The Economic Consequences of Increasing Sleep Among the Urban Poor*

Pedro Bessone† Gautam Rao‡ Frank Schilbach
Heather Schofield Mattie Toma

January 2020

Abstract

This paper measures sleep among the urban poor in India and estimates the economic returns to increased sleep. Adults in Chennai have strikingly low quantity and quality of sleep relative to typical guidelines: despite spending 8 hours in bed, they achieve only 5.6 hours per night of sleep, with 32 awakenings per night. A three-week treatment providing information, encouragement, and sleep-related items increased sleep quantity by 27 minutes per night without improving sleep quality. Increased night sleep had no detectable effects on cognition, productivity, decision-making, or psychological and physical well-being, and led to small decreases in labor supply and thus earnings. In contrast, offering high-quality naps at the workplace increased productivity, cognition, psychological well-being, and patience. Taken together, the returns to increased night sleep are low, at least at the low-quality levels typically available in home environments in Chennai. We find suggestive evidence that higher-quality sleep improves important economic and psychological outcomes. *JEL Codes: O1, I1, C9, D9*

*We gratefully acknowledge generous funding and support by the Government of Tamil Nadu, the Abdul Latif Jameel Poverty Action Lab, IFMR LEAD, the William F. Milton Fund, the Harvard Global Health Institute, the Program on the Global Development of Aging, the SHASS Research Fund, the Mind, Brain, and Behavior Interfaculty Initiative, the Institute for Translational Medicine and Therapeutics (ITMAT), the Weiss Family Program Fund for Research in Development Economics, and the Pershing Square Venture Fund for Research on the Foundations of Human Behavior. Research reported in this publication was also supported by the National Institute on Aging of the National Institutes of Health under award number K01AG055691, as well as the Program for Training in Sleep, Circadian and Respiratory Neurobiology in the Division of Sleep Medicine at Harvard Medical School and Brigham & Women's Hospital, which is supported by a Ruth L. Kirschstein NRSA (T-32 grant) from the National Heart, Lung and Blood Institute within the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank numerous seminar audiences and colleagues for helpful comments and generous feedback. We thank Alosias A, Srinivas Balasubramanian, Phoebe Cai, Stephanie Chan, Fiona Chen, Thomas Escande, Gabriel Jardanovski, Gunjita Gupta, Erik Hausen, Yihong Huang, Dexin Li, Andrew Locke, Maddie McKelway, Adrien Pawlik, João Pugliese, Jane von Rabeau, Cory Rand, Maya Roy, Krishna Prasad Srinivasan, Nikkil Sudharsanan, Ziqing Yan, and Yanzun Yang for excellent research assistance. We thank all our study participants and staff, all those who took the time to participate in our expert survey, and sleep medicine experts Dr. Dinges, Dr. Basner, Dr. Redline, and Dr. Perlis for their help and guidance on this study. We received IRB approval from Harvard University, protocol number 14-2294, the Institute for Financial Management and Research in India, and the Indian Council of Medical Research. The experiment was pre-registered with a pre-analysis plan on the AEA registry, number AEARCTR-0002494.

†Bessone: MIT (tepedino@mit.edu); Rao: Harvard and NBER (grao@fas.harvard.edu); Schilbach: MIT and NBER (fschilb@mit.edu); Schofield: UPenn (hschof@wharton.upenn.edu); Toma: Harvard (mattietoma@g.harvard.edu)

‡Corresponding author. Address: 1805 Cambridge St, Cambridge, MA 02138. Phone: +1-734-846-7754

1 Introduction

Sleep is the most time-consuming activity of our lives, taking a third of our time. Lab experiments and observational studies in rich countries show that sleep deprivation has severe impacts on cognitive function, health, and psychological well-being (Lim and Dinges, 2010; Banks and Dinges, 2007; Strine and Chapman, 2005). Based on this research, sleep experts recommend 7 to 9 hours of sleep per night (Hirshkowitz et al., 2015). Since many do not follow this advice, experts warn of a sleep "epidemic" and predict that increasing sleep will yield profound benefits (Walker, 2017; Rosekind et al., 2010). They argue that people sleep too little due to a lack of awareness of the effects of sleep deprivation (Van Dongen et al., 2003), self-control problems (Avery et al., 2019), and a disrupted circadian rhythm due to exposure to artificial light (Touitou et al., 2017).

Sleep could be particularly poor in cities in developing countries, where living environments are replete with irritants like noise, heat, and physical discomfort. Indeed, using wristwatch-like sensors which infer sleep from movement, we find alarmingly low quantities of sleep among low-income workers in Chennai, India: just 5.6 hours per night, with 95% sleeping less than 7 hours on average. Since the returns to sleep are thought to be diminishing, these low levels may imply high marginal benefits to increasing sleep. Consistent with this view, 118 experts from economics and sleep science predicted sizable economic benefits of increasing sleep by half an hour in this setting.

However, the returns to increasing sleep in field settings such as ours may be lower than found in sleep laboratories. First, people may choose their sleep quantity optimally, implying that marginal benefits equal to marginal opportunity costs. Second, the benefits of sleep may depend not just on quantity, but also on its quality (Ohayon et al., 2017). We document that sleep quality in Chennai is strikingly low: Sleep efficiency – time asleep divided by time in bed – is only 70% in our sample, much lower than in US populations or even those with disorders such as sleep apnea (Roure et al., 2008; Walker, 2017). Sleep is also highly fragmented, with 32 awakenings per night, of which 9 are longer than 5 minutes, in a typical night. Despite the low baseline quantity of sleep, additional low-quality sleep may not have the rejuvenating effects of normal sleep.

To investigate the economic impacts of increased sleep outside the lab, we recruited 452 low-income adults in Chennai for a full-time data-entry job for one month. We cross-randomized the sample into two types of interventions designed to increase sleep: (i) night sleep treatments that offered participants information about their sleep, verbal and/or financial encouragement to increase night sleep,[1] and items to improve their home-sleep environments,[2] and (ii) a nap treatment that offered participants the daily opportunity for a half-hour nap in the afternoon in a quiet space in the office.[3]

[1] In most places in the analysis, we pool across two treatments, one which provides only verbal encouragement, and the other which additionally provides daily financial incentives. To avoid mechanical differences in income effects across groups due to the financial incentives, participants in all other groups received identical streams of daily payments, unrelated to their sleep.

[2] Participants could take home as many items as they wanted from the following list: mattress, cot, pillow, eye shades, earplugs, and a table fan. In practice, use of the devices at home was low. Part of the control group also received placebo items unrelated to sleep, which had no effect on outcomes.

[3] The nap treatment is compared to two types of randomized controls: taking a compulsory break of the same

The night-sleep interventions increased sleep by an average of 27 minutes per night, which represents a large gain compared to alternative interventions such as sleeping pills (Riemann and Perlis, 2009). This increase in time asleep was entirely driven by greater time spent in bed—on average 38 additional minutes per night—rather than improved sleep efficiency or gains in other measures of sleep quality. The efficacy of the night-sleep treatments did not vary systematically with baseline sleep quantity, quality, or other observables.

Naps were also effective at increasing sleep. Eighty-eight percent of individuals fell asleep at some point during their allotted nap time, yielding an average of 13 minutes of nap sleep per day.[4] This nap sleep was higher quality than night sleep, using comparable measures of sleep efficiency and awakenings.

Increased night sleep did not have any measurable positive impacts on a range of outcomes.[5] The night-sleep treatment groups did not exhibit significantly higher productivity, despite working on a relatively cognitively-demanding task—data entry—intended to be sensitive to sleep deprivation. The increased time in bed among the night-sleep treatment group came at the cost of lowering labor supply by nine minutes per day, lowering earnings by 3% compared to the control group. We can clearly reject the median expert prediction of a 7% increase in output, since earnings increases of 1% or higher are outside the estimated 95% confidence interval. We also do not detect any impacts on the other outcomes, including psychological well-being, physical health, and time, risk, and social preferences. Nor do we find effects on a standard test of attention designed by sleep researchers to detect sleep deprivation.

Taken together, we find no evidence that participants under-invest in the quantity of night sleep on average, despite sleeping only 5.6 hours per night, substantially less than recommended guidelines. Given their current sleeping conditions, the short- and medium-run marginal benefits of sleep are low. In contrast, the opportunity costs are high: it takes 86 more minutes in bed to produce an hour of sleep. Of course, there could be longer-term benefits from increased sleep on health, cognition, or productivity, which we do not capture.

In contrast, naps increased work productivity by 2.3%, boosted a measure of attention (+0.17 s.d.), and raised an index of psychological well-being (+0.13 s.d.). Naps also increased patience as measured both by reduced present bias in a real-effort task ($\beta = 0.98$ versus $\beta = 0.92$) and 14% higher deposits in a savings account. We strongly reject that the nap treatment had no effect on all outcomes in a permutation-based joint test ($p < 0.001$, see Figure VII). These findings serve as a proof of concept that sleep *can* affect important economic outcomes relatively quickly. As a result of the productivity gains, naps increased workers' earnings compared to those who were randomized

length, or working through the break.

[4] Crowd-out of sleep from the interventions was modest: on average, the nap treatment crowded out five minutes of night sleep the following night, while those assigned to night sleep treatment groups napped just as much as others when offered naps.

[5] To test whether the pooled night-sleep treatment affects *any* outcome, we conduct a permutation-based Wald test of joint significance against the null of zero effect on each of the main outcomes, in the spirit of Young (2019). We fail to reject the null ($p = 0.85$) when excluding labor supply as an outcome, but reject it ($p = 0.014$) when including labor supply (Figure VII). Since labor supply was *reduced* by the night-sleep treatments, we interpret these results as evidence of the lack of any positive effect of the night-sleep treatments.

to take an enforced break of the same length. However, naps entail significant opportunity costs, and reduced total output and therefore earnings compared to working through the same period.[6]

We investigate the role of sleep quality in explaining our findings. First, while the average individual sees no increase in productivity from increasing night sleep, those with higher-quality baseline sleep *do* benefit from additional sleep. Treated participants with above-median baseline sleep quality reduced the time working by 11 minutes but see productivity increases of 3.7%. These changes offset each other, leaving earnings unchanged despite of the reduction in labor supply. This increase in productivity is substantial, given the relatively inelastic nature of this task, in which quadrupling the piece-rate increases productivity by only 10%.

The importance of quality is also consistent with the effectiveness of naps, which are provided in a relatively comfortable office environment in which measured sleep quality is substantially higher than in home environments. Together, these results suggest that high sleep quality may be essential to unlock the benefits of sleep. However, we cannot rule out other interpretations of the surprising lack of benefit of increased night sleep despite the low baseline levels. Little field evidence on the economic returns to sleep exists even from rich countries where sleep quality is high. It could be that the returns are low everywhere, and that the existing laboratory evidence from induced sleep deprivation does not generalize to increases in sleep in the field or to economically-relevant outcomes with meaningful stakes.

Our paper makes four contributions. First, we point to the value of taking an economic perspective on sleep, as pioneered by Biddle and Hamermesh (1990). Night sleep is endogenously chosen by individuals. Under standard economic assumptions, individuals should not leave high-return sleep unexploited. Indeed, returns to an additional 30 minutes of sleep are low for our study participants. In contrast, the nap treatment is effective and proves that additional sleep can positively impact economic outcomes. Since naps of this quality are typically unavailable at home and in the control group, the high returns do not pose an economic puzzle. To clarify the underlying economic logic, we provide a theoretical model in which sleep quantity and quality both affect productivity, and agents must choose how to allocate time to work, labor and leisure. The model illustrates how sleep quantity responds to changes in quality, and contrasts the effects of increasing quality versus providing incentives for increased sleep quantity. Our approach differs from sleep science, which typically focuses on identifying sleep levels that maximize performance, without weighing the opportunity costs or allowing for endogenous sleep choice.

Second, we provide causal estimates of the effect of sleep on economic outcomes in the field. We study policy-relevant increases in sleep, in contrast to experiments in sleep science that typically induce acute sleep deprivation. We build on a recent economics literature which uses natural experiments to demonstrate that sleep can have sizable effects on wages (Gibson and Shrader, 2018), hospitalizations (Jin and Ziebarth, 2020), accidents (Smith, 2016), and civic behaviors (Holbein et al., 2019). The differences of our findings from these papers may be due to our focus on a developing country, where sleep quality is low due to poor living environments. Our paper complements Jagnani

[6]This negative impact on earnings declined over the course of the treatment period, with nearly equal earnings (despite lower labor supply) between the nap and work groups by the end of the study.

(2018), who exploits variation in sunset times in India to show that less time in bed is associated with worse educational outcomes. The relatively large returns to sleep in this context suggest that children may be farther away from optimal sleep levels, or that sleep quantity may matter more for children and for learning outcomes.

Third, we employ objective measurements of sleep to measure sleep in a developing country, and document low quantity and quality of sleep, consistent with limited recent evidence (Schokman et al., 2018; Castro et al., 2013). Our findings contrast with studies using self-reported measures of sleep, which find only moderate prevalence of sleep deprivation in the developing world (Stranges et al., 2012; Gildner et al., 2014; Simonelli et al., 2018). Self-reports of sleep have been found to be unreliable in rich-country settings (Lauderdale et al., 2008; Girschik et al., 2012). Our results point to the value of collecting objective measures of sleep quantity and quality in developing-country populations.

Finally, we contribute to a recent literature in behavioral development economics which studies how the everyday experience of living in poverty affects psychology and economic behaviors (Kremer et al., 2019; Dean et al., 2019). Recent work has studied how experiencing financial constraints (Mani et al., 2013; Kaur et al., 2019), alcohol consumption (Schilbach, 2019), malnutrition (Schofield, 2014), or noise (Dean, 2019) affect economic outcomes. We document a novel aspect of the lives of the poor, low levels of sleep, and provide proof of concept that increasing sleep – in the form of naps – can affect economic and psychological outcomes. Yet, we find no evidence that individuals under-invest in sleep at the margin they can control.

The rest of this paper proceeds as follows: Section 2 discusses different methods to measure sleep and presents our findings on baseline levels of sleep quantity and quality among adults in Chennai. Section 3 describes the experimental design. Section 4 reports the effect of our interventions on sleep. We then describe the impact of our interventions on work outcomes (Section 5), well-being and cognitive function (Section 6), and preferences and decision-making (Section 7). Section 8 discusses our findings and concludes.

2 Measuring Sleep in Chennai

2.1 Measuring Sleep Outside the Lab

The gold standard for measuring sleep is polysomnography (PSG), which records brain waves, blood oxygen levels, eye movements, and body movements to determine sleep/wake cycles and stages of sleep (Marino et al., 2013). While highly accurate, this technology is impractical for field studies as it is bulky and requires multiple wire attachments to the participant. Thus, its use is restricted to the lab. Lab studies using PSG are typically of short duration, often lasting just one night or a week or two at most. In addition, they typically feature small sample sizes due to the challenging logistics and high cost.

In contrast to polysomnography, self-reported sleep provides a notoriously unreliable measure that usually only correlates moderately with objective sleep measures. Individuals asked to report their sleep instead tend to report the hours spent in bed, frequently leading to over-reporting of sleep duration (Lauderdale et al., 2008; Schokman et al., 2018). Self-reported measures of sleep quality similarly correlate only modestly with objective measures (Girschik et al., 2012). Since experimental interventions can create experimenter demand effects (potentially causing people to over-report sleep) or other biases (e.g. treated participants may pay more attention to awakenings), objective measures of sleep are essential to measure sleep rigorously in experimental trials.

Technological advances in sleep-measurement techniques greatly facilitate experimental field studies. Wristwatch-like devices, known as actigraphs, can be worn in one's natural home environments and objectively measure sleep by inferring wake/sleep states from movement. Actigraph measures of sleep *quantity* are considered highly reliable: Comparisons between actigraph and PSG measures show high degrees of accuracy (86%) and sensitivity (97%) in sleep-wake detection, with 90% minute-by-minute agreement between the two (Marino et al., 2013; Sadeh et al., 1995). Actigraphs have been found to provide valid and clinically useful measures of sleep quantity even in people with poor sleep (Kushida et al., 2001; Smith et al., 2018), and consistently capture treatment effects on sleep of various interventions (Sadeh, 2011).

Actigraphs also provide several measures of sleep *quality*. We focus on three widely-used measures in this paper: sleep efficiency, number of awakenings, and longest sleep episode (Ohayon et al., 2017; Ancoli-Israel et al., 2003). First among those, sleep efficiency is defined as time asleep divided by time in bed. This measure is available since – in addition to measuring the number of hours an individual is asleep – actigraphs also detect when an individual is in bed, but not asleep.[7] Second, by examining the minute-by-minute readings of wakefulness, we also measure the number of awakenings, i.e. instances of substantial movement during sleep. Actigraph-recorded sleep efficiency and awakenings data do not significantly differ from PSG data, and the National Sleep Foundation takes such variables to be appropriate measures of sleep quality (Kushida et al., 2001; Ohayon et al., 2017). Finally, while actigraphs are unable to measure sleep stages, they can capture the longest sleep episode in a given night, where longer episodes indicate less interrupted sleep. Less interrupted sleep has been shown to be beneficial even holding fixed total sleep quantity, potentially because it interferes with the natural progression of stages of sleep (Stepanski, 2002).

2.2 Sleep Deprivation Around the World

Sleep experts recommend that adults sleep 7 to 9 hours per night (Watson et al., 2015).[8] Relative to this guideline, sleep deprivation is common in high-income countries. Different samples of US adults have found average night sleep of 6 to 7 hours per night when measured objectively using actigraphs (Lauderdale et al., 2008; Jackson et al., 2018; Cespedes et al., 2016). Similarly, in the United Kingdom, the Mental Health Foundation finds that only 38% of adults are "good sleepers" (UK National Health Services, 2011). Sleep deprivation is therefore often described as a public-health

[7]We also collect daily self-reports on time in bed. These largely line up with the actigraph measures of time in bed, see Appendix Figure A.I.

[8]This guideline refers to actual time asleep, not merely time in bed. Those two numbers tend to be fairly close in healthy rich-country populations, where sleep efficiency typically exceeds 85 to 90 percent (Walker, 2017).

imperative in rich countries (Luyster, 2012; Walker, 2017).

