an interaction between the default amount (Rs. 0 or Rs. 40) and the sleep treatment, our key outcome of interest. Since we are adding a variable that takes the *value* of the default, rather than a 0-1 variable, the coefficients associated with it should be interpreted as differences in pass-through[2] of defaults between treatment and control groups.

Intuitively, we want to measure whether participants in the sleep treatment save less on days when the payment is defaulted into savings than do control participants, i.e. whether they are less susceptible to the default.

The estimation builds on the RF and IV specifications outlined in the benchmark section, but the main variable of interest is the interaction of either the sleep treatment assignment (RF case) or the sleep duration variables (IV cases) with $\textit{Default Amount}_{it}$

We will run the RF regression as below:

$$\begin{aligned} y_{it} =&\alpha_0 + \alpha_1 \text{Treat}_i + \alpha_2 \text{Treat\_Nap}_i + \alpha_3 \text{Default}_{it} \\ &+\alpha_4 \text{Treat}_i * \text{Default}_{it} + \alpha_5 \text{Treat\_Nap}_i * \text{Default}_{it} \\ &+ \boldsymbol{\theta_1}' \textbf{Baseline}_i + \theta_2 \text{Female}_i + \theta_3 \text{Age}_i + \theta_4 \text{Other\_vars}_{it} \\ &+ \text{Date FE}_t + \text{Day in Study FE}_{it} + \varepsilon_{it}, \end{aligned} \tag{3}$$

where

- $\textit{Treat}_i$: A dummy variable equal to one if the participant is either in the sleep treatment group or in the sleep and incentives treatment group. $\textit{Treat\_Nap}_i$ is a dummy variable equal to one if the participant is in the nap treatment group.
- $\textit{Default}_{it}$: A variable taking the value 40 or 0, depending on the amount defaulted into the participant's saving account.
- $y_{it} = \textit{Daily Net Savings}_{it}$, where "Daily Net Savings" is defined as "Daily Deposits" minus "Daily Withdrawals," as above.
- $\boldsymbol{\textit{Baseline}}_i = \frac{1}{\tau} \sum_{t=1}^{\tau} y_{it}$, i.e. participant i's average daily net savings during his or her τ baseline days.
- $\textit{Other\_vars}_{it}$ = $\textit{Fraction High Pay}_{it}$, $\textit{Max Pay Cognitive Tasks}_{it}$, $\textit{Present Bias}_{it} * \textit{Piece Rate}_{it}$, $\textit{Risk and Social day}_{it}$, $\textit{Interest Rate}_{it}$, $\boldsymbol{\textit{Surveyor FE}}_{it}$. The variable definitions are as above.

The IV specification is analogous but with $\textit{SleepTime}_{it}$ in place of the treatment variables.

[2] The difference between the treatment and control groups in the percent of the defaulted payment that the participants save due to the default effect alone

Specification 2

In the second specification, the key outcome of interest is whether the sleep treatment (or additional sleep in the IV regression) affects an individual's propensity to be an active saver, i.e. to make decisions defying the defaulted amount.

The specification we employ to capture that uses an indicator of whether the participant followed the default amount as an outcome variable (saving an amount other than Rs. 0 or Rs. 40, depending on whether the payment was defaulting into cash versus savings). We will run the RF regression as below:

$$\begin{aligned} y_{it} =& \alpha_0 + \alpha_1 \text{Treat}_i + \alpha_2 \text{Treat\_Nap}_i \\ &+ \alpha_3 \text{Default Type}_{it} + \alpha_4 \text{Treat}_i * \text{Default}_{it} + \alpha_5 \text{Treat\_Nap}_i * \text{Default}_{it} \\ &+ \boldsymbol{\theta_1}' \textbf{Baseline}_i + \theta_2 \text{Female}_i + \theta_3 \text{Age}_i + \theta_4 \text{Other\_vars}_{it} \\ &+ \text{Date FE}_t + \text{Day in Study FE}_{it} + \varepsilon_{it} \end{aligned} \tag{4}$$