Evidence on sleep is more limited in developing countries. The WHO-SAGE study of over 40,000 adults from rural areas in 8 countries documents a moderate overall prevalence of self-reported sleeping problems (16.6%), with large variation across countries and a higher prevalence of sleeping problems among women, the elderly, and and lower-education individuals (Stranges et al., 2012; Gildner et al., 2014). In the Indian component of this study, 4,500 older Indian adults from largely rural self-report average sleep duration of 7.1 hours (Gildner et al., 2014). About 30% of these individuals report sleeping six or fewer hours per night (Selvamani et al., 2018). Given that people tend to substantially overestimate their own sleep, such self-reported measures likely underestimate the true extent of sleep deprivation (Lauderdale et al., 2008). Additional challenges to sleep among the urban poor – including heat, noise, light, crowding, or physical discomfort – raise the concern that sleep deprivation may be even more widespread in urban areas.

A small number of recent studies have measured sleep using actigraphy in low- to middle-income countries. Schokman et al. (2018) finds an average of only 6.0 hours per night among 175 adults from Sri Lanka, while Knutson (2014) finds nearly 7 hours per night among 58 adults in Haiti. However, both studies document very low sleep quality compared to standard guidelines. As we describe in the next subsection, we find low levels of both sleep quality and quantity in two samples in Chennai.

2.3 Sleep in Chennai

We measure sleep in two samples of adults in Chennai. The first is our experimental sample. Details of the recruitment of this sample are deferred to Section 3.4, where we lay out the full experimental design. To summarize, the sample comprised of 452 adults between 25 and 55 years of age who were recruited to participate in a month-long study where their primary task was doing data-entry work (see Table A.II for sample characteristics). We measured their sleep throughout this period using actigraphs, and observe 8 nights of baseline sleep data for each participant before they were randomized to a treatment or control condition. The second source – our "Sleep Survey" – collects data on a broader survey of 3,833 individuals living in Chennai, with 439 participants also completing three nights of actigraphy. This data is described in greater detail below.

Baseline Sleep Patterns. For the experimental sample, we capture the following measures related to sleep quantity: (i) time in bed, (ii) minutes asleep at night, (iii) nap-sleep minutes, and (iv) 24-hour sleep (night sleep plus nap sleep). In addition, we use the following measures of sleep quality: (v) sleep efficiency (night sleep divided by time in bed), (vi) number of awakenings per hour of sleep, and (vii) longest uninterrupted episode of sleep. Each of these measures was captured via actigraphy. In addition, participants self-reported measures (i) through (iii), and 'self-reported' efficiency can be calculated from self-reported time in bed and time asleep. Unless otherwise noted, all sleep results presented below refer to actigraphy-based measures.

Figure I illustrates three features of the sleeping patterns of participants in our sample: (a) time in bed, (b) time asleep, and (c) sleep efficiency (time asleep/time in bed), both self-reported (upper panel) and objectively measured using actigraphs (lower panel).

Time in Bed. Workers in our sample spend considerable time in bed, as indicated by nearly identical averages of self-reports and actigraph measures of time in bed. At baseline, the average study participant spent roughly 8 hours per night in bed (Figures Ia and Ib). These averages resemble reported hours in bed in US samples: Kurina et al. (2015), for instance, found in their study of older Americans that the average time in bed is 8.4 hours, while Jackson et al. (2018), in a multi-ethnic study of adults across several US states, found that the average time in bed per night is 7.2 hours.

Self-Reported Sleep. Average baseline self-reported sleep duration in our study is 7.2 hours (Figure Ic). This average is slightly above the average of 7.1 hours found in the representative WHO-SAGE survey among the elderly in rural India described in Gildner et al. (2014). For comparison, averages of self-reported sleep duration in US range from 6.8 to 7.9 hours per night (Jackson et al., 2018; Lauderdale et al., 2008; Watson et al., 2015). Based on the self-reports, 41% of workers in our sample sleep on average fewer than 7 hours per night.

Actigraph-Measured Sleep. In contrast, the actigraphs provide evidence of striking levels of sleep deprivation relative to guidelines (Figure Id). Only 5% of participants slept more than 7 hours per night on average, the lower bound of recommended sleep level for adults, and 71% slept less than 6 hours (Hirshkowitz et al., 2015). Average sleep duration is only 5.6 hours per night. Variation around the mean is modest, with an average within-person standard deviation of 0.8 hours per night. This duration is significantly lower than typical actigraph-measured sleep duration in the US (6 to 7 hours per night, Lauderdale et al. (2008); Jackson et al. (2018); Cespedes et al. (2016)).

Low Sleep Quality. Sleep quality in our study population is alarmingly low, as measured by sleep efficiency, number of awakenings, and longest sleep episode per night. Figure II illustrates this low sleep quality by providing two examples from our actigraph data of minute-by-minute sleep/wake status. A typical night in our sample, shown in the upper panel, features sleep efficiency of around 70 percent, is highly fragmented with about 32 awakenings per night, and exhibits few uninterrupted sleep episodes that are longer than half an hour. In contrast, few study participants enjoy the “normal" sleep efficiency of 85 to 95 percent illustrated using a selected observation in the lower panel.

Average baseline sleep efficiency in our sample is only 70% (Figure If), much lower than typical estimates of sleep efficiency in rich countries between 85 to 95 percent (Yoon et al., 2003; Carrier et al., 2001; Cole et al., 1992; Walker, 2017). It is also far below recommended levels from an expert panel of sleep scientists, who judged sleep efficiency above 85% to indicate high sleep quality (Ohayon et al., 2017). Indeed, sleep efficiency in our sample is substantially below that even of US-based patients suffering from sleep disorders such as sleep apnea (Roure et al., 2008).[9]

[9]Low sleep efficiency in our sample is not just a result of poor sleep in the early evenings or late mornings. Rather,

The other metrics of sleep quality also indicate poor sleep quality (Table A.VII). The average participant experiences about 32 awakenings in an average night, translating to 4 awakenings per hour of sleep achieved, comparable to the levels observed in sleep-disordered populations (e.g. insomniacs) in the United States (Lichstein et al., 2006). Taking a more conservative definition of awakening where the disruption must last at least 5 minutes rather than 1 minute, we still find an average of about 10 such awakenings per night, compared to expert guidelines of 4 or fewer per night (Ohayon et al., 2017).

Similarly, the average longest sleep episode—the duration of the participant's longest episode of uninterrupted sleep—in a night is 55 minutes, substantially shorter than among individuals with disrupted sleep in the United States (in a sample of incontinent and physically restrained nursing home residents in the US, for example, the average longest sleep episode was 84 minutes) (Alessi et al., 2005). A complete sleep cycle of non rapid eye movement (NREM) followed by rapid eye movement (REM) sleep takes from 70 to 120 minutes (Walker, 2017). Fragmented sleep which is punctuated by awakenings is thought to reduce the benefits of different stages of sleep (Stepanski, 2002). Even the longest stretch of uninterrupted sleep in our sample is not sufficient, on average, to accommodate a full sleep cycle.

Sleep Survey in Broader Population. To investigate the representativeness of our RCT sample, we conducted a supplementary "Sleep Survey" with 3,833 individuals across different randomly-sampled neighborhoods in Chennai. In addition, a subset of 439 of these participants completed three nights of actigraph measurements. Participants in the sleep survey were not screened on any of the criteria used for the RCT. Yet, the nighttime sleep quantity and quality in this broader sample resembles to that in the RCT, with an average of 5.5 hours of sleep per night and 71% sleep efficiency. In addition, naps are relatively common in this population, with 37% of individuals napping on any given day. Sleep quantity does not vary substantially by household income, education, or employment status, with the caveat that the sample includes few high-income households. The Sleep Survey is discussed in greater detail in Appendix F.

Summary. Taken together, we find clear evidence of low quantity of sleep relative to typical guidelines among adults in Chennai. The primary driver of limited sleep appears to be strikingly low sleep efficiency rather than little time in bed. Other measures of sleep quality are similarly poor. These findings hold both in our experimental sample and in broader sample of low- and middle-income adults in Chennai. Our data are consistent with the intuition that the urban poor in developing countries face challenging sleep environments. Survey responses from daily surveys highlight the importance of environmental factors in interfering with nighttime sleep, with over half of the respondents indicating that cold or heat, noise, and/or light disrupt their sleep (Figure A.III). The low levels of sleep imply widespread sleep deprivation and raise the possibility of substantial economic and health benefits of increasing sleep.

---

sleep efficiency remains around 70% even between 1 and 5 am (when almost everyone is in bed), consistent with very poor and disrupted sleep throughout the night (Appendix Figure A.II).

3 Experimental Design and Empirical Framework

Figure III provides an overview of the experimental design and timeline of the study. We recruited 452 adults to participate in our study. Participants worked for 28 days in an office in central Chennai, spending most of their workdays doing paid data-entry work. Enrollment took place on a rolling basis over the course of 18 months. The office contained computer work-stations for data-entry, a break room, booths for surveys and experimental tasks, and nap stations on a separate floor.

3.1 Interventions to Increase Sleep

For the first eight days of the study, all participants remained in a control condition, allowing us to collect baseline data on their sleep, work output and other outcome variables. Then, we cross-randomized two sets of treatments to increase sleep: (a) the Night-Sleep Treatments, designed to increase night sleep among individuals in their home environments, and (b) the Nap Treatment, designed to increase daytime sleep by offering individuals the opportunity to nap at the workplace.

Night-Sleep Treatments. At the end of day eight of the study, each participant was randomly assigned to one of three experimental groups. Randomization was stratified by baseline productivity and sleep to ensure balance on these key variables. Participants had equal probability of being assigned to each group.

1. *Sleep Devices:* Participants in this group were provided with a bundled intervention to increase their night sleep. Surveyors offered individuals: (i) loaned devices to improve their sleep environment, (ii) information regarding the benefits of sleep (in particular, health benefits), and (iii) encouragement to increase their sleep as well as daily feedback on the total duration of their previous night's sleep as measured by the actigraph. The offered devices included eye shades, earplugs, a cot, a mattress, sheets, pillows, and a fan (see Figure IVa).[10] Participants were asked to return the devices at the end of the study; virtually all complied.

2. *Sleep Devices + Incentives:* Participants in this group received the same interventions as the Sleep Devices Group *plus* financial incentives to increase their actigraph-measured sleep during the treatment period relative to their baseline-period sleep. Each day, the participants were paid Rs. 1 per minute of increased sleep for up to two hours of increased sleep (Rs. 120, about \$1.70). They were not penalized if they slept less than the baseline amount.[11]

   To control for any income effects due to the sleep-incentive payments, participants in the other experimental groups were randomly and anonymously matched to participants in the

[10] Participants were permitted to take more than one of each device, as piloting had suggested that the devices were often shared with family members.

[11] One concern is that participants could game the incentives by strategically reducing their baseline sleep. Of course, participants do not know their treatment status, or even the details of the different treatments, during the baseline period. Nor do control group participants increase their sleep after treatment assignment, as would be expected if they were artificially suppressing their sleep in the baseline period. Finally, as described in Section 2.3, baseline sleep is very similar to levels seen in the broader Sleep Survey.

Sleep Devices + Incentives group and received the same stream of payments, independent of their own sleep.

3. *Control:* Participants in the control group did not receive any of the treatments discussed above. To deal with the concern that loaning items might generate reciprocity effects or impact reported well-being directly, we offered placebo household goods, unrelated to sleep. The total value of these goods was roughly the same as that of the sleep devices and included items such as small kitchen devices, a chair, decorative figurines, a shoulder bag, kitchen utensils, and a flashlight. These goods were also returned at the end of the study.[12]

Given the difficulty of increasing sleep in the field—at least compared to reducing sleep experimentally in lab settings—our treatments took a bundled approach that may increase sleep through multiple channels. While the loaned devices could plausibly improve both sleep quality and the utility of spending time in bed (e.g. due to a more comfortable mattress), we expected the information, encouragement, and financial incentives to operate solely through increased time in bed.

Nap Treatment. In addition to the treatments to increase night sleep, we cross-randomized study participants into a nap treatment. Starting on day 9 of the study, a random subset of individuals were given the opportunity to take a short afternoon nap every day between 1:30 pm and 2 pm at their workplace. The nap treatment was motivated by existing evidence that naps can be effective in boosting cognitive function, including among sleep-deprived samples (Lovato and Lack, 2010).

The private nap space was located in a quiet and gender-separated area of the study office and included a bed, blanket, pillow, table fan, ear plugs, and eye shades as depicted in Figure IVb. While roughly 90% of individuals in this condition did indeed sleep during their allotted nap time, study participants who did not want to nap were asked to sit quietly or rest in their nap area. The nap group did not have the option to work during this time.

Those participants who were not assigned to take a nap were instead randomized each day with equal probability to either (a) a work condition, in which we allowed them to work through the 'nap period', or (b) a break condition, in which we enforced a half-hour break from data entry. Break participants were allowed to engage in any leisure activity they chose, including sitting in a comfortable break room in the office. By comparing the nap and break conditions, we isolate the effect of a nap relative to a break of the same length. By instead comparing the nap and work conditions, we can estimate the net effect of naps on work output, including the lost work time.[13]

[12] We found no detectable difference in treatment effects when conditioning on the subset of the control group offered the placebo goods. However, we should caution that the control group participants were not randomly selected to receive the placebo item. Instead, we decided to incorporate the placebo items halfway through the experiment, and all control group participants after that were offered them.

[13] Of course, nap and break participants may adjust their hours at the office in response to the lost time. Since we precisely observe work hours, we can estimate such labor supply effects.

3.2 Modeling the Impacts of Sleep Quantity and Quality

In Appendix C, we present a model to clarify our argument regarding how sleep relates to productivity, time use and earnings. The model is a standard labor supply/time allocation model augmented to incorporate the decision to sleep, in the spirit of Biddle and Hamermesh (1990). Extending their work, we incorporate sleep quality in addition to sleep quantity. Our interventions are modeled as potentially altering both the incentive to sleep – for example, through encouragement to sleep – and the quality of sleep – for example, through the provision of sleep devices or high-quality nap environments. The model predicts how changes in either the incentive to sleep or the quality of sleep may impact productivity and time allocation.

The model provides two key insights. First, increasing the incentive to sleep should (weakly) increase productivity, but its effect on time spent working and on earnings is ambiguous. The ambiguity results from the fact that sleep quantity affects hours working through two opposing forces. On the one hand, as sleep increases productivity, it also increases the incentive to work, as a standard price effect in labor supply models. On the other hand, increased sleep crowds out time available for leisure, increasing the shadow price of work. Thus, unless the impact of increased sleep on productivity is large, labor supply will decrease, with an ambiguous overall impact on earnings.

Second, increasing sleep quality while holding quantity fixed should increase labor supply and earnings. When sleep quality increases, it reduces the time cost of sleep by translating time in bed more efficiently into time asleep. This efficiency gain frees up time for all other activities, decreasing the shadow cost of work. Moreover, sleep quality may increase work productivity directly, holding sleep quantity fixed. Both effects work in the direction of increasing productivity, hours working, and earnings. [14]

3.3 Overview of Timing and Outcome Measures

Study participants were primarily engaged in data-entry work. Compensation consisted of a mixture of per-minute wage and piece-rate payments per correct character entry (minus a penalty for incorrect characters), as described in detail in Section 5.1. On most days, participants could choose their work hours freely: they were not penalized if they systematically showed up late or left work early. We observe the timing and content of each keystroke entered via the data entry platform, providing unusually fine measures of labor supply and work output.

Sleep was measured throughout the study via actigraphy. Participants received a modest daily incentive of Rs. 10 to wear actigraphs, which they forfeited if they removed the actigraph at any point during the day. Compliance was high in all experimental arms, with approximately 6% of participants removing the device on any given day.

[14] Increasing sleep quality has ambiguous effects on sleep quantity. First, more efficient sleep has a lower time cost, creating an incentive to sleep more. Second, depending on whether sleep quality and quantity are substitutes or complements in work productivity, increasing quality pushes quantity to decrease or increase. Finally, sleep quantity is also influenced by how hours worked react to sleep quality. It is theoretically possible that an increase in sleep quality could increase sleep so much that it actually crowds out work and reduces earnings. The conditions for this are unlikely to be met in our setting, as we discuss in Appendix C.

In addition to sleep and data-entry work outcomes, study participants completed a series of short surveys and experimental tasks—such as incentivized measures of time, risk and social preferences—throughout the study, as detailed in Table A.XIX. Participants also completed a baseline survey on Day 1 and an endline survey on Day 28, where demographics and health outcomes were measured. Finally, we measured participants' willingness to pay for the sleep devices at endline. Additional details of each task are provided in the relevant sections.

3.4 Study Population, Recruitment, and Balance

3.4.1 Recruitment and Selection

Recruitment followed two strategies: First, recruiters went to low-income neighborhoods in Chennai and spread information about the study. Advertisements for the study offered a one-month data-entry job, and recruiters provided interested individuals with additional information. Second, past participants could refer potential new participants to the study. In both cases, recruiters approached interested individuals to interview them and determine their eligibility to participate in the study.

Screening. Interested individuals participated in a two-stage screening process, involving an unpaid survey and a home visit to check whether the individual met study's eligibility criteria: (i) being between 25 and 55 years of age; (ii) fluency in Tamil (the local language); (iii) the ability to read and write numbers; (iv) having worked fewer than 5 days per week in the previous month; (v) earning less than Rs. 700 ($10) per day in the previous month; (vi) living in a dwelling able to accommodate the sleep devices used in night sleep treatments; (vii) ownership of three or fewer of the sleep devices being offered in the study; (viii) the intention to be in Chennai for the following 5 weeks; and (ix) no children in the household younger than 3 years.

Informed Consent. All participants who passed the screening process were then asked to engage in an informed-consent process. During this process, we explained to participants that if they chose to enroll, they would also be participating in a study and would be randomized among treatment conditions and asked to engage in a number of study-related activities. Participants were told that the research team was studying poverty and economics in Chennai, and the goal of the research was "to understand what difficulties underprivileged people in India face, and how these problems affect their lives." The consent form did not explicitly link the research to questions regarding sleep, although it did note that sleep would be measured and that participants may be randomly assigned to receive sleep aids.

Selection. At each recruitment and screening stage, the majority of individuals were able and willing to proceed to the subsequent stage. First, 62% of individuals surveyors approached on the street agreed to take the eligibility survey. Second, 57% of these individuals passed the initial screening test for eligibility in the study. Third, 72% of such individuals passed the home screening, and 95% passed the final office screening and completed informed consent.

How might these selection criteria affect our sample? Most importantly, they do not seem to select participants on average levels of sleep quantity and quality given that we find very similar patterns of sleep among individuals in Chennai in the broader Sleep Survey, as described in detail in Appendix F. Of course, it could be that treatment effects are different in a broader sample. Homeless individuals, for instance, would not be able to accept the sleep devices such as mattresses, and parents of young children may be unable to avoid awakenings at night.

3.4.2 Sample Characteristics and Balance Checks

Tables A.II and A.III show sample characteristics and balance checks for the experiment. Sixty-six percent of study participants were female, and the typical study participant was roughly 35 years old with 1 to 2 children . Study participants had about 10 years of education on average. While only about 30% of participants had prior experience with computers, participants were eager to learn and their performance on the data entry work improved quickly during the baseline period.

We test for imbalances in baseline characteristics across the experimental conditions using the following specification:

$$y_i^B = \beta_1 T_i^{\mathrm{D}} + \beta_2 T_i^{\mathrm{I}} + \beta_3 T_i^{Nap} + \varepsilon_i, \tag{1}$$

where y_i^B is a participant-level baseline variable.[15] The treatment variables T_i^{D}, T_i^{I}, and T_i^{Nap} indicate whether an individual is part of the sleep devices, the sleep devices + incentives, or the nap treatment, respectively. Given the cross-randomized design of the study, we show the results separately for the night sleep and nap treatments.

The treatment groups were well-balanced across key characteristics, indicating that the randomization was successful. Each treatment arm passed a joint F-Test, indicating that there are no systematic differences on observable characteristics across groups. As is expected given the large number of comparisons, a few statistically significant differences across treatment groups did emerge. Most notably among those, we find differences in baseline age across night sleep treatment groups and in productivity between the nap and the no nap groups. We control for both variables and for participant-level baseline average of the outcome variable, so these imbalances should not substantially affect our results.

3.5 Empirical Framework

Primary Specification. Our benchmark estimating equation is:

$$y_{itd} = \beta_1 T_i^{NS} + \beta_2 T_i^{Nap} + \beta_3 S_{it}^W + \gamma_1 \bar{y}_i^B + \gamma_2' X_{it} + \delta_t + \delta_d + \varepsilon_{itd} \tag{2}$$

where y_{itd} is the outcome variable for participant i on her t^{th} day in the study and on calendar date d. T_i^{NS} is an indicator of whether the participant was assigned to one of the two night-sleep

[15] Unless otherwise stated, y_i^B is a baseline-wide average whenever a variable is measured multiple times during baseline.

(NS) treatments. We pool these two treatments in order to boost statistical power and for ease of exposition and discussion.

The average treatment effect of the night-sleep and nap treatments is captured by β_1 and β_2, respectively. The variable S_{it}^{W} is a dummy capturing whether participant i was assigned to the Work condition on date t, such that β_2 captures the nap treatment in comparison to taking break. The excluded group varies with the specification and the relevant counterfactual but is noted in the table notes.

Following McKenzie (2012) and as pre-specified, we control for the average baseline value of the outcome variable $\bar{y}_{ib}$ in all specifications, excluding the baseline days from the analysis. When analyzing the nap treatment, we also exclude day 28, since the participants are not provided with a nap on the last day in study. We also control for various baseline covariates, including participants' age (quartiles) and sex. X_{it} also includes additional task-relevant controls specific to each outcome, as we describe below. Finally, we include day-in-study and date fixed effects, captured by δ_t and δ_d, respectively. In some specifications, we include the interaction between the night-sleep and nap treatments.

Pre-Analysis Plan (PAP). This study was pre-registered on both the AEA RCT Registry and ClinicalTrials.gov prior to endline data collection.[16] This analysis plan is followed throughout except when noted. The most substantial deviations are: 1) pooling the sleep devices and the the sleep devices + incentives treatment groups in most regressions, for ease of exposition and to increase power, 2) omitting days when participants are absent to improve power (attendance is balanced across groups), 3) assessing earnings in levels rather than with an inverse hyperbolic sine transformation (IHS) because results were similar but easier to interpret in levels (Table A.XXV), and 4) using deposits and interest earned rather than daily net savings due to a sub-optimal design feature discovered during the study (and described in the savings section). A comprehensive description of all deviations is presented in Appendix G.

Multiple Hypotheses Testing Corrections. Following the PAP, we adjust p-values for multiple hypotheses testing corrections within two families of outcomes: work-related outcomes (productivity, hours worked, earnings) and decision-making outcomes (savings, default effects, time preference, inattention, risk preferences, and social preferences). Specifically, we run a simulation to control the Family-Wise Error Rate (FWER). Adjusted p-values, which are reported in square brackets, can be found in Tables II, III, VII and A.XVIII, reflect the percentage of iterations in the simulation for which at least one test within the family of outcomes would have been (falsely) rejected at the critical value actually observed in our data. For more details, see Appendix E.

Joint Test of Significance

Another approach to inference when dealing with multiple outcome measures is to conduct joint tests of the null hypothesis that all treatment effects are equal to zero for a given treatment. We

[16] IDs AEARCTR-0002494 and NCT03322358, respectively.

conduct this test for both the pooled night-sleep treatment and the nap treatment, in the spirit of Young (2019), who suggests the use of randomization inference for such joint significance tests. We consider 10 outcomes variables for this test: productivity on the typing task, hours typing, savings, time, risk and social preferences, two measures of attention, and physical and psychological well-being. Since we are largely interested in positive effects of our treatments, we also consider specifications of the joint test in which we exclude labor supply, which is negatively affected by the night-sleep treatment. The results of these joint tests are depicted in Figure VII.

3.6 Survey of Experts

In order to benchmark our results and quantify how they compare existing scientific knowledge, we conducted surveys of experts in sleep science and economics which elicited their prior beliefs about the treatment effects of this study (DellaVigna et al., 2019). The surveys were distributed via a combination of direct emails from the principal investigators to experts in labor, development, and behavioral economics and mailing lists to experts and practitioners in sleep science.

The survey provided information on the design of the study, the magnitude of the effect of increased night sleep, and the outcome measures that were collected.[17] Three versions of the survey were tailored to different audiences: development and labor economists; behavioral economists; and sleep medicine experts. A total of 28 labor and development economists, 19 behavioral economists, and 77 sleep medicine experts responded to the survey. All experts made predictions on labor-supply and work-output effects. Only behavioral economists were asked to predict changes in present bias, while only sleep experts were asked to predict changes in sustained attention and physical health. The expert predictions are referenced when presenting results. Further details are provided in Appendix B.1.

4 Treatment Effects on Sleep

All of our treatments were highly effective at increasing study participants' sleep. On average, the two night-sleep treatments increased night sleep by 27 minutes. The median nap duration was 15 minutes.

Night-Sleep Treatments. Both the sleep devices and the devices + incentives treatments increased sleep markedly and immediately, as measured both by self-reports and objective sleep measurements using the actigraphs (Figure V and Table I). On average, individuals in the sleep devices and the sleep incentives treatment groups increased their time asleep by 20 and 33 minutes compared to the control group, respectively (Table I, Column 1).

[17] The expert surveys were conducted after over half the RCT sample had been acquired, in order to provide respondents with information on the treatment effects on sleep themselves. However, the paper had not been publicly presented or circulated with results at that time. We did not elicit predictions about the effects of the nap treatment in an effort to keep the survey short.

The increase in time asleep was almost exclusively due to additional time in bed rather than improved sleep quality. Both night-sleep treatment groups increased their time in bed significantly throughout the treatment period—31 minutes for the Sleep Devices group and 46 minutes for the Sleep Incentives group (Figure Va and Table I, Column 2). They did not significantly change sleep efficiency compared to the control group (Figure Vd and Table I, Column 3), even in the middle of the night (Appendix Figure A.IIa). Thus, while the treatments were effective at increasing sleep quantity, they entailed substantial opportunity costs of time by increasing the time spent in bed.

Similarly, the treatments were generally unable to improve the severely fragmented nature of sleep experienced by most participants (Table A.VII). There is no significant change in the number of awakenings per hour across either treatment group, and the longest sleep episode only increases by a small and marginally significant amount.

The Sleep Devices treatment was designed to improve individuals' sleep environments with the goal of increasing sleep efficiency, and 87% of treated participants took at least two of the offered sleep-aid devices home. Why then did efficiency not increase? Either the devices we provided were not sufficient to overcome the environmental challenges interfering with sleep, or other barriers to sleep (such as mosquitoes, heat, crowding, or psychological distress) were binding. One clue is that the self-reported usage of the devices was, on average, low (see Appendix Figure A.IV).

The self-reported changes in sleep are broadly consistent with the estimates based on actigraphs (Figure Vc). Participants in the sleep treatments report increasing their time in bed by 1 to 1.2 hours and time asleep by 0.9 to 1.1 hours. This over-estimation of treatment effects relative to the actigraph data suggests either difficulties in reporting sleep accurately—possibly due to confusion between time in bed and time asleep—or potential experimenter demand effects, both of which highlight the importance of objective sleep measurement. As in the actigraph-based data, the changes in time in bed and time asleep did not result in a significant change in sleep efficiency.[18]

Nap Treatment. The nap intervention was effective at increasing study participants' daytime sleep. The vast majority of participants in the Nap Group (92%) self-reported falling asleep during a nap. These reports are confirmed by actigraph data which recorded that 88% of participants were able to fall asleep during their nap. The median time asleep during the nap window was 15 minutes, and the (unconditional) mean time asleep was 13 minutes.

The quality of the naps provided in the study office appears to be significantly higher than the quality of night sleep. For instance, the average number of awakenings per hour is lower for naps. Moreover, sleep efficiency during naps in the office (85%) is higher than efficiency in night sleep (70%) and in naps at home (72%, similar to night sleep).[19] The higher quality of nap sleep is consistent with two important factors affecting sleep quality. First, naps in our treatment group take place in our office, which is devoid of many of the environmental factors (excessive noise, heat,

[18] At roughly 90% self-reported sleep efficiency is significantly higher than measured sleep efficiency.

[19] Sleep efficiency is tricky to define for naps in a way comparable to night sleep. We calculate time in bed for naps as beginning with the minute the participant is first detected to fall asleep, and ending with the end of the nap. The same definition is applied for naps at the office and at home.

mosquitoes, crowded conditions) that drive poor sleep in participants' typical sleep environments. Second, the naps in our study are timed to coincide with the circadian dip in the mid afternoon, when individuals are prone to sleepiness and impaired performance. A short burst of sleep during the circadian dip has been shown to be particularly valuable (Takahashi, 2003).

Interactions and Heterogeneity. Interaction effects between the night-sleep and nap treatments were modest. Participants randomized to the night-sleep treatments did not nap any less when offered a nap. Those treated with naps did sleep five minutes less at night, implying a partial crowd-out of night sleep by naps (Table I, Columns 1 and 2). Both treatments increased total time asleep in 24 hours. Finally, the impacts of both treatments on sleep quantity and quality were largely homogeneous on baseline sleep quantity and quality, as well as other covariates such as sex, age, and earnings (Table A.IV).

5 Labor Outcomes

5.1 Design and Outcomes

Data-Entry Work. Study participants were employed full-time as data-entry workers throughout the study. The data-entry task consisted of digitizing text and numeric data designed to mimic a real-world, data-entry job, using the interface represented in Figure A.V.[20] Fields had to be entered sequentially.[21] Participants spent 63% of their work time at the office doing data-entry work, which constituted 57% of their earnings in the study. The rest of the time was spent answering surveys or completing other experimental tasks.

Work Hours. The participants were generally free to choose their work hours, including their arrival, departure, and break times (except those associated with the nap). Most days ("regular days"), participants could arrive and depart as they choose between 9:30 am and 8 pm. On a subset of days ("short days"), work hours were limited to 11 am to 5 pm, in an effort to provide clean estimates of productivity, unconfounded by potential changes in labor supply.[22]

Data-Entry Incentives. The payment for the data-entry work had two components. First, the participants received Rs. 21.60 or roughly $0.30 per hour of active typing. If a participant spent two

[20]The data to be digitized had been artificially generated. By generating the data, we had ready access to the correct data, allowing us to measure the accuracy of the work immediately. Study participants were unaware of the artificial nature of the data. Since accuracy is often measured in "real" data-entry jobs via double-entry they had no reason to doubt that their work was not "real" and useful.

[21]To facilitate learning, during the first three days of participation in the study, a pop-up box appeared after each page was submitted. This box informed participants of their earnings and the number of mistakes they had made on that page. After dismissing the message, the software provided a new batch of data for entry. Starting on day four, this information was no longer provided.

[22]To encourage presence during these hours, we paid a bonus of Rs. 50 to anyone present during the entire period. Since there is little difference in labor supply across night-sleep treatment arms, we do not separate the analysis of work outcomes by short vs. long days.

consecutive minutes without typing, the software would automatically pause and the participant did not receive the time-based payment until they began typing again. Second, participants received a performance-based payment consisting of a piece rate per correct character and a penalty per mistake.[23] Each half hour, piece-rates were randomly assigned between a high value (Rs. 20 per 1,000 correct characters) and a low value (Rs. 5 per 1,000 correct characters) with equal probability. The penalty rate remained constant throughout at Rs. 1 per 10 mistakes. The data-entry screen informed workers about their current piece-rate as described in Section 6.1. Payments were made daily before the participant left the study office.

5.1.1 Work Outcomes

We focus on three key performance measures:

1. **Labor Supply**. We use two measures of labor supply: (a) time at the office, measured as the difference between their departure and arrival times at the office, and (b) active typing time, captured via the data-entry software by adding all periods of active typing within the day.

2. **Earnings**. Data-entry earnings are a combination of attendance pay and performance earnings as described above. On average, 65% of the typing compensation was performance pay.

3. **Productivity**. Our primary measure of worker productivity is output divided by active typing hours. Output is the number of correct entries minus 8 times the number of mistakes.[24]

Since the treatments had no impact on labor supply at the extensive margin (days worked, see Figure A.VI), we focus on labor supply at the intensive margin (hours worked), as well as productivity and earnings conditional on visiting the office.

5.1.2 Expert Survey Predictions

The expert survey predictions show a clear pattern (Table A.I): most experts in both economics and sleep science predicted that a 32 minute increase in night-time sleep would have strong positive impacts on labor market outcomes. The median expert predicted that output would increase by 7% and hours typing would increase by 4%.[25]

5.2 Impact of Increased Night Sleep on Work Outcomes

The night-sleep treatment increased productivity by 1.3% (s.e. 1.1) in comparison to the control group (Table II, column 1), a small increase compared to expert predictions. The night-sleep treatments reduced labor supply on the intensive margin by approximately 10 minutes per day (column

[23] Following Augenblick et al. (2015), we measured mistakes using the distance between the text entered and the text prompted as measured by the Levenshtein distance, defined as the minimum number of single-character edits (insertions, deletions, or substitutions) required to change one word into the other.

[24] The relative weight was derived from the ratio of the average piece-rate (Rs. 1.25) and error rate (Rs. 10) per 100 characters.

[25] More precisely, experts were asked to make predictions on the number of correct entries. The correlation between correct entries and the pre-registered measure of output, described in Section 3, is 0.99.

4). Through the lens of our model in Appendix C, this implies that work and sleep are substitutes in equilibrium in our setting.The small impact of increased night sleep on productivity is not enough to outweigh the reduction in labor supply, leading to a small and non-significant decrease in earnings (column 7).

The lack of positive impacts on earnings or output is in sharp contrast to the expert predictions of output increases of 7%. This discrepancy can in part be explained by experts over-estimating the productivity impacts of increased sleep, and in part by a prediction that sleep quantity and work hours would be complements, as most economists and sleep medicine experts predicted that increased night sleep would raise labor supply by at least 4%. Instead, some of the increased time in bed caused by the night-sleep treatments translated into reduced time at the office and thus a 4% reduction in typing time (see Table A.VI for a more detailed decomposition of labor supply effects).

Why do expert predictions diverge from our results? We hypothesize that low sleep quality in our setting explains this result. First, sleep quality may directly affect productivity, encouraging substitution from work to leisure at low levels of quality. Second, individuals with higher sleep quality—as measured by sleep efficiency—need to spend less time in bed to increase their sleep, implying a smaller opportunity cost. When sleep quality is low, the opportunity cost is high and the substitution effect is in favor of leisure, leading to net reductions in labor supply and output.

To test this hypothesis, we interact the night-sleep treatment with baseline sleep quality.[26] For participants with high baseline sleep quality, the night-sleep treatment increased productivity by 3.7% on average (Table II, column 2). In comparison, participants with low baseline sleep quality had a negative treatment effect of about 1% (not significant). Both groups saw similar reductions in labor supply of 16 and 18 minutes respectively (column 5). Putting these effects on productivity and labor supply together, the night-sleep treatment had no effect of earnings for participants with high baseline sleep quality, despite reducing their work time. But it reduced the earnings by 5.4% for participants with low sleep quality (column 8). Taken together, these results point to the importance of sleep quality for work productivity.

The night-sleep treatment was also more effective for participants with lower baseline sleep quantity, consistent with diminishing returns to sleep (Table II, columns 3, 6, and 9). Participants with below-median sleep quantity at baseline increased their productivity by 3% and reduced labor supply by only 5 minutes per day in response to treatment. In contrast, those with above-median sleep quantity had no effect on productivity and reduced labor supply by 16 minutes per night. Taken together, the earnings losses are much larger for participants with above-median baseline sleep quantity (earnings decrease of 7.6%) compared to those with below-median sleep quantity (non-significant increase of 1.7%).