where

- $Treat_i$ and $Treat\_Nap_i$ are defined as in Specification 1 above.
- Default Type$_{it}$ is a dummy variable equal to one if the daily survey payment is defaulted into savings account rather than given to participants in cash.
- y_{it} is a dummy variable equal to 1 if the participant makes an active savings decision.
- $\boldsymbol{Baseline}_i = \frac{1}{\tau} \sum_{t=1}^{\tau} y_{it}$, i.e. participant i's average daily net savings during his or her τ baseline days.
- $Other\_vars_{it}$ = Fraction High Pay$_{it}$, Max Pay Cognitive Tasks$_{it}$, Present Bias$_{it}$ $*$ Piece Rate$_{it}$, Risk and Social day$_{it}$, Interest Rate$_{it}$, ***Surveyor FE***$_{it}$. The variable definitions are as above.

Again, the IV specification is analogous but with $SleepTime_{it}$ in place of the treatment variables.

5.3. Present Bias

5.3.1 Overview.

In the present bias task, we will elicit participants' choices over how many pages of data they would like to enter for different piece rates. One of these choices will be randomly selected to be implemented as extra work after their regular work day is finished. We elicit choices on two dates: once either one or two days before the participant completes the extra work; and once just before the extra work time. If the choice implemented was elicited on date t and the participant successfully completes the number of pages they had committed to, they are paid for the present bias task on date $t+3$. The payment consists of a lump-sum payment of Rs. 100 plus the number of pages completed times the piece rate implemented. If the participant does not complete the number of pages they had committed to, they receive nothing for the task.

We have two goals with the Present Bias task:

1. Estimate how sleep deprivation impacts time preferences and time consistency.
2. Estimate individual parameters of present bias and long-term discounting to correlate with other outcomes in the study.

5.3.2 Empirical Framework.

We will implement three empirical strategies to achieve our goals.

Semi-parametric estimation of present bias

To achieve goal 1 semi-parametrically, we will estimate the linear model below, where the unit of observation is the decision of how many pages to work for, given a piece-rate offer.

$$\begin{aligned} y_{itc} = {} & \alpha_0 + \boldsymbol{\alpha_1}\textbf{Prospective}_{itc} + \alpha_2 T_i + \alpha_3 Nap_i \\ & + \boldsymbol{\alpha_4}\textbf{Prospective}_{itc} \cdot T_i + \boldsymbol{\alpha_5}\textbf{Prospective}_{itc} \cdot Nap_i \\ & + \theta_1 \text{Baseline}_{ic} + \theta_2 \text{Female}_i + \theta_3 \text{Age}_i + \theta_4 w_{itc} \\ & + \text{Date FE}_t + \text{Day in Study FE}_{it} + \textbf{Surveyor FE}_{it} + \varepsilon_{itc}, \end{aligned} \tag{5}$$

where

- y_{itc}: The number of pages chosen on decision c;
- Prospective_{itc} : A vector of dummy variables indicating whether decision c is a decision made 1 day before or 2 days before the extra work day. The excluded group is decisions made on the day of the extra work.
- T_i : A dummy variable equal to one if the participant is either in the sleep treatment group or in the sleep and incentives treatment group. Nap_i is a dummy variable equal to one if the participant is in the nap treatment group.
- w_{itc} is the piece rate (wage) offered at decision c.
- Baseline_{ic} : Is the number of pages chosen in the baseline period for the prospective date

with wage offer w_{itc}, if the participant was offered that piece-rate during baseline. If he/she wasn't, this variable takes value 0.

The main parameter of interest in this regression is

$$\alpha_4 = (\mathbb{E}[y|T=1,P=1] - \mathbb{E}[y|T=1,P=0]) - (\mathbb{E}[y|T=0,P=1] - \mathbb{E}[y|T=0,P=0])$$

which measures the difference in time-consistency between sleep and control groups. If sleep deprivation decreases time-consistency, we expect $\alpha_4 < 0$.

Similarly for α_5, we can measure the difference in time consistency for the nap treatment.

The difference in the components of the vector of the parameters $\boldsymbol{\alpha_1}$, $\boldsymbol{\alpha_4}$,and$\boldsymbol{\alpha_5}$ can be used to test whether the long-term discount rate varies across treatment groups.

Additional specifications include (i) interacting the treatment variables and prospective dummies with a dummy for wage offered; (ii) interacting the former two variables with individuals fixed effect to estimate the distribution of the reduced form parameters in the population; (iii) collapsing the prospective dummies into a single dummy indicating whether decision was taking prospectively.

We will also try to run an IV specification. Here the IV is trickier, because the mediating variable we have been using so far, total sleep on the last night and naps, varies within day. However, the decisions on this task are made in different dates, and we would like to have a proxy for sleep deprivation that does not change from day to day. Thus, we will explore cumulative measures of sleep as endogenous variables for the IV.

Structural Estimation

The structural estimation has two goals. First, it estimates parameters with a structural interpretation, which is useful when comparing our results with the estimates from the literature in different settings. Second, one of the versions of the model will be estimated with individual-specific short-term and long-term discounting parameter. This allows us to achieve goal 2 described above.

We estimate a model assuming that participants' choice of pages come from maximizing the utility function given by[3]

$$U(e, w, k, t, T) = -\beta^{\{k>t\}}\delta^{t-k}C(e, \gamma) + D_m(T-k)U_m(e \cdot w), \tag{6}$$

where

[3]Our structural estimation closely follows the empirical framework in Augenblick and Rabin (2017).

- e is the number of pages chosen at piece rate w.
- T is the date of payment, t is the date of the work, and k is the date of the choice.
- $C(e, \gamma)$ is a cost of effort function parametrized by γ and $U_m(\cdot)$ is a function that captures the utility from money (Augenblick and Rabin (2017); Augenblick (2017)).
- The β and δ parameters are the standard quasi-hyperbolic time preference parameters (Laibson (1997); O'Donoghue and Rabin (1999)).
- $D_m(\cdot)$ is the monetary discounting function.

For the estimation, we will assume specific parametric forms for both utility functions and for the monetary discounting function. We impose a minimum and a maximum bound for the number of pages participants can commit to work on. For that reason, we will also use censored regression techniques to deal with corner solutions.

We will run two specifications of the structural model:

1. Estimating a set of parameters for the treatment group (either incentive or no incentive sleep groups), and a different set of parameters for the control group, assuming all the heterogeneity in preferences in the population is captured by treatment status.

2. Assume individual-specific parameters to estimate the distribution of each parameter in the population.

Remark. We will not be able to estimate the structural parameters for individuals who show significant non-monotonicity with respect to wages, as the individual-specific parameters will not converge. The same is true for individuals who choose strictly decreasing number of pages per wage offered. This follows Augenblick and Rabin (2016).

From the first procedure, we will recover two sets of estimates of β and δ, one for the treatment group, $\hat{\beta}_T$, $\hat{\delta}_T$ and the other for the control group $\hat{\beta}_C$, $\hat{\delta}_C$.

Then we will test the null hypotheses

$$\mathbf{H}_0^{\beta} : \beta_T = \beta_C$$

$$\mathbf{H}_0^{\delta} : \delta_T = \delta_C$$

using the estimation from the structural model.

5.4. Limited Attention

5.4.1 Overview

We would like to test whether sleep affects how much individuals pay attention to information that affect earnings and are therefore relevant to decision making.

To do this, we created two ways in which piece rates in the data-entry task are displayed to participants: a high-salience and a low-salience condition. The key hypotheses to test are, first, whether responses to incentives are tempered when participants are in the low-salience compared to the high-salience condition (a check on whether the salience variation works “as intended”); and second, whether the sleep treatments increase individuals’ reaction to changes in incentives thus reducing the discrepancy between the high-salience and low-salience conditions.