The differential earnings effects of treatment on those with high quality and quantity of sleep are not driven by a differential impact of treatment on sleep itself, as discussed in Section 4. Our

[26] We construct the quality variable as follows. First, we standardize three measures of sleep quality – sleep efficiency, longest sleep episode, and number of awakenings per hour – by their mean and standard deviation, changing the sign where needed so positive values indicate higher sleep quality. We then define the quality index as the simple average of these three standardized variables. Section 2.3 contains the details about the measures of sleep quality and Tables A.IX-A.XI present a breakdown of Table II into each component of the quality index.

results instead are consistent both with diminishing marginal benefits to sleep quantity and of complementarity between sleep quality and quantity in terms of work productivity.

5.3 Impact of Naps on Work Outcomes

In contrast to the average effects of night sleep, naps increased productivity significantly. Despite spending an average of only 13 minutes asleep, participants who were randomized to naps were 2.3% more productive on average across the day (Table III, column 1) with similar impacts relative to both the break and work counterfactual.[27] These impacts are sizable, given that the work task itself is relatively inelastic – quadrupling the piece rate increases productivity by only 10%.

By design, nap participants had a 30-minute block during which they could not work. Individuals were free to adjust their labor supply outside of this period, but we find no evidence of significant adjustments (columns 3 and 4). The labor supply of the nap group is only two minutes greater than that of the break group and 26 minutes less than the group that was eligible to work through that period due to a slight increase in labor at the end of the day among nap participants.

Given the substantial differences in labor supply, the impact of naps on earnings depends on the comparison group. Compared to taking a break, naps increased overall earnings by about Rs. 11 per day, a sizable increase of 4.1% (column 6). However, the opportunity costs of taking time to nap lowered earnings by Rs. 23 (8.3%) for the nap group when compared to working through the break.

Importantly, this negative effect on earnings appears to diminish over time. During "regular" work days (Figure A.VIIa), when participants are not restricted to artificially short work hours, there is no detectable difference in earnings between the nap group and the work group by the end of the study. This effect is partly driven by the nap group working longer hours over time (Figure A.VIIc). Nap participants—who are mechanically prevented from typing for 30 minutes in the afternoon—typed roughly 40 minutes per day less than the work group in the first week of treatment, but only 15 minutes less by the end of the study. They stay in the office 10 minutes per day more than break participants by the end of the study. They also appear to become more productive over time, although these estimates are noisier (Figure A.VIIb).

5.4 Summary of Work Outcomes

Taken together, the productivity benefits of increasing night sleep are low in our setting, despite the low baseline levels of sleep. A 30-minute increase in night sleep did not produce significant gains in productivity and even slightly reduced earnings via reduced labor supply. Yet, it is not the case that sleep does not matter for productivity. Additional sleep *did* produce large productivity gains when sleep quality was high – both during the night and in the form of afternoon naps.

[27] Productivity effects of the naps are more strongly concentrated in the afternoon period following the nap, however, point estimates are positive (but insignificant) in the morning suggesting potential anticipatory effects of the nap.

6 Cognition and Well-being

We now turn to the effects of our treatments on cognitive function, psychological well-being and physical health. These outcomes are of interest in themselves, and may be potential mechanism for impacts on work outcomes discussed in the previous section.

6.1 Sleep and Limited Attention

Sleep scientists have documented a strong relationship between sleep and attention in laboratory studies in rich countries (Lim and Dinges, 2010; Killgore, 2010). Because attention is a critical cognitive resource, it is likely to be direct contributor to productivity. Some scholars argue that limited attention may underlie a number of behavioral biases and anomalies (Gabaix, 2019). We test whether increased sleep improves a measure of attention used in the sleep literature, and whether it improves participants' ability to attend to important aspects of their work environment, serving as a potential channel through which productivity effects may operate.

Lab Measure of Simple Attention. Each day in the office, participants completed the Psychomotor Vigilance Task (PVT), a standard measure of alertness and attention used in sleep medicine (Basner et al., 2011; Basner and Dinges, 2011). Participants are asked to react to a series of visual stimuli shown on a computer screen over ten minutes by pressing a key as soon as they see a stimulus appear on the screen. The test measures the speed and accuracy with which subjects respond to the visual stimuli on the screen and has been shown to be highly responsive to experimentally-induced sleep deprivation (Dinges et al., 1997). The task was incentivized with payments reflecting both speed and accuracy.

Results. Similar to the observed labor market responses and in contrast to expert survey predictions, the night-sleep treatments had no significant effect on PVT performance. In contrast, naps increased payment received for the task, the pre-registered outcome, by 0.169 standard deviations ($p<0.01$, Table IV, Panel A, column 3), working through both reduced lapses in attention and improved reaction time (Table IV, Panel A, columns 4 and 5). We find similar effects for other indices constructed with standardized measures of the individual components. Motivated by this finding, we next examine whether these impacts on attention translate to the participant's ability to attend to important aspects of their work environment.

Attention in the Work Environment. To test whether sleep impacts the ability to react to pay attention to shrouded incentives in one's work environment, we randomized the visual salience of piece rates (performance pay) across days within individuals starting on day 6 of the baseline period.

- In the *salient condition*, the current piece-rate was highlighted and readily available to study participants at all times. A low piece rate was highlighted in blue at the bottom of the screen (panel A.Vb), while a high piece rate was highlighted in green (panel A.Vc). In addition, the screen blinked twice to indicate the beginning of a new 30-minute slot, thereby drawing attention to a possible switch in incentives. We consider this condition the "full-attention" benchmark, as in Chetty et al. (2009).

- In contrast, in the *non-salient condition*, noticing and remembering the piece-rate was more challenging. First, the bottom of the screen was uncolored for both piece-rates in the non-salient condition. In addition, the piece-rate was only visible for the first 15 seconds of a slot (as illustrated by panels A.Ve and A.Vf) and faded in and out slowly.

Empirical Strategy. Our empirical strategy to identify attention to wage incentives follows Chetty et al. (2009). We estimate the ratio of the productivity reaction to the increased incentives in the non-salient periods to the reaction in the salient periods.[28] We interpret this ratio as the deviation from the "full-attention benchmark" caused by inattention to non-salient incentives. The closer the ratio is to 1, the more participants attend to the shrouded incentives.

Results. Consistent with limited attention, work output of the control group reacted 16% *less* to high incentives when piece-rates were *non*-salient (Table IV, Panel B, column 1). Participants in the night-sleep treatments behave quite similarly, reacting 15% less to incentives when they were non-salient. These results are consistent with the lack of effect of the night-sleep treatment on PVT performance.

In contrast, and consistent with the PVT results, the Nap Group was nearly fully attentive to non-salient incentives. Specifically, we cannot reject $\theta = 1$ for the Nap Group, who was 10 percentage points closer to the "full-attention benchmark" than the Control Group, highlighting the improved attentional resources provided by naps in a real-world work environment. Consistent with evidence from the sleep literature that short afternoon naps boost attention (Lovato and Lack, 2010), the increases in attentiveness almost entirely arose in the afternoon (i.e. post nap time).[29]

Summary. Using both a standard lab task used by sleep scientists and a novel measure of attention to work incentives, we find consistent evidence that naps increase participants' attention. These effects may serve as one of the underlying drivers of the productivity gains in the Nap Group. In contrast, increases in night sleep have no detectable effect on attention, either as measured via the laboratory task form the sleep literature on the work environment.

Other Aspects of Cognition. The cognitive benefits of naps appear to be concentrated on attention rather than effect other aspects of cognition.[30] In contrast to the existing literature regarding sleep deprivation from laboratory experiments (Lim and Dinges, 2010; Killgore, 2010), we find no evidence of impacts of any of the treatments on memory and inhibitory control, the ability to override impulses (Table A.XIII).[31]

[28] A detailed description of this estimation is provided in Appendix B.2.
[29] Results are similar when examining labor supply changes driven by breaks in response to the salience of incentives.
[30] For a more detailed overview of these tasks, see Figure A.X and Dean et al. (2019).
[31] We find suggestive evidence that naps may have slightly reduced reaction times in the inhibitory control task

6.2 Well-Being

A substantial literature in sleep science argues for the centrality of sleep to health and well-being (Cappuccio et al., 2010; Liu, 2013; Banks and Dinges, 2007; Strine and Chapman, 2005). Such effects may produce direct utility benefits of sleep and are a potentially important component of the overall welfare effects of additional sleep.

6.2.1 Psychological Well-Being

Design. We collect data on participants' depression, happiness, sense of life possibilities (Cantril Scale), life satisfaction, and stress, as described in more detail in the appendix. As pre-registered, we examine these variables both as indices and individually.

Results. Mirroring the general pattern of results described above, naps improved psychological well-being, while increased night sleep did not (Table V). The nap treatment increased the index of psychological well-being index by 0.08 to 0.13 standard deviations. While the estimates for all individual components of well-being are positive, napping appears to have the strongest treatment effects on happiness, life satisfaction, and life possibilities. As a point of comparison, the increase in life satisfaction among individuals in the nap treatment is roughly one-third to one-half the size of the gap in life satisfaction between individuals earning $10,000 per year and those earning $20,000 per year (Kahneman and Deaton, 2010). In contrast to the naps and the existing literature, the night-sleep treatments did not have a positive (and possibly even a slightly negative) impact on any measure of subjective well-being or on the overall index (Strine and Chapman, 2005; Kahneman and Krueger, 2006).

6.2.2 Physical Well-Being

Design. We collected a variety of measures of physical activity and physical health. Our pre-registered measures are as follows: (i) performance in a stationary biking task; (ii) reported days of illness; (iii) self-reported pain; (iv) activities of daily living; and (v) blood pressure. As above, we examine these variables both as indices and independently. We also measured daily steps via the actigraph, which we include as an additional measure of physical activity.

Results. Neither the night-sleep treatments nor the nap treatment led to increases in physical well-being (Table VI, columns 1 to 3). We find positive, not statistically significant point estimates of increased night sleep on several of the outcomes. These impacts are offset, however, by the negative impact of night sleep on total daily steps, which appears to be driven by a simple incapacitation effect: time spent in bed reduces time available to walk.[32] These results are perhaps not surprising given

(0.14 SD). However, this change was not large enough to impact overall payments for performance on the task.

[32] If we consider just steps during the hours when all participants would be expected to be awake, the night-sleep treatment groups accrue more steps during this time. The null result of the nap treatment on the biking task may be at least partially explained by the fact that this task was administered at endline, on a day when participants in the

that many health outcomes are likely to be slow-moving. It is possible that greater improvements would be observed in a longer study.

7 Preferences and Decision-making

Sleep may impact preferences and decision-making. In order to examine such effects we study time preference via changes in savings rates and choices on a real-effort task, as well as risk and social preferences via standard measures. Finally, we elicit willingness to pay for the sleep devices provided in the experiment as a summary measure of demand for additional sleep.

7.1 Time Preferences: Savings Task

Design. We measured savings behavior by offering individuals the opportunity to save money in a locked box at the study office, as in Schilbach (2019). This allows us to observe a part of participants' savings behavior precisely, rather than relying on often noisy self-reported measures. At the end of each work day after receiving their earnings, individuals had the opportunity to deposit or withdraw money from their savings box. Participants were randomly assigned to receive *daily* interest rates between of 0 and 2% for any money saved in the box.[33] For participants receiving the positive interest rate, at least, the savings account we offered was quite attractive, even if they had other lower-interest savings opportunities outside the office. Deposits were capped at Rs. 600 per day in order to ensure that participants did not make large deposits from other sources to leverage the high interest rates.[34]

Results. The night-sleep treatments did not meaningfully affect savings behavior (Table VII, Panel A). In contrast, the nap treatment is associated with economically meaningful increases in daily deposits and daily net savings (deposits minus withdrawals, columns 1 and 2 of Table VII, Panel A). The Nap Group deposited an additional Rs. 16 per day, a 15% increase relative to the Control Group. The impact of naps on daily net savings is similar in relative terms (13.5%), but less precisely estimated, which is at least in part explained by large, one-time withdrawals before the end of the study (Figure A.VIII).[35]

We consider two versions of a third measure savings. First, we consider the interest accrued over the course of the study, excluding participants assigned to zero interest rate. Nap participants

nap treatment did not have the opportunity to nap.

[33]The interest rates changed during the study two times. Details are provided in Appendix G.

[34]The deposit ceiling was Rs. 400 for roughly the first 4 months of the study. Because participants were frequently reaching this cap, we raised the limit to Rs. 600.

[35]We pre-registered daily net savings as our main variable of interest. However, this variable is not an ideal measure of savings behavior since large withdrawal before participants' end of their participation in the study add noise to this measure (Figure A.VIII). Unlike in Schilbach (2019), our study features a daily interest rate rather than a matching contribution at the end of the study, thus making large withdrawals less costly and frequent. Therefore, deposits and especially interest accrued in the study are preferable outcome measure since they are less noisy and capture savings behavior more comprehensively. Because these very large withdrawals are an artificial feature of our study design, we believe deposits more accurately reflect differences in savings behavior.

enjoy a 21.4% increase in interest accrued, indicating higher average savings balances throughout the study (column 3). Although the magnitude of this estimate is economically meaningful, it is under-powered (and only marginally significant) in part because participants assigned to zero interest rates by definition earn no interest. To account for this, we create a variable that captures the counterfactual interest accrued assuming all participants were assigned a 1% daily interest rate, absent behavioral responses to this interest-rate change (see Section B.3 for details). Naps increased the counterfactual interests accrued by 17.7%, significant at the 10% level (column 4).

These effects are not driven by mechanical wealth effects. On average, participants in the nap group earned *less* than the relevant comparison group (i.e. a 50/50 mix of working and taking a break). Similarly, changes in risk preferences are unlikely to explain these effects, as naps have no detectable impact on risk preferences (Section 7.3). Another possibility is that sleep improves either attention to the future or self-regulatory capabilities (Christian and Ellis, 2011; Barber and Munz, 2011), decreasing the extent to which self-control issues reduce participants savings. Consistent with this hypothesis, we find that naps reduced present bias in a real-effort task, as discussed next.[36]

7.2 Time Preference: Effort Discounting

Design. Following Augenblick and Rabin (2019a) and Augenblick et al. (2015), we measure present bias using real-effort choices. Participants made decisions about how many pages to type in the evenings on a fixed date (referred to as "work day") under different piece rates. Differences in committed choices for a future work day vs. the updated choices made when that day actually arrived identify present bias. The task mimics regular data-entry work in the study, using shorter pages to allow for a finer choice set. Using effort choices elicited for the future and present both at baseline and at least once during the treatment period for each participant, we structurally estimate individual-level present bias parameter β_i for the baseline and for the treatment period. A complete description of the task is in Appendix B.6.

Results. In preferred specification, we estimate a mean present-bias parameter β of 0.92 in the control group (Table VII, Panel B, column 1).[37] Reassuringly, the estimated β is predictive of other behaviors conceptually related to time preference. For example, participants with a lower estimated β arrive at work later and save less (Table A.XVI). Similar to the other outcomes described above, the night-sleep treatments did not significantly affect the present-bias parameter. In contrast, the nap treatment increased the estimated β by 0.06 (column 1).[38] For robustness we also consider the ratio between pages chosen for work "now" versus "later" as dependent variable (columns 3 and 4).

[36] We also used the savings task to study whether the sleep treatments reduced participants' propensity to be subject to 'default effects' in savings decisions. This measure ended up being under-powered, with correspondingly imprecise estimates. We therefore relegate its detailed discussion to the Appendix in Section B.4.

[37] Our estimate is slightly higher than the average present bias found in papers using a similar methods. Augenblick et al. (2015) for example find an average effort discounting of 0.89, while Augenblick and Rabin (2019) find an average β ranging from 0.79 to 0.84.

[38] For participants who did an updated version of the present bias experiment, results are even larger (column 2), although they are not significant for this smaller sample. See Appendix B.6 for details.

The results are similar to those with the structural parameter, although less precise.[39]

7.3 Social and Risk Preferences

Design. We measure risk and social preferences via standard tasks in the behavioral economics literature. Risk preferences and loss aversion are captured via a multiple price list elicitations similar to those in Holt and Laury (2002), Sprenger (2015), and Charness et al. (2013). Social preferences are measured via dictator, ultimatum, and trust games (Camerer, 2003).

Results. Neither the night-sleep treatments nor the nap treatment significantly altered risk aversion or loss aversion in the standardized tasks (Table A.XVII), in contrast to findings of McKenna et al. (2007). While the results are not precise enough to detect very small effects, we are able to rule out changes greater than 0.2 standard deviations for each of these outcomes and treatments, suggesting that if effects are present, they would be small. We find similarly precise and null results when examining behavior in the dictator, ultimatum, and trust games, in both sender and receiver positions, where applicable (Table A.XVII). These results are also in contrast with earlier work, though we study different populations with different treatment and baseline sleep conditions (Dickinson and McElroy, 2017; Anderson and Dickinson, 2010).

7.4 Willingness to Pay for Sleep Devices

Design. At the conclusion of the study, we elicited participants' willingness to pay for a subset of the devices provided in the night-sleep treatments using an incentive-compatible BDM mechanism (Becker et al., 1964). The valuation captures both any direct hedonic effects of the devices as well as any expected benefits of additional sleep. To ensure that participants were not liquidity-constrained in these purchase decisions, their bonus payments (e.g. for wearing the actigraph) accrued throughout the study were paid out on the same day.[40]

Results. Willingness to pay for these devices is, on average, relatively low. The average participant is willing to pay roughly one-third of the market value of the devices. In addition, exposure to these goods, either via the night-sleep treatment or the nap treatment, does not impact willingness to pay for them. These results are broadly consistent with the limited impacts of additional night sleep described above and the fact that access to the devices does not result in improved sleep quality.[41]

[39] In our primary specification, consistent with the structural estimate, we estimate the model among participants for whom the structural estimator converges. This approach is commonly used in papers relying on this estimation strategy such as (Augenblick and Rabin, 2019b). We also estimate the model including the additional 46 participants for whom the estimator does not converge, resulting in an estimate roughly only half as large (columns 5 and 6). We believe this difference is caused by the high frequency of corner and non-monotonic choices among these 46 participants, which impedes our ability to estimate present bias.

[40] We also elicited participants' beliefs regarding the impact of the sleep devices treatment on attention, earnings, labor supply, and savings. However, the responses gathered – which included a high proportion of inconsistent responses and unreasonable values – suggested that participant understanding of the task was quite low, so the results are not reported here.

[41] Low willingness to pay could also be consistent with beliefs that the devices themselves are not productive in

8 Discussion and Conclusion

This paper provides experimental evidence on the economic consequences of increasing sleep. Using objective measurement devices, we find that the urban poor sleep little and poorly, both in our experimental sample and in a broader survey of the population of Chennai. The *average participant* sleeps only 5.6 hours per night and experiences a quality of sleep below typical sleep apnea patients in the United States. In addition to potentially reducing the restorative benefits of sleep, the poor quality of sleep also implies high opportunity costs of sleep: it takes over 8 hours in bed each night to produce 5.6 hours of highly interrupted sleep.

Our night-sleep intervention increased sleep by nearly half an hour over the course of three weeks, without any discernible impact on sleep quality. Surprisingly, we find no impact of this additional sleep on a wide range of outcomes including productivity, earnings, well-being, attention, or time preference, as summarized in Figure VIa. We cannot reject the null of no effect of the night-sleep intervention on all outcomes excluding labor supply, which falls (Figure VII). These results are in stark contrast to expert predictions from economists and sleep scientists as well as an extensive sleep-science literature that all suggest that additional sleep from a low baseline should have substantial benefits.

We provide suggestive evidence that low average sleep quality explains these surprising results. Treated participants with higher baseline sleep quality do see large increases in productivity, which offset reductions in labor supply, while those with low-sleep quality benefit less and earn less when induced to sleep more. Similarly, high-quality naps provided in the study office substantially improve a range of outcomes, as summarized by Figure VIb. Napping improves psychological well-being, decreases present bias, increases savings, and improves attention to important features of one's work environment. We strongly reject the null hypothesis that naps have no effect on all outcomes ($p < 0.001$, in Figure VII).

While these results point to the centrality of sleep quality, other factors may also play a role in explaining some part of our findings. First, naps may be effective relative to night sleep not just because of their high quality, but because their timing coincides with a pronounced circadian dip in the middle of the day. Second, it is possible that the effects of increased night sleep would manifest over a longer time horizon. However, sleep laboratory experiments find effects on cognitive function in well under three weeks (Van Dongen et al., 2003; Belenky et al., 2003; Lim and Dinges, 2010). We detect no such effects, and can rule out relatively small effects. Nor do we find any evidence of increasing treatment effects of night sleep over time. Third, unlike most experimental studies to date, we increase sleep at the margin rather than experimentally reducing sleep sharply. The effects of altering sleep in these different ways may not be symmetric or linear (Belenky et al., 2003). Finally, given that participation in our study constitutes the participants' livelihood, incentives for performance in this experiment are significantly greater than is typical in the sleep literature.

Taken together, we find no evidence that individuals in this population are under-investing in sleep by spending too little time in bed. Night sleep is low-quality, making it time-intensive

generating additional sleep.

and limiting the benefits of additional sleep. The high-quality nap environment used in the study is unlikely to be available to most participants at home. Because little evidence is available on methods to increase sleep quality in this environment, it is difficult to say whether participants under-invest in sleep quality. However, given that the available and reasonably inexpensive sleep devices did not impact quality, it is likely that much more substantial (and costly) improvements to urban environments are needed.

Our findings do not imply that sleep is not important or has no benefits even in our setting. There may be substantial long-run health or cognitive benefits to additional sleep that we do not capture. Yet, the average participant's choices are unsurprising given the observable trade-offs: the opportunity costs of (even) more time in bed are clear and immediate while any potential health benefits are amorphous and far in the future.

Similarly, our study does not imply low marginal benefits of sleep in rich countries, where sleep quality is typically much higher. Indeed, natural experiments such as Gibson and Shrader (2018) suggest substantial wage returns to sleep in the US. Relative to the massive body of laboratory evidence, more field evidence is sorely needed from different contexts, including more objective measurement of sleep in additional populations and causal evidence on the economic impacts of sleep in the field. Such efforts should also include uncovering how to increase sleep quality in cities in developing countries. Moreover, measuring sleep and testing interventions to improve sleep among children in developing countries is another exciting avenue for research, given the evidence on the learning effects of sleep deprivation (Jagnani, 2018; Minges and Redeker, 2016).

Finally, while naps are a common feature of daily life around the world, they are especially common in tropical countries (Dinges, 1992). Shifts towards western and industrialized work schedules crowd out napping in developing countries. Given the promising effects of naps in our study, understanding how these effects generalize to other work settings and other outcomes would be valuable.

References

Alessi, C., J. Schnelle, P. MacRae, J. Ouslander, N. Al-Samarrai, S. Simmons, and S. Traub, "Does Physical Activity Improve Sleep in Impaired Nursing Home Residents?," *Journal of the American Geriatrics Society*, 2005, *43* (10), 1098–1102.

Allcott, Hunt and Dmitry Taubinsky, "Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market," *American Economic Review*, 2015, *105* (8), 2501–2538.

Ancoli-Israel, Sonia, Roger Cole, Cathy Alessi, Mark Chambers, William Moorcroft, and Charles P. Pollak, "The Role of Actigraphy in the Study of Sleep and Circadian Rhythms," *Sleep*, 2003, *26* (3), 342–392.

Anderson, Clare and David L. Dickinson, "Bargaining and Trust: The Effects of 36-h Total Sleep Deprivation on Socially Interactive Decisions," *Journal of Sleep Research*, 2010, *19* (1-Part-I), 54–63.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Augenblick, Ned and Matthew Rabin, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," *Review of Economic Studies*, 2019, *86* (3), 941–975.

_ and _ , "An experiment on time preference and misprediction in unpleasant tasks," *Review of Economic Studies*, 2019, *86* (3), 941–975.

_ , Muriel Niederle, and Charles Sprenger, "Working Over Time: Dynamic Inconsistency in Real Effort Tasks," *Quarterly Journal of Economics*, 2015, *130* (3), 1067–1115.

Avery, Mallory, Osea Giuntella, and Peiran Jiao, "Why Don't We Sleep Enough? A Field Experiment among College Students," 2019.

Banks, Siobhan and David Dinges, "Behavioral and Physiological Consequences of Sleep Restriction," *Journal of Clinical Sleep Medicine*, 2007, *3* (5), 519–528.

Barber, Larissa K. and David C. Munz, "Consistent-sufficient sleep predicts improvements in self-regulatory performance and psychological strain," *Stress and Health*, 2011, *27* (4), 314–324. 00073.

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

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

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak, "Measuring Utility by a Single-response Sequential Method," *Behavioral Science*, 1964, *9* (3), 226–232.

Belenky, Gregory, Nancy Wesensten, David Thorne, Maria Thomas, Helen Sing, Daniel Redmond, Michael Russo, and Thomas Balkin, "Patterns of Performance Degradation and Restoration during Sleep Restriction and Subsequent Recovery: A Sleep Dose-Response Study," *Journal of Sleep Research*, 2003, *12*, 1–12.

Biddle, J. E. and D. S. Hamermesh, "Sleep and the Allocation of Time," *Journal of Political Economy*, 1990, *98* (5), 922–943.

Buysse, Daniel J, Charles F Reynolds, Timothy H Monk, Susan R Berman, David J Kupfer et al., "The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research," *Psychiatry res*, 1989, *28* (2), 193–213.

Camerer, Colin F., *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton, NJ: Russell Sage Foundation/Princeton University Press, 2003.

Cappuccio, F., D. Cooper, L. D'Elia, P. Strazzullo, and M. Miller, "Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies ," *European Heart Journal*, 2011, *32* (12), 1484–1492.

Cappuccio, Francesco P., L. D'Elia, P. Strazzullo, and M. Miller, "Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies," *Sleep*, 2010, *33* (5), 585–592.

Carrier, Julie, Stephanie Land, Daniel J. Buysse, David J. Kupfer, and Timothy H. Monk, "The Effects of Age and Gender on Sleep EEG Power Spectral Density in the Middle Years of Life (Ages 20–60 Years Old)," *Psychophysiology*, 2001, *38* (2), 232–242.

Castro, L., D. Poyares, D. Leger, L. Bittencourt, and S. Tufik, "Objective Prevalence of Insomnia in the Sao Paulo, Brazil Epidemiologic Sleep Study," *Annals of Neurology*, 2013, *74* (4), 537–546.

Cespedes, Elizabeth M, Frank B Hu, Susan Redline, Bernard Rosner, Carmela Alcantara, Jianwen Cai, Martica H Hall, Jose S Loredo, Yasmin Mossavar-Rahmani, Alberto R Ramos et al., "Comparison of self-reported sleep duration with actigraphy: results from the Hispanic Community Health Study/Study of Latinos Sueño Ancillary Study," *American journal of epidemiology*, 2016, *183* (6), 561–573.

Charness, Gary, Uri Gneezy, and Alex Imas, "Experimental Methods: Eliciting Risk Preferences," *Journal of Economic Behavior & Organization*, 2013, *87*, 43–51.

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

Christian, Michael S. and Aleksander P. J. Ellis, "Examining the Effects of Sleep Deprivation on Workplace Deviance: A Self-Regulatory Perspective," *Academy of Management Journal*, October 2011, *54* (5), 913–934. 00390.

Cohen, Jonathan, Keith Marzilli Ericson, David Laibson, and John Myles White, "Measuring Time Preferences," *Journal of Economic Literature*, 2019.

Cohen, Sheldon, Tom Kamarck, and Robin Mermelstein, "A global measure of perceived stress," *Journal of Health and Social Behavior*, 1983, *24* (1), 385–396.

Cole, Roger J., Daniel F. Kripke, William Gruen, Daniel J. Mullaney, and J. Christian Gillin, "Automatic Sleep/Wake Identification From Wrist Activity," *Sleep*, 1992, *15* (5), 461–469.

Dean, Emma, Frank Schilbach, and Heather Schofield, "Poverty and Cognitive Function," in Carter Michael R. Barrett, Christopher B. and Jean-Paul Chavas, eds., *The Economics of Poverty Traps*, NBER/University of Chicago Press, 2019, chapter 2, pp. 57–118.

Dean, Joshua T, "Noise, cognitive function, and worker productivity," Technical Report, mimeo 2019.

DellaVigna, Stefano, Devin Pope, and Eva Vivalt, "Predict science to improve science," *Science*, 2019, *366* (6464), 428–429.

Dickinson, David L. and Todd McElroy, "Sleep Restriction and Circadian Effects on Social Decisions," *European Economic Review*, 2017, *97*, 57–71.

Dinges, David F., "Adult Napping and Its Effects on Ability to Function," in Claudio Stampi, ed., *Why We Nap*, Boston, MA: Birkhäuser, 1992, chapter 9, pp. 118–134.

_ , Frances Pack, Katherine Williams, Kelly A. Gillen, John W. Powell, Geoffrey E. Ott, Caitlin Aptowicz, and Allan I. Pack, "Cumulative Sleepiness, Mood Disturbance, and Psychomotor Vigilance Performance Decrements During a Week of Sleep Restricted to 4-5 Hours per Night," *Sleep*, 1997, *20* (4), 267–277.

et al. Liu, Yong, "Sleep duration and chronic diseases among US adults age 45 years and older: evidence from the 2010 Behavioral Risk Factor Surveillance System," *Sleep*, 2013, *36* (10), 1421–1427.

F., P. Strollo P. Zee J. Walsh Luyster, "Sleep: A Health Imperative," *Sleep*, 2012, *35* (6), 727–734.

Gabaix, Xavier, "Behavioral Inattention," in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics: Foundations and Applications*, Vol. 2, Amsterdam: Elsevier/North-Holland, 2019.

Gibson, Matthew and Jeffrey Shrader, "Time Use and Labor Productivity: The Returns to Sleep," *Review of Economics and Statistics*, 2018, *100* (5), 783–798.

Gildner, Theresa, Melissa Liebert, Paul Kowal, Somnath Chatterji, and Josh Snodgrass, "Associations between Sleep Duration, Sleep Quality, and Cognitive Test Performance among Older Adults from Six Middle Income Countries: Results from the Study on Global Ageing and Adult Heath (SAGE)," *Sleep*, 2014, *10* (6), 613–621.

Girschik, Jennifer, Lin Fritschi, Jane Heyworth, and Flavie Waters, "Validation of self-reported sleep against actigraphy," *Journal of epidemiology*, 2012, *22* (5), 462–468.

Hirshkowitz, Max, Kaitlyn Whiton, Steven M. Albert, Cathy Alessi, Oliviero Bruni, Lydia DonCarlos, Nancy Hazen, John Herman, Eliot S. Katz, Leila Kheirandish-Gozal, David N. Neubauer, Anne E. O'Donnell, Maurice Ohayon, John Peever, Robert Rawding, Ramesh C. Sachdeva, Belinda Setters, Michael V. Vitiello, J. Catesby Ware, and Paula J. Adams Hillard, "National Sleep Foundation's Sleep Time Duration Recommendations: Methodology and Results Summary," *Sleep Health*, 2015, *1* (1), 40–43.

Holbein, John B, Jerome P Schafer, and David L Dickinson, "Insufficient sleep reduces voting and other prosocial behaviours," *Nature human behaviour*, 2019, *3* (5), 492.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Jackson, Chandra L., Sanjay R. Patel, W. Braxton Jackson II, Pamela L. Lutsey, and Susan Redline, "Agreement between Self-reported and Objectively Measured Sleep Duration among White, Black, Hispanic, and Chinese Adults in the United States: Multi-Ethnic Study of Atherosclerosis," *Sleep*, 2018, *41* (6), 1–12.

Jagnani, Maulik, "Poor Sleep: Sunset Time and Human Capital Production," *mimeo*, 2018.

Jin, Lawrence and Nicolas R Ziebarth, "Sleep, health, and human capital: Evidence from daylight saving time," *Journal of Economic Behavior & Organization*, 2020.

Kahneman, Daniel and Alan Krueger, "Development in the Measurement of Subjective Well-Being," *Journal of Economic Perspectives*, 2006, *20* (1), 3–24.

_ **and Angus Deaton**, "High income improves evaluation of life but not emotional well-being," *Proceedings of the national academy of sciences*, 2010, *107* (38), 16489–16493.

Kaur, Supreet, Sendhil Mullainathan, Frank Schilbach, and Suanna Oh, "Does financial strain lower worker productivity?," Technical Report, Working Paper 2019.

Killgore, William, "Effects of Sleep Deprivation on Cognition," *Progress in Brain Research*, 2010, *185*, 105–129.

Knutson, Kristen L, "Sleep duration, quality, and timing and their associations with age in a community without electricity in Haiti," *American Journal of Human Biology*, 2014, *26* (1), 80–86.

Kremer, Michael, Gautam Rao, and Frank Schilbach, "Behavioral development economics," *Handbook of Behavioral Economics*, 2019, *2.*

Kurina, Lianne M., Ronald A. Thisted, Jen-Hao Chen, MMartha K. McClintock, Linda J. Waite, and Diane S. Lauderdale, "Actigraphic Sleep Characteristics among Older Americans," *Sleep Health*, 2015, *1* (4), 285–292.

Kushida, Clete A., Arthur Chang, Chirag Gadkary, Christian Guilleminault, Oscar Carrillo, and William C. Dement, "Comparison of Actigraphic, Polysomnographic, and Subjective Assessment of Sleep Parameters in Sleep-disordered Patients," *Sleep Medicine*, 2001, *2* (5), 389–396.

Lauderdale, Diane S., Kristen L. Knutson, Lijing L. Yan, Kiang Liu, and Paul J. Rathouz, "Sleep Duration: How Well Do Self-Reports Reflect Objective Measures? The CARDIA Sleep Study," *Epidemiology*, 2008, *19* (6), 838–845.

Lichstein, K., K. Stone, J. Donaldson, S. Nau, J. Soeffing, D. Murray, K. Lester, and N. Aguillard, "Actigraphy Validation with Insomnia," *Sleep*, 2006, *2* (29), 232–239.

Lim, Julian and David F. Dinges, "A Meta-Analysis of the Impact of Short-Term Sleep Deprivation on Cognitive Variables," *Psychological Bulletin*, 2010, *136* (3), 375–389.

Lovato, Nicole and Leon Lack, "The Effect of Napping on Cognitive Functioning," in Gerard A. Kerkhof and Hans P.A. Van Dongen, eds., *Progress in Brain Research*, Vol. 185, Elsevier B.V., 2010, chapter 9, pp. 155–166.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao, "Poverty impedes cognitive function," *science*, 2013, *341* (6149), 976–980.

Marino, Miguel, Yi Li, Michael N. Rueschman, J.W. Winkelman, J.M. Ellenbogen, J. Solet, Hilary Dulin, Lisa F. Berkman, and Orfeu M. Buxton, "Measuring Sleep: Accuracy, Sensitivity, and Specificity of Wrist Actigraphy Compared to Polysomnography," *Sleep*, 2013, *36* (11), 1747–1755.

Martinez, Seung-Keun, Stephan Meier, and Charles Sprenger, "Procrastination in the Field: Evidence from Tax Filing," *New York, NY: Columbia Business School. doi*, 2017, *10*, 2466.

McKenna, Benjamin S., David L. Dickinson, Henry J. Orff, and Sean P.A. Drummond, "The Effects of One Night of Sleep Deprivation on Known-risk and Ambiguous-risk Decisions," *Journal of Sleep Research*, 2007, *16* (3), 245–252.

McKenzie, David, "Beyond Baseline and Follow-up: The Case for more T in Experiments," *Journal of Development Economics*, 2012, *99*, 210–221.

Meier, Stephan and Charles Sprenger, "Present-biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2010, *2* (1), 193–210.

Minges, Karl E and Nancy S Redeker, "Delayed school start times and adolescent sleep: A systematic review of the experimental evidence," *Sleep medicine reviews*, 2016, *28*, 86–95.

Ohayon, Maurice, Emerson M Wickwire, Max Hirshkowitz, Steven M Albert, Alon Avidan, Frank J Daly, Yves Dauvilliers, Raffaele Ferri, Constance Fung, David Gozal et al., "National Sleep Foundation's sleep quality recommendations: first report," *Sleep Health*, 2017, *3* (1), 6–19.