To implement these tests, we created variation in the salience of incentives. There is a bar at the bottom of the screen during the typing task that indicates how much participants earn for each correct entry. During the first 5 days of the study, individuals are in the “high-salience" condition. The bar indicating payments is green when the piece rate is high (2 paise for each correct entry) and white when the piece rate is low (.5 paise for each correct entry). Furthermore, the screen flashes twice every 30 minutes, when the incentive scheme has a 50% probability of changing, making participants highly aware of the change if it happens. The bar flashes every time, even if the incentive scheme does not change.

Starting on day 6, participants are randomly assigned to a “low-salience" condition on half the days. During the remaining days, participants work in the “high-salience" condition as before. In the “low-salience" condition, the piece-rate bar is grey throughout the day, so the changes in the incentive scheme are less salient on these days. Moreover, the screen does not flash to indicate possible changes in the incentive scheme.

We will consider the response to incentives in the sleep treatment versus the control group, on either side of the incentive switches, and under the salient versus non-salient conditions.

The idea is that the response to incentives under the salient condition reveals the “optimal" response. We will compare this with the response to incentives in the non-salient incentives condition. The comparison allows us to estimate the degree of inattention, as in Chetty et al. (2009) or DellaVigna (2009). We will further compare the degree of inattention across the sleep treatment and control groups.

We do not fully pre-specify this analysis since the appropriate analysis will depend upon first establishing that the salience variation works “as intended” – that is, the response to incentives is stronger under high-salience relative to low-salience. Moreover, the precise form of the response to salience – e.g. whether individuals notice and respond to the incentive change in 5 minutes versus 30 minutes – remains unknown at the time of this pre-registration, making it difficult to write down the full contingent analysis plan.

5.5. Risk and Social Preferences

5.5.1 Overview

Participants complete the following risk and social preferences activities twice throughout the study, once during the baseline period and once during the treatment period:

1. **Dictator Game:** A standard dictator game in which participants choose how much of a Rs. 50 endowment to give to an unknown recipient.
2. **Ultimatum Game:** A standard ultimatum game in which subjects choose how much of a Rs. 50 endowment to offer to an unknown recipient, with the possibility that this recipient will reject the offer. There are two versions of this activity, one where the participant acts as the sender and one where she acts as the receiver. When the participant acts as the receiver, she will make accept/reject decisions for each amount that she might receive from a sender, before she actually knows how much was sent.
3. **Trust Game:** A standard trust game in which participants choose how much of Rs. 50 they want to offer to an unknown recipient. The offered amount that the recipient receives is tripled. The recipient then decides how much of the money they received to send back to the sender. The Rs. 50 used in this task are taken from a participant's daily earnings. Again, there are two versions of this activity, one in which the participant acts as the sender and one in which she acts as the receiver.
4. **Risk Preference Task:** Here participants are presented with a series of choices in which they indicate whether they prefer a risky or a safe option. Participants who choose the safe option A get Rs. 50 for sure while those who choose the risky option B are faced with a 50-50 coin toss that dictates their earnings. The choices the participants face vary in terms of the lower bound that they might earn after choosing the risky option B. That is to say, if the coin comes up Tails in the risky option the participant always earns Rs. 100, but if it comes up Heads she might gain anywhere from Rs. 0 to Rs. 40, depending on the choice in question. We are interested in the point at which the participant switches from choosing the safe option to the risky option.
5. **Loss Aversion Task:** Here participants are asked to choose whether they would accept a series of gambles in which they have a 50-50 chance of losing versus some gaining some

amount. If they reject the gamble, they neither lose nor win money. If they accept the gamble, there is a 50% chance that they gain Rs. 50 and a 50% chance that they lose Rs. 10 to Rs. 60, depending on the choice in question, as above. Again, we are interested in the point at which the participant switches to rejecting versus accepting the gamble

5.5.2 Regressions

1. Dictator Game

   We will use RF and IV to estimate the reduced-form and IV regressions for the amount sent in dictator games.