Riemann, Dieter and Michael L Perlis, "The treatments of chronic insomnia: a review of benzodiazepine receptor agonists and psychological and behavioral therapies," *Sleep medicine reviews*, 2009, *13* (3), 205–214.

Rosekind, Mark R, Kevin B Gregory, Melissa M Mallis, Summer L Brandt, Brian Seal, and Debra Lerner, "The cost of poor sleep: workplace productivity loss and associated costs," *Journal of Occupational and Environmental Medicine*, 2010, *52* (1), 91–98.

Roure, Nuria, Silvia Gomez, Olga Mediano, Joaquin Duran, Monica de la Peña, Francisco Capote, Joaquin Teran, Juan Fernando Masa, Maria Luz Alonso, Jaime Corral et al., "Daytime sleepiness and polysomnography in obstructive sleep apnea patients," *Sleep medicine*, 2008, *9* (7), 727–731.

Sadeh, A., P. Hauri, D. Kripke, and P. Lavie, "The Role of Actigraphy in the Evaluation of Sleep Disorders," *Sleep*, 1995, *18* (4), 288–302.

Sadeh, Avi, "The Role of Actigraphy in Sleep Medicine: An Update," *Sleep Medicine Review*, 2011, *15*, 259–267.

Schilbach, Frank, "Alcohol and self-control: A field experiment in India," *American economic review*, 2019, *109* (4), 1290–1322.

Schofield, Heather, "The economic costs of low caloric intake: Evidence from India," *Unpublished Manuscript*, 2014.

Schokman, Aaron, Yu Sun Bin, Guido Simonelli, Jonathon Pye, Richard Morris, Athula Sumathipala, Sisira H Siribaddana, Matthew Hotopf, Fruhling Rijsdijk, Kaushalya Jayaweera et al., "Agreement between subjective and objective measures of sleep duration in a low-middle income country setting," *Sleep health*, 2018, *4* (6), 543–550.

Selvamani, Y., Perianayagam Arokiasamy, Mamta Chaudhary, and Himanshu, "Association of Sleep Problems and Sleep Duration with Self-rated Health and Grip Strength among Older Adults in India and China: Results from the Study on Global Aging and Adult Health (SAGE)," *Journal of Public Health*, 2018, *26* (6), 697–707.

Simonelli, Guido, Nathaniel S Marshall, Antigone Grillakis, Christopher B Miller, Camilla M Hoyos, and Nick Glozier, "Sleep health epidemiology in low and middle-income countries: A systematic review and meta-analysis of the prevalence of poor sleep quality and sleep duration," *Sleep Health*, 2018, *4* (3), 239–250.

Smith, Austin C, "Spring forward at your own risk: Daylight saving time and fatal vehicle crashes," *American Economic Journal: Applied Economics*, 2016, *8* (2), 65–91.

Smith, Michael T, Christina S McCrae, Joseph Cheung, Jennifer L Martin, Christopher G Harrod, Jonathan L Heald, and Kelly A Carden, "Use of actigraphy for the evaluation of sleep disorders and circadian rhythm sleep-wake disorders: an American Academy of Sleep Medicine clinical practice guideline," *Journal of Clinical Sleep Medicine*, 2018, *14* (07), 1231–1237.

Sprenger, Charles, "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Journal of Political Economy*, 2015, *123* (6), 1456–1499.

Stepanski, Edward J, "The effect of sleep fragmentation on daytime function," *Sleep*, 2002, *25* (3), 268–276.

Stranges, Saverio, William Tigbe, Francesc Gomze-Olive, Margaret Thorogood, and Ngianga-Bakwain Kandala, "Sleep Problems: An Emerging Global Epidemic? Findings From the INDEPTH WHO-SAGE Study Among More Than 40,000 Older Adults From 8 Countries Across Africa and Asia," *Sleep*, 2012, *35* (8), 1173–1181.

Strine, T. and D. Chapman, "Associations of Frequent Sleep Insufficiency with Health-Related Quality of Life and Health Behaviors," *Sleep Medicine*, 2005, *6* (1), 23–27.

Takahashi, M, "The Role of Prescribed Napping in Sleep Medicine," *Sleep Medicine Reviews*, 2003, *7* (3), 227–235.

Touitou, Yvan, Alain Reinberg, and David Touitou, "Association between light at night, melatonin secretion, sleep deprivation, and the internal clock: Health impacts and mechanisms of circadian disruption," *Life sciences*, 2017, *173*, 94–106.

UK National Health Services, "Sleep Problems in the UK Highlighted," 2011.

Van Dongen, Hans P.A., Greg Maislin, Janet M. Mullington, and David F. Dinges, "The Cumulative Cost of Additional Wakefulness: Dose-Response Effects on Neurobehavioral Functions and Sleep Physiology from Chronic Sleep Restriction and Total Sleep Deprivation," *Sleep*, 2003, *26* (2), 117–126.

Walker, Matthew P., *Why We Sleep: Unlocking the Power of Sleep and Dreams*, New York: Scribner, 2017.

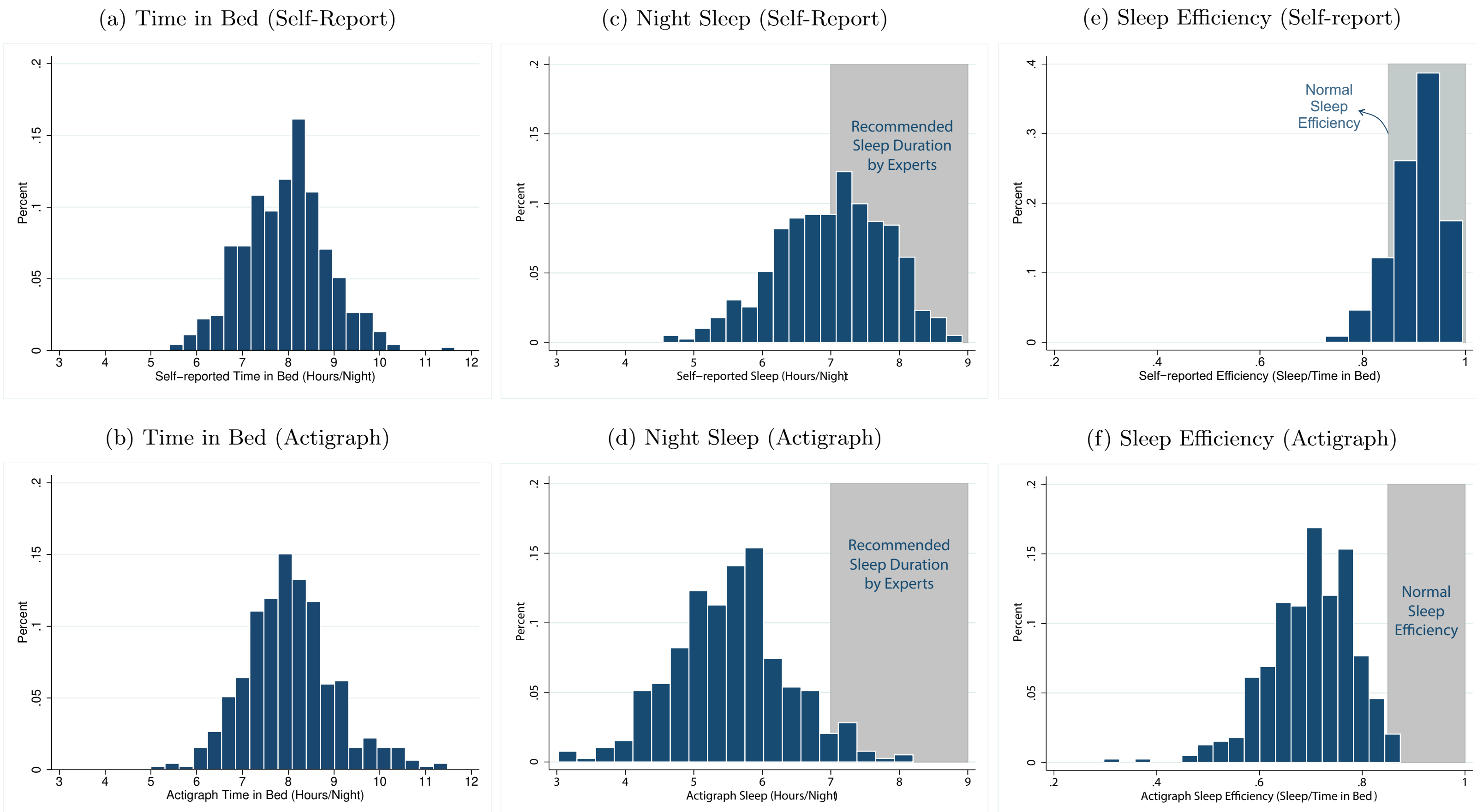
Watson, Nathaniel F., Safwan M. Badr, Gregory Belenky, Donald L. Bliwise, Orfeu M. Buxton, Daniel Buysse, David F. Dinges, James Gangwisch, Michael A. Grandner, Clete Kushida, Raman K. Malhotra, Jennifer L. Martin, Sanjay R. Patel, Stuart F. Quan, and Esra Tasali, "Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society on the Recommended Amount of Sleep for a Healthy Adult: Methodology and Discussion," *Sleep*, 2015, *38* (8), 1161–1183.

Yoon, In-Young, Daniel F. Kripke, Shawn D. Youngstedt, and Jeffrey A. Elliott, "Actigraphy Suggests Age-related Differences in Napping and Nocturnal Sleep," *Journal of Sleep Research*, 2003, *12*, 87–93.

Young, Alwyn, "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *The Quarterly Journal of Economics*, May 2019, *134* (2), 557–598. 00000.

Tables and Figures

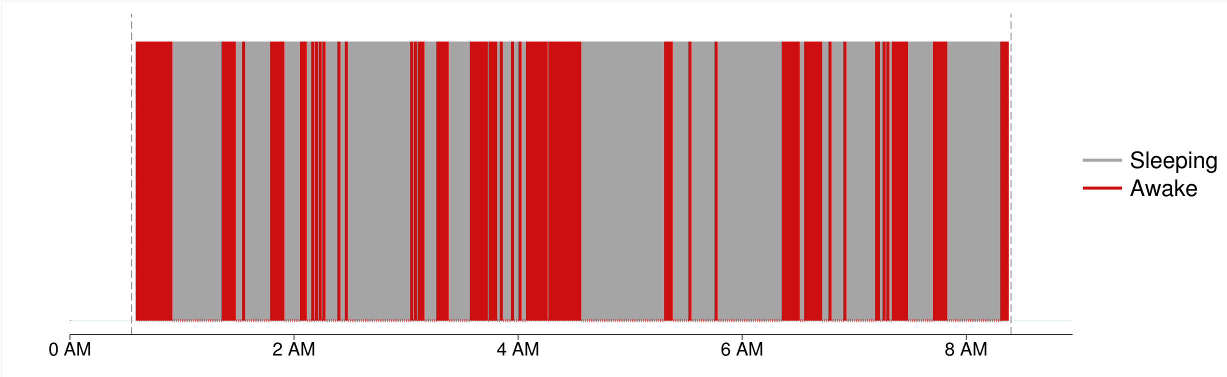
FIGURE I: Baseline Distributions of Sleep-Related Variables



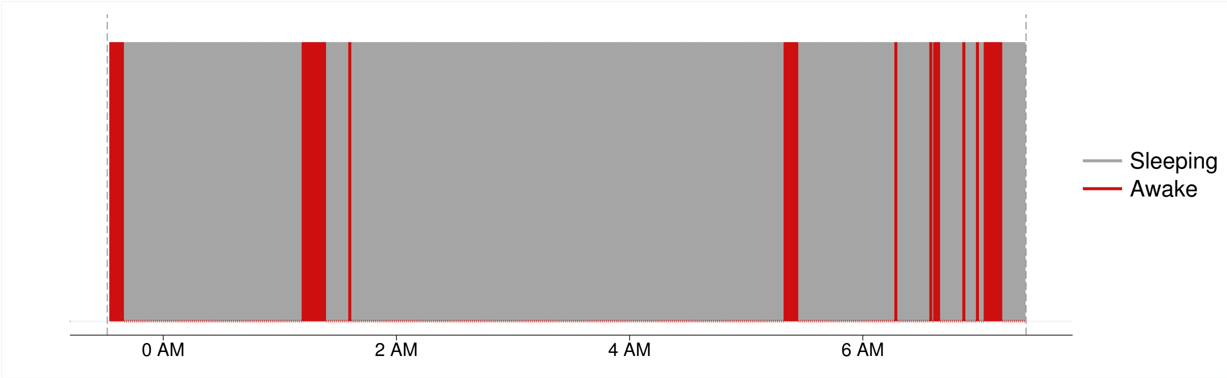
Notes: This figure shows the distribution of the participant-level baseline average of sleep-related variables. In panels (a) and (b), we plot hours in bed as reported by the participant and measured by the actigraph, respectively. Panels (c) and (d) plot self-reported and actigraph-measured night sleep. Panels (e) and (f) plot sleep efficiency (night time sleep / time in bed) via self-reports and as measured by the actigraph.

FIGURE II: Regular and Good Nights of Sleep

(a) Regular

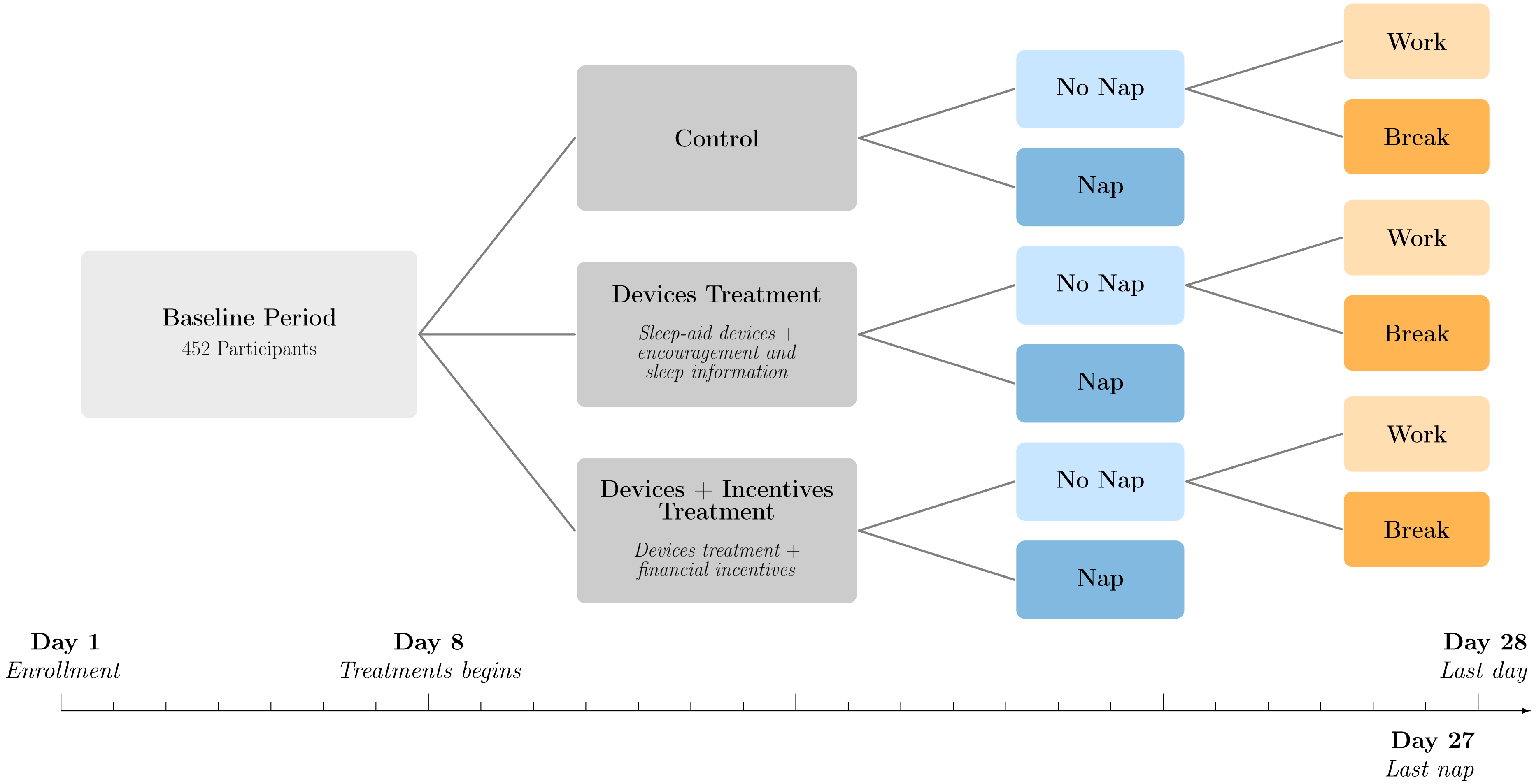


(b) Good



Notes: This figure represents actual two nights of participants sleep: an "average" night and a "good" night. Gray areas indicate one-minute epochs where the participant is asleep while red areas indicate epochs where the participant awake. The light gray dashed lines indicate when the participant get into or out of bed. Figure IIa depicts a "regular" night of sleep in our sample: the participant stays in bed for 7 hours and 45 minutes but only actually sleeps for 5 hours and 20 minutes, resulting in a sleep efficiency of roughly 69%. Moreover, this participant awakes 31 times during the night and the longest sleep episode he achieves lasts only 45 minutes. On the other hand, Figure IIb depicts a "good" night of sleep. The participant stays in bed for 7 hours and 53 minutes and sleeps for 7 hours and 8 minutes, resulting in a sleep efficiency of roughly 90%. In contrast to the "typical" night of sleep, in the "good" night, the participant only awakes 9 times during the night and the longest sleep episode lasts 202 minutes.

FIGURE III: Experimental Design and Timeline



Notes: This figure presents an overview of the timeline and experimental design of the study. After the 8 baseline days, the 452 participants are first divided in 3 groups: control, sleep devices, and sleep devices plus incentives. Participants in each of these groups were further randomized between a nap group, which was allowed and encouraged to use a nap station in the office in the early afternoon, and a no nap group. While all these randomization's occurred between participants, participants in the no nap group were further randomized on a daily level either to being allowed to work during or to take a mandatory pause during the nap period. The nap treatment ends at day 27, and the participants return the sleep devices on day 28. Finally, endline surveys occur on day 28 or shortly thereafter.

FIGURE IV: Treatments to Improve Sleep

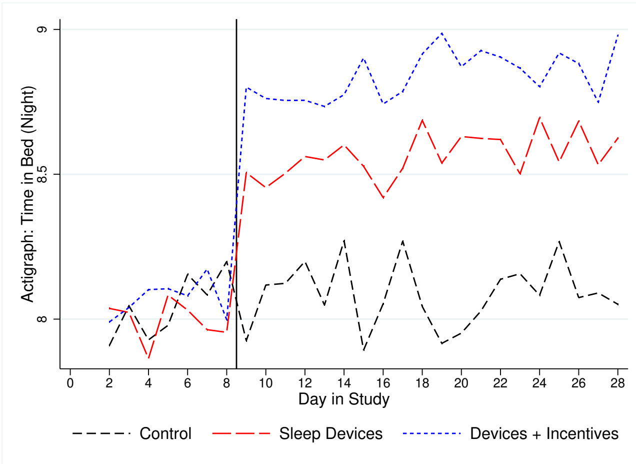
(a) Devices to Improve Night Sleep Environments

(b) Nap Station

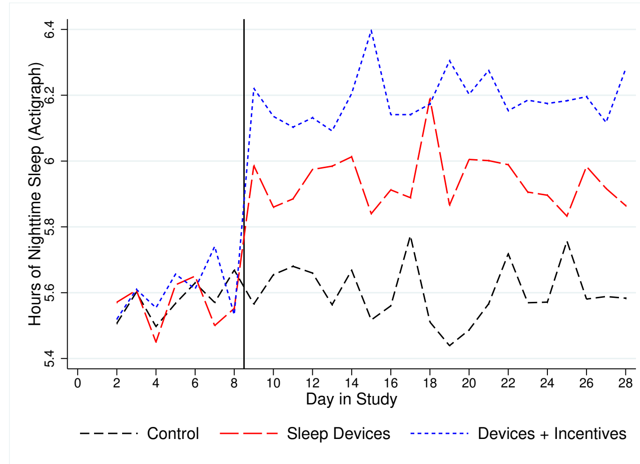
Notes: Panel (a) of this figure displays the items offered to individuals in the sleep devices group. These items were loaned to the participants, who could borrow as many units of the items as they wished. The items were brought to the participant's home on day 8 and retrieved on day 28 by surveyors. A subset of the participants in the control group also received household goods unrelated to sleep in order to allow us to test for (and if needed, estimate) experimental demand or reciprocity effects. Panel (b) of this figure shows the nap station where participants in the nap group were allowed and encouraged to sleep, for up to 30 minutes, in the early afternoon. The participants in the no nap group were not allowed to use this nap station, which was located on a separate floor at the study office.

FIGURE V: Impacts of Night Sleep Treatments on Sleep

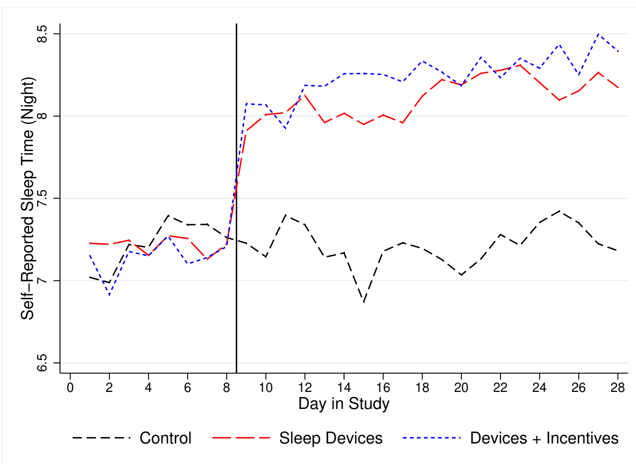
(a) Impact of Treatments on Time in Bed (Actigraph)



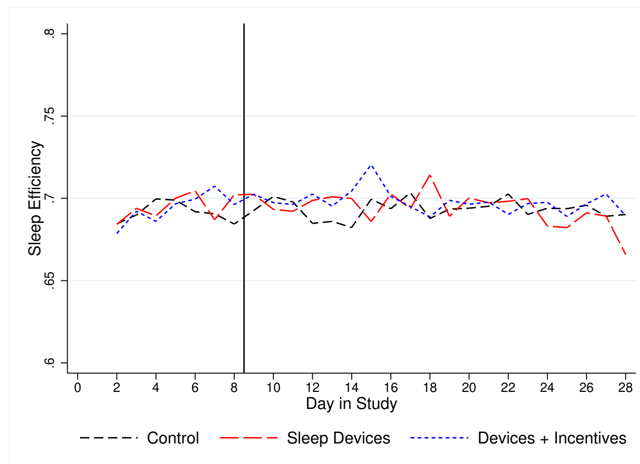
(b) Impact of Treatments on Night Sleep (Actigraph)



(c) Impact of Treatments on Night Sleep (Self-Reported)



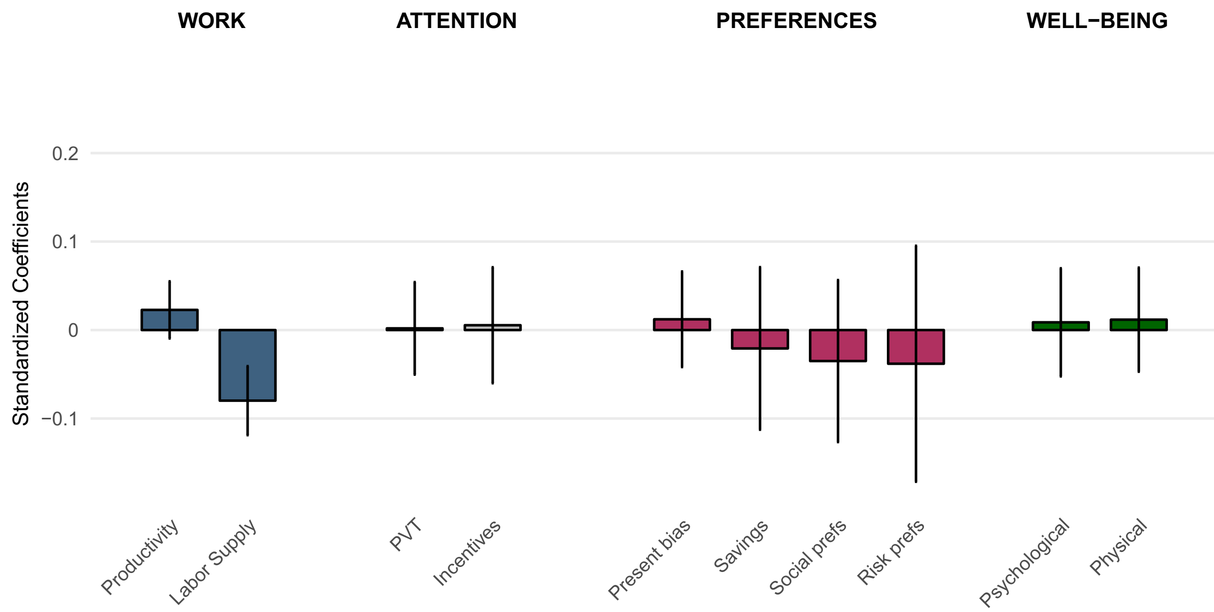
(d) Impact of Treatments on Sleep Efficiency (Actigraph)



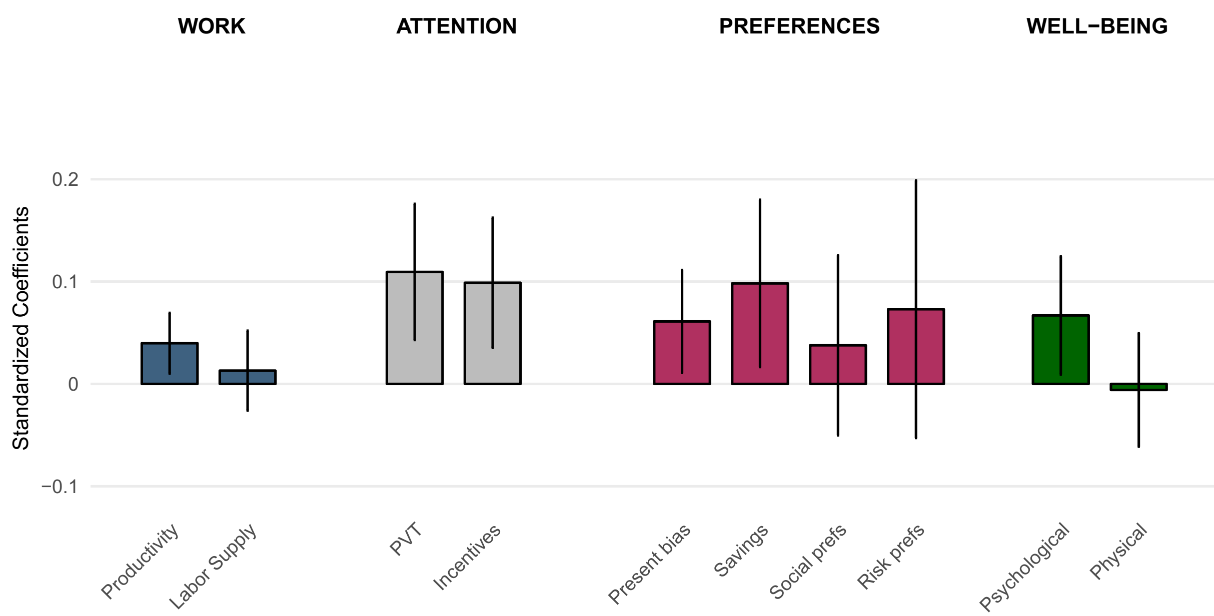
Notes: This figure shows the average of different sleep-related variables for each night sleep treatment group by day in study. In panels (a) and (b), we plot, respectively, the series for hours of night sleep and hours in bed as measured by the actigraph. In panel (c), we plot self-reported night sleep. In panel (d), we plot the series for Sleep Efficiency (Nighttime Sleep / Time in Bed) as measured by the actigraph. The sample is restricted to long days, where participants can work until 8 PM.

FIGURE VI: Summary of Treatment Effects

(a) Summary of Effects for Night Sleep Treatments



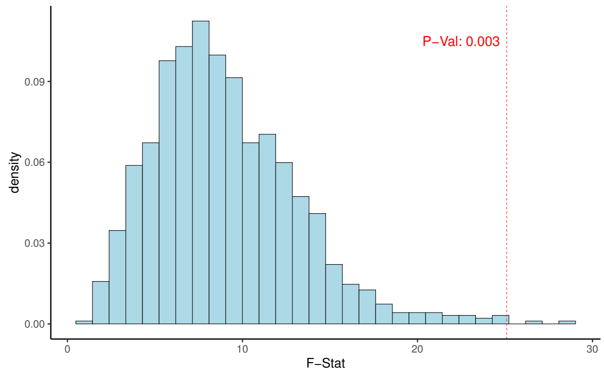
(b) Summary of Effects for Nap Treatment



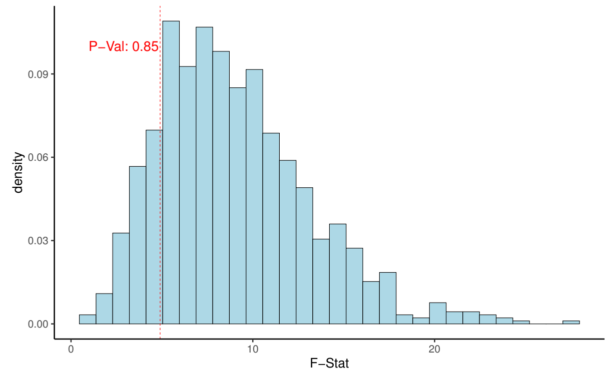
Notes: This figure summarizes the treatment effects for the night sleep treatments and the nap treatment. We plot the point estimates and confidence intervals with respect to the night sleep interventions in Panel (a), and nap intervention in Panel (b). All outcomes are standardized measures with the exception of "Incentives", which corresponds to differences between treatment and control coefficients as shown in Table IV, panel B, column 1. The coefficient for "Savings" corresponds to deposits (net savings winsorized at zero) and the coefficient for present bias corresponds to the structurally estimated β in the full sample. Outcomes with multiple components are condensed in indices.

FIGURE VII: Joint Test of Significance of All Treatment Effects

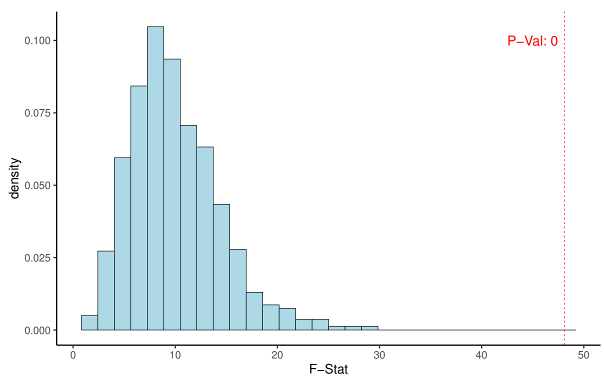
(a) Nap Without Labor Supply



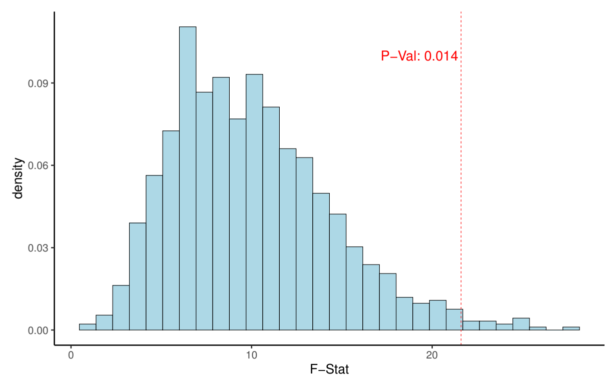
(b) Night-Sleep Without Labor Supply



(c) Nap



(d) Night-Sleep



Notes :

- *The blue bars represent the distribution of the Wald (F) statistics for joint significance of the treatment effect under the null that all treatment effects are equal to zero for a given treatment. This distribution is obtained through 1000 permutations of the treatment assignment. The red dotted line shows the original Wald statistic, i.e., the one for the treatment effects we estimate with the full sample. This test for joint significance follows Young (2019).*
- *In panels (a) and (b) we consider 9 outcome variables for the joint test: productivity in the typing task, savings, time, risk and social preferences, two measures of attention, and physical and psychological well-being. In panels (c) and (d) we consider the same outcome variables and also include hours typing.*
- *In panels (a) and (c), we show the distribution for the nap treatment, while in panels (b) and (d), we show the distribution for the night-sleep treatment.*

Table I: Treatment Effects on Sleep

	Night Sleep (1)	**Time in Bed** (2)	**Sleep Efficiency** (3)	**Nap Sleep** (4)	**24 Hr Sleep** (5)
Devices Treat	0.331***	0.515***	-0.00440	-0.00275	0.274***
	(0.0559)	(0.0619)	(0.00471)	(0.00552)	(0.0561)
Devices + Incentives Treat	0.555***	0.761***	0.00219	-0.00167	0.482***
	(0.0574)	(0.0679)	(0.00490)	(0.00542)	(0.0577)
Nap Treat	-0.0842*	-0.166***	0.00273	0.242***	0.122**
	(0.0481)	(0.0525)	(0.00394)	(0.00458)	(0.0482)
Baseline	0.788***	0.596***	0.744***		0.814***
	(0.0302)	(0.0344)	(0.0317)		(0.0308)
Control Mean	5.611	8.067	0.699	0	5.616
Control SD	1.199	1.368	0.113	0	1.189
N	8454	8454	8454	7191	8018
Participants	451	451	451	450	448

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table considers the treatment effect of the two night sleep and nap interventions on sleep patterns.

- Night Sleep, Time in Bed, Nap Sleep and 24-Hour Sleep (columns 1, 2, 4 and 5, respectively) are measured in hours. Sleep Efficiency (column 3) is the ratio between Night Sleep and Time in Bed. 24-Hour Sleep adds Nap Sleep to Night Sleep.
- Each column shows the OLS estimates of equation (2) separating the two night sleep treatments, controlling for the average baseline measure of the dependent variable (ANCOVA), age, sex, and day-in-study and date fixed effects.
- Differences in number of observations between the three first columns and the last two stem from two sources. First, column 4 and 5 does not include day 28 of the sample, a day in which there is no nap sleep in the office.
- Standard errors are clustered at the participant level.

Table II: Night Sleep Treatment Effects on Data-Entry Work Performance

	Productivity			Hours Typing			Earnings		
		Quality	Quantity		Quality	Quantity		Quality	Quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Night Sleep Treat	46.30	-36.93	107.26**	-0.17***	-0.16**	-0.08	-7.64	-14.97**	4.46
	(39.79)	(52.31)	(50.52)	(0.05)	(0.07)	(0.06)	(5.09)	(6.12)	(6.39)
	[0.778]			[0.008]			[0.562]		
Above Median		-139.44**	93.56		0.01	0.13		-5.44	19.54**
		(68.18)	(68.47)		(0.09)	(0.08)		(8.77)	(9.22)
Night Sleep Treat × Above Median		168.79**	-122.18		-0.02	-0.19*		14.42	-24.29**
		(78.81)	(80.19)		(0.10)	(0.10)		(10.15)	(10.53)
Baseline	1.08***	1.08***	1.08***	0.50***	0.50***	0.50***	1.14***	1.14***	1.14***
	(0.01)	(0.01)	(0.01)	(0.17)	(0.17)	(0.17)	(0.04)	(0.04)	(0.04)
Control Mean	3533.39	3496.50	3253.24	4.07	4.06	4.06	283.32	277.03	262.22
Control SD	2031.42	2053.18	1930.44	2.14	2.14	2.12	218.86	214.08	194.52
N	7350	7350	7350	7351	7351	7351	7351	7351	7351
Participants	451	451	451	451	451	451	451	451	451

**** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level*

Notes: This table considers the treatment effect of the night sleep interventions on work outcomes.