   - y_{it} = Amount Sent_{it}
   - Baseline_i = The amount that participant i sent in the ultimatum game in the baseline period.
   - $Other\_vars_{it}$ = $\textbf{Surveyor FE}_{it}$: A vector of dummies capturing which surveyor conducted the savings survey for participant i on day t.

2. Ultimatum Game Sender:

   We will use RF and IV to estimate the reduced-form and IV regressions for the amount sent as the sender in ultimatum games, where:

   - y_{it} = Amount Sent_{it}
   - Baseline_i = The amount that participant i sent in the ultimatum game in the baseline period.
   - $Other\_vars_{it}$ = $\textbf{Surveyor FE}_{it}$: A vector of dummies capturing which surveyor conducted the savings survey for participant i on day t.

3. Ultimatum Game Recipient:

   We will treat each choice for each level of sent amount by each recipient as a separate observation and run the reduced-form regression as below. We will also look separately at the smaller amounts received where the effect is likely to be the strongest.

$$\begin{aligned} \text{y}_{itc} = {} & \alpha_0 + \alpha_1 \text{Treat\_S}_i + \alpha_2 \text{Treat\_S\&I}_i + \alpha_3 \text{Treat\_Nap}_i \\ & + \alpha_4 \text{Treat\_S}_i * \text{Treat\_Nap}_i + \alpha_5 \text{Treat\_S\&I}_i * \text{Treat\_Nap}_i \\ & + \theta_1 \text{Baseline}_{ic} + \theta_2 \text{Female}_i + \theta_3 \text{Age}_i + \boldsymbol{\gamma_1}' \textbf{Amount Received}_{itc} \\ & + \boldsymbol{\gamma_2}' \textbf{Date FE}_t + \boldsymbol{\gamma_3}' \textbf{Day in Study FE}_{it} + \boldsymbol{\gamma_4}' \textbf{Surveyor FE}_{it} + \varepsilon_{itc} \end{aligned}$$

   - y_{itc}: an indicator variable equal to one when the participant rejects the amount sent.

- Baseline_{ic} : a dummy equal to one when participant i rejected the sender's offer c in the baseline period.
- $\textbf{Amount Received}_{itc}$, a vector of levels of sent amount by the ultimatum sender.

The IV specification is analogous but with the sleep time substituting the treatment variables.

4. Trust Game Sender:

   We will use RF and IV to estimate the reduced-form and IV regressions for the amount sent as the sender in trust games, where:

   - $y_{it} = \text{Amount Sent}_{it}$
   - Baseline_i = The amount that participant i sent in the trust game in the baseline period.
   - $Other\_vars_{it}$ = $\textbf{Surveyor FE}_{it}$: A vector of dummies capturing which surveyor conducted the savings survey for participant i on day t.

5. Trust Game Recipient:

   We will treat each choice for each level of sent amount by each recipient as a separate observation and run the reduced-form regression as below. We will also look separately at the smaller amounts received where the effect is likely to be the strongest.

$$\begin{aligned}
\text{y}_{itc} &= \alpha_0 + \alpha_1\text{Treat_S}_i + \alpha_2\text{Treat_S\&I}_i + \alpha_3\text{Treat_Nap}_i \\
&+ \alpha_4\text{Treat_S}_i * \text{Treat_Nap}_i + \alpha_5\text{Treat_S\&I}_i * \text{Treat_Nap}_i \\
&+ \theta_1\text{Baseline}_{ic} + \theta_2\text{Female}_i + \theta_3\text{Age}_i + \boldsymbol{\gamma_1}'\textbf{Amount Received}_{itc} \\
&+ \boldsymbol{\gamma_2}'\textbf{Date FE}_t + \boldsymbol{\gamma_3}'\textbf{Day in Study FE}_{it} + \boldsymbol{\gamma_4}'\textbf{Surveyor FE}_{it} + \varepsilon_{itc}
\end{aligned}$$

   - y_{itc}: the amount that participant i send back when given c on day t.
   - Baseline_{ic} : the amount that participant i sent back when given c in the baseline period.
   - $\textbf{Amount Received}_{itc}$, a vector of levels of sent amount by the trust sender.

   The IV specification is analogous but with the sleep time substituting the treatment variables.

6. Risk Aversion:

   The outcome variable is the participant's switching point, defined as the riskiest option for which he or she chooses the lottery rather than the certain outcome. We will estimate the reduced-form regression at the day and individual level as below.

$$\begin{aligned}
\text{y}_{it} &= \alpha_0 + \alpha_1\text{Treat_S}_i + \alpha_2\text{Treat_S\&I}_i + \alpha_3\text{Treat_Nap}_i \\
&+ \alpha_4\text{Treat_S}_i * \text{Treat_Nap}_i + \alpha_5\text{Treat_S\&I}_i * \text{Treat_Nap}_i \\
&+ \theta_1\text{Baseline}_i + \theta_2\text{Female}_i + \theta_3\text{Age}_i \\
&+ \boldsymbol{\gamma_2}'\textbf{Date FE}_t + \boldsymbol{\gamma_3}'\textbf{Day in Study FE}_{it} + \varepsilon_{it}
\end{aligned}$$

- y_{it}: The value of the switch point, as defined above.
- Baseline_i : The value of the switch point in the baseline period.

The IV specification is analogous but with the sleep time substituting the treatment variables.

7. Loss Aversion:

   The RF and IV specifications for the loss aversion task are same as the risk aversion task.

6. Cognitive Outcomes

We will study the participants' performance in two cognitive tasks: (i) inhibitory control as measured by the Hearts and Flowers Task, and (ii) memory as measured by the Corsi Block-Tapping Test. We vary the incentives for each of these tasks. Individuals face either high or low incentives, where the incentives are randomly assigned at the day-task level. In the high-incentives condition, the payments are twice as high as they are in the low-incentives condition.

6.1. Inhibitory Control

Inhibitory control will be measured using the Hearts and Flowers Task. The task consists of three different rounds, and only the third round is relevant to the participants' payments. In the third round, either a flower or a heart will appear on one side of the screen (40 times). The participant then has to press the correct button corresponding to the shape that appears (participants press a button on the same side of the screen for a heart and on the opposite side for a flower). Inhibitory control is reflected in the participant's ability to press the button that corresponds to the shape that appears on the screen. For each round, a response is correct only if the shape is still present on the screen when the correct button is pressed; reaction time is therefore also important in this task.

Participants are paid for this task based on their accuracy (fraction correct) and speed (mean response time, conditional on being correct) in the task. The payment structure for the high and low incentives versions of Hearts and Flowers is based on participants' performance in pilot studies. We use standardized outcome measures for the payment scheme, both to ensure that the incentives correspond to the outcome measures in which we are interested and to ensure that the payment range that lines up with that in the other cognitive tasks: Rs. 5 to Rs. 20 under low incentives.

We will use RF and IV to estimate the reduced-form and IV regressions for task performance in Hearts and Flowers.

- y_{it} = H&F performance$_{it}$, defined as the variable payment that participants would earn if they faced low incentives.
- **Baseline**$_i$ = $\frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$, i.e. participant i's average performance on this task during τ baseline days.
- $Other\_vars_{it}$ = High Task Pay$_{it}$, a dummy equal to one if participant i faces high incentives in day t for the Hearts and Flowers Task.

6.2. Memory

The Corsi task serves to measure participants' short-term working memory. In the task, blocks appear one after another in a pattern on the computer screen. The task starts with a short sequence of 2 blocks appearing in a row and increases one block at a time until it reaches a sequence of 9 blocks. The participant is asked to reproduce the pattern once it has played out on the screen by touching the place where each block appeared on the screen, in the order the blocks appeared. In other words, if two blocks appear one-by-one in a pattern on the screen, the participant must first touch the place on the screen where the first block appeared and then touch the place where the second block appeared. For each sequence length (1-9), participants attempt the task two times.

Participants are paid based on the maximum sequence of blocks (1-9) that they correctly reproduce.

As in the Hearts and Flowers task, the incentive schemes are constructed based on standardized outcome measures in previous pilot data such that the range of payments under low incentives is between Rs. 5 and Rs. 20.

6.2.1 Regressions.

We will use RF and IV to estimate the reduced-form and IV regressions for the Task Performance in Corsi tasks.

- y_{it} = Corsi performance$_{it}$, defined as the variable amount that participants would earn if they faced low incentives.
- **Baseline**$_i$ = $\frac{1}{\tau} \sum_{t=1}^{\tau} y_{it}$, i.e. participant i's average performance during τ baseline days.
- $Other\_vars_{it}$ = High Task Pay$_{it}$, a dummy equal to one if participant i faces high incentives in day t for the Corsi Task.

Additional specifications will include interactions with High Task Pay$_{it}$ for both H&F and Corsi.

7. Happiness and Subjective Well-Being

7.1. Overview.

We measure participants' happiness and subjective well-being via two key survey questions:

1. *Today, would you say you are:*

0. Not at all happy.
1. A little unhappy.
2. Somewhat happy.
3. Happy.
4. Very happy.

2. *Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time? Choose answer among options ranging from 0 to 10.*

7.2. Regressions.

We will use RF and IV to estimate the reduced-form and IV regressions for the two above outcomes.

- y_{it} = Happiness Score_1$_{it}$ (captured from survey question 1 above), Happiness Score_2$_{it}$ (captured from survey question 2 above).
- **Baseline**$_i$ = $\frac{1}{\tau}\sum_{t=1}^{\tau} y_{it}$, i.e. participant i's average happiness score on the corresponding survey question during his or her τ baseline days.

References

Augenblick, N, "Short-Term Time Discounting of Unpleasant Tasks," 2017. Working Paper.

— **and M Rabin**, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," 2017. Working Paper.

Basner, M and DF Dinges, "Maximizing Sensitivity of the Psychomotor Vigilance Test (PVT) to Sleep Loss," *Sleep*, 2011, *34* (5), 581–591.

— **, D Mollicone, and DF Dinges**, "Validity and Sensitivity of a Brief Psychomotor Vigilance Test (PVT-B) to Total and Partial Sleep Deprivation," *Acta Astronautica*, 2011, *69* (11-12), 949–959.

Blackwell, T, S Redline, S Ancoli-Israel, JL Schneider, S Surovec, NL Johnson, JA Cauley, and KL Stone, "Comparison of Sleep Parameters from Actigraphy and Polysomnography in Older Women: the SOF Study," *Sleep*, 2008, *31* (2), 283–291.

Chetty, R, A Looney, and K Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

DellaVigna, S, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 2009, *47* (2), 315–372.

Jean-Louis, G, H von Gizycki, F Zizi, A Spielman, P Hauri, and H Taub, "The Actigraph Data Analysis Software: I. A Novel Approach to Scoring and Interpreting Sleep-wake Activity," *Perceptual and Motor Skills*, 1997, *85* (1), 207–16.

Laibson, D, "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

Lim, J and DF Dinges, "Sleep Deprivation and Vigilant Attention," *Annals of the New York Academy of Sciences*, 2008, *1129* (1), 305–322.

O'Donoghue, T and M Rabin, "Doing it Now or Later," *American Economic Review*, 1999, *89* (1), 103–124.

Schutte-Rodin, S, L Broch, D Buysse, C Dorsey, and M Sateia, "Clinical Guideline for the Evaluation and Management of Chronic Insomnia in Adults," *Journal of Clinical Sleep Medicine*, 2008, *4* (5), 487–504.