- Each column shows the OLS estimates of equation (2) pooling the two night sleep treatments, controlling for the baseline averages of the dependant variable (ANCOVA), age, sex, a dummy indicating whether the day is a "regular day", the fraction of High piece rates randomized for that day, and day-in-study and date fixed effects. Standard errors are clustered at the participant level.
- Columns (1), (4), and (7) present results for the night sleep treatment on work outcomes. Columns (2), (5), and (8) interact the night sleep treatment with a dummy indicating whether or not the participant's baseline sleep quality is above the sample average baseline quality median. Columns (3), (6), and (9) interact the night sleep treatment with a dummy indicating whether or not the participant's baseline sleep quantity is above the sample average baseline quantity median. Quality is defined as a standardized index of number of awakenings, sleep efficiency, and longest sleep episode.
- The differing number of observations is due to a participant being present at the office one day without working positive hours, resulting in one less observation for productivity (column 1).
- Corrected p-values that control for the Family-Wise Error Rates are included in brackets in columns 1, 4 and 7. A full description of our approach to multiple hypothesis corrections can be found in Appendix E.

Table III: Nap Treatment Effects on Data-Entry Work Performance

	Productivity	Labor Supply			Earnings	
Comparison	Output/Hour (1)	Days in Office (2)	Hours in Office (3)	Hours Typing (4)	Performance (5)	Overall (6)
Nap Treat vs. Break	82.45**	0.79	0.06	0.03	9.79**	11.28**
	(36.33)	(0.49)	(0.04)	(0.05)	(4.18)	(4.87)
	[0.128]			[0.985]		[0.116]
Nap Treat vs. No Break	74.61**	-0.33	0.06	-0.44***	-14.18***	-22.67***
	(36.64)	(0.44)	(0.04)	(0.05)	(4.27)	(4.96)
	[0.212]			[0.000]		[0.000]
Control Mean	3401.36	16.48	6.71	4.05	184.30	273.69
Control SD	1834.62	4.00	2.93	2.07	165.90	198.31
N	7350	452	7348	7351	7351	7351
Participants	451	452	451	451	451	451

*** *Significant at the 1 percent level,* ** *Significant at the 5 percent level,* * *Significant at the 10 percent level*

Notes: This table considers the treatment effect of the nap treatment on work outcomes.

- Row 1 shows the treatment effect of the nap intervention in comparison to participants required to take a break during the nap time. Row 2 is analogous, but the comparison group consists of participants who were allowed to work during the nap period.
- The dependent variables are: productivity, defined as output per hour typing; labor supply outcomes capturing, respectively, the number of days present in the office, overall hours in the office, and hours actively typing; earnings from data-entry work capturing performance earnings and overall earnings (adding payments for time working as well), respectively.
- Each column shows the OLS estimates of equation (2), controlling for baseline average values of the dependant variable (ANCOVA), age, sex, a dummy indicating whether the day is a "regular day", fraction of High piece-rate sessions, and day in study and date fixed effects. Standard errors are clustered at the participant level.
- The differing number of observations across regressions comes from three sources. First, in one day a participant was present at the office but did not work, resulting in one less observation for productivity (column 1). Second, check-in and check-out times are missing for three observations, which are used to compute hours in office (column 3). Furthermore, one participant was absent in all postline period although he kept using the watch, resulting in 451 participants (instead of 452 as in column 2).
- Corrected p-values that control for the Family-Wise Error Rates are included in brackets in columns 1, 4 and 6. A full description of our approach to multiple hypothesis corrections can be found in Appendix E.

Table IV: Treatment Effects on Attention

Panel A: PVT

	Indices			Individual Components		
	Anderson (1)	Average (2)	Payment (3)	Inverse RT (4)	Minor Lapses (5)	False Starts (6)
Night Sleep Treat	0.00193	0.00181	0.0130	-0.0171	0.0245	-0.0220
	(0.0318)	(0.0348)	(0.0425)	(0.0417)	(0.0405)	(0.0463)
Nap Treat	0.109***	0.117***	0.169***	0.111**	0.151***	0.00133
	(0.0404)	(0.0436)	(0.0555)	(0.0466)	(0.0527)	(0.0508)
Baseline		0.182*	0.410***	0.579***	0.424***	0.0392*
		(0.104)	(0.0576)	(0.0194)	(0.0256)	(0.0237)
Raw Control Mean			13.17	2.808	3.523	1.738
Raw Control SD			1.503	0.287	4.592	5.131
N	6962	6962	6962	6962	6962	6962
Participants	452	452	452	452	452	452

Panel B: Incentives

	Overall		Morning		Afternoon	
	Output (1)	Minutes (2)	Output (3)	Minutes (4)	Output (5)	Minutes (6)
Night Sleep Treat	0.85	0.80	0.83	0.94	0.85	0.80
	(0.03)	(0.13)	(0.05)	(0.54)	(0.04)	(0.12)
Nap Treat	0.94	0.99	0.85	0.84	0.97	0.96
	(0.04)	(0.16)	(0.05)	(0.61)	(0.06)	(0.14)
Control	0.84	0.77	0.80	0.65	0.86	0.75
	(0.04)	(0.15)	(0.05)	(0.38)	(0.06)	(0.13)
p-value NS vs Control	0.891	0.830	0.576	0.569	0.877	0.677
	[1.000]					
p-value Nap vs Control	0.011	0.110	0.282	0.710	0.023	0.060
	[0.130]					
N	71596	71596	29241	29241	42355	42355
Participants	451	451	450	450	451	451

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table considers the treatment effects of the night sleep and nap interventions on alertness and attention.

- Panel A: Psychomotor Vigilance Task (PVT)
  - Each column shows the OLS estimates of equation (2), controlling for baseline average values of the dependant variable (ANCOVA), age, sex, whether participants faced high or low incentives for the task (which varied randomly within-participant each day), whether participants were randomized to the Work group in the afternoon, and day in study and date fixed effects.
  - Column 1 to 3 show different indices of performance, while columns 4 to 6 show results for each component outcome of the PVT task.
  - Column 1 index averages the standardized outcomes accounting for correlation across measures (Andersen 2008). Column 2 index is a simple average of the standardized outcome variables. Column 3 is the overall payment for the task, which was determined before the study began and weights inverse reaction times more than the other components.
  - Column 4 shows treatment effects for the inverse reaction time (a higher inverse reaction time implies greater alertness). Column 5 depicts minor lapses (significant delays between when the signal appears and the participant acts). The outcome variable in column 6 is the number of false starts (when the participant acts before the signal is displayed). Signs are flipped for columns 5 and 6 such that positive values indicate *fewer* minor lapses and false starts.
  - Standard errors are clustered at the participant level. Non-standardized means and standard deviations for the control group are displayed at the bottom of the table.
- Panel B: Attention to Incentives
  - Each column shows the OLS estimates of a modified equation (2), where we control for date, day in study and participant fixed effects instead of following the usual ANCOVA specification.
  - Panel B display the attention parameter θ – which captures attention as the ratio of the reaction to high incentive between the non-salient and the salient conditions – across treatment groups. Details of the definition and estimation of θ are presented in Section 6.1.
  - Columns 1 and 2 include the entire day, while columns 3 and 4 only include windows prior to the nap (morning), and 5 and 6 only include windows after nap time (afternoon).
  - The two dependent variables are Output and Minutes Actively Typing, all at the 30-minute incentive-window level.
  - We consider attention for three groups: First, the control group, which in this regression consists of individuals in the intersection between the no sleep devices and no nap groups. Second, the night sleep treatment group, which pools both sleep devices and devices + incentive groups. Third, the nap treatment group.
  - Rows 4 and 5 depict the p-value of a test of differences between the coefficients estimated for each group.
  - Corrected p-values that control for the Family-Wise Error Rates are included in brackets in Column 1 in Panel B. A full description of our approach to multiple hypothesis corrections can be found in Appendix E.
  - Standard errors are clustered at the participant level.

Table V: Treatment Effects on Psychological Well-Being

	Indices			**Standardized Components**				
	Anderson (1)	Average (2)	Pre-Reg. (3)	Depression (4)	Happiness (5)	Life Possibility (6)	Life Satisfaction (7)	Stress (8)
Night Sleep Treat	-0.02	-0.02	-0.05	-0.12	0.05	-0.01	0.01	-0.08
	(0.05)	(0.05)	(0.06)	(0.10)	(0.06)	(0.08)	(0.08)	(0.08)
Nap Treat	0.09	0.08*	0.13**	0.06	0.18***	0.24***	0.13*	0.01
	(0.05)	(0.05)	(0.05)	(0.10)	(0.06)	(0.07)	(0.07)	(0.08)
Baseline		0.78***	0.73***	0.25***	0.68***	0.67***	0.63***	0.54***
		(0.04)	(0.04)	(0.06)	(0.03)	(0.04)	(0.04)	(0.04)
N	442	442	442	445	7690	2217	2217	2217
Participants	442	442	442	445	452	446	446	446

Notes: This table considers the treatment effect of the night sleep and nap interventions on measures of psychological well-being.

- All variables are standardized by the control group's average and standard deviation. Signs are flipped when needed such that higher outcomes always indicate more desirable outcomes.
- The outcome variables in columns 1 and 2 are weighted averages of the 5 standardized well-being outcomes. Column 1 averages the outcomes accounting for correlation across measures (Anderson, 2008), while column 2 is a simple average. Column 3 is a simple average of the three pre-registered outcomes - depression, happiness, and life possibilities.
- The outcomes in columns 4 to 8 are (4) self-reported depression, at endline (we use PHQ-8, excluding the sleep-related question to avoid mechanical correlations); (5) self-reported happiness, where a score of 1 means "not at all happy" while a score of 5 means "very happy"; (6) ladder of life possibility (Cantril Scale), where participants were asked, "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?"; (7) life satisfaction (from Gallup Survey), where participants are asked, "All things considered, how satisfied are you with your life as a whole?" (0 Dissatisfied to 10 Satisfied); (8) self-reported stress (part of Cohen et al. (1983)), where an answer of 1 means "none of the time" while 6 means "a lot of the time."
- We pool the two night sleep treatments and control for age, sex, day-in-study and date fixed effects, and a baseline measures of the dependant variable (ANCOVA) when available. When the dependent variables comprise an index, we control for the index at baseline.

Table VI: Treatment Effects on Physical Well-being

	Indices			Standardized Components					
	(Index 1 - Manski LB)	(Index 1 - Manski UB)	(Index 2)	Biking	Illness	Pain	Daily Act	BP	Steps
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Night Sleep Treat	0.04	0.06	0.06	0.10	0.07	0.09	0.10	0.02	-0.09***
	(0.03)	(0.03)	(0.04)	(0.10)	(0.06)	(0.09)	(0.09)	(0.03)	(0.03)
Nap Treat	0.04	0.04	0.01	-0.14	0.06	-0.06	0.11	0.05*	0.02
	(0.03)	(0.03)	(0.04)	(0.09)	(0.04)	(0.08)	(0.09)	(0.03)	(0.02)
Baseline			0.33***	0.13	-0.01	0.19***	0.22***	0.77***	0.86***
			(0.04)	(0.10)	(0.03)	(0.05)	(0.05)	(0.02)	(0.01)
N	439	439	439	367	439	439	439	2214	8615
Participants	439	439	439	367	439	439	439	443	449

Notes: This table considers the treatment effect of the night sleep and nap interventions on physical well-being outcomes.

- All variables are standardized by the control group's average and standard deviation. Signs are flipped when needed such that higher values always indicate more desirable outcomes.
- The two indices present the aggregated results for pre-registered outcomes (all components shown here with the exception of daily steps). Columns 1 and 2 display Manski bounds (lower and upper, respectively) applied to the Anderson index (Anderson, 2008) to account for missing biking task data (some participants opted out of the task). In Column 3, Index 2 was set equal to the average of the remaining standardized health outcomes in the case of missing biking data.
- The remaining dependent variables are: (4) biking task performance, (5) days in the last week with self-reported illness, (6) self-reported pain on a scale from 1 to 10, (7) score of days in the last week that health impaired daily activities, (8) an average of standardized, winsorized systolic and diastolic blood pressure, and (9) total number of daily steps (which is not a pre-registered outcome). More details of the health outcomes can be found in Section 6.2.2 and Appendix B.5.
- Each column shows the OLS estimates of an equation similar to Equation (2). In the regression, we pooled the two night sleep treatments and controlled for age, sex, day in study and date fixed effects, and baseline measure of the dependent variable (ANCOVA) when available. When the dependent variable is an index, we control for the index at baseline.

Table VII: Treatment Effects on Savings and Time Preferences

Panel A: Savings

	Savings		Interest Accrued	
	Deposits	Net Savings	Real Pos. Rates	Hypothetical 1%
	(1)	(2)	(3)	(4)
Night Sleep Treat	-3.46	-9.64	0.47	-0.12
	(9.30)	(11.80)	(1.74)	(0.98)
	[1.000]			
Nap Treat	16.37**	9.72	2.28	1.54*
	(8.28)	(11.11)	(1.60)	(0.88)
	[0.434]			
Interest Rate	34.42***	39.43***	17.25***	3.91***
	(8.61)	(11.27)	(3.52)	(0.93)
Baseline	0.83***	0.57***	0.42***	0.25***
	(0.05)	(0.08)	(0.05)	(0.02)
Control Mean	113.29	71.97	10.63	8.70
Control SD	166.68	325.68	19.66	15.43
N	8574	8574	5534	8574
Participants	452	452	292	452

Panel B: Present Bias

	Structural Beta (β)		Ratio Now vs. Later			
			Restricted		Unrestricted	
	Full Sample	New Version	Full Sample	New Version	Full Sample	New Version
	(1)	(2)	(3)	(4)	(5)	(6)
Night Sleep Treat	0.012	0.053	0.012	0.048	0.015	0.047
	(0.033)	(0.047)	(0.038)	(0.059)	(0.036)	(0.055)
	[1.000]					
Nap Treat	0.061**	0.079*	0.056	0.089	0.030	0.048
	(0.031)	(0.046)	(0.036)	(0.057)	(0.036)	(0.055)
	[0.416]					
Baseline Beta	0.041	0.073	0.015	0.038	0.006	0.028
	(0.043)	(0.064)	(0.041)	(0.058)	(0.040)	(0.056)
Control Mean	0.92	0.89	0.87	0.81	0.88	0.81
Control SD	0.34	0.34	0.39	0.46	0.38	0.45
N	352	214	352	214	398	252

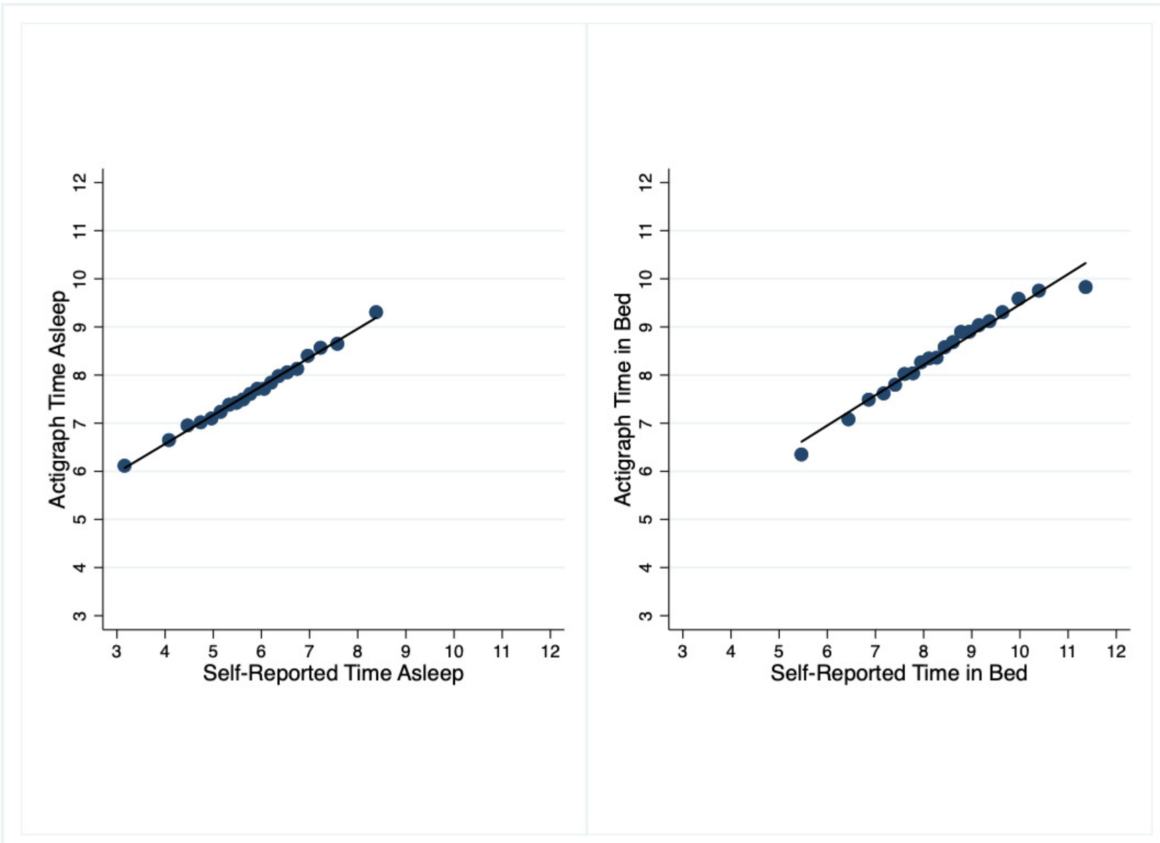
**** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level*

Notes: This table considers the treatment effects of the night sleep and the nap interventions on time preferences.

- Panel A: Savings
  - The dependent variable in column 1 captures daily deposits (which is equivalent to winsorizing daily net savings at Rs. 0) at the study office. Column 2 shows daily net savings (difference between deposits and withdrawals). Columns 3 and 4 show daily interest accrued on the participants savings, with column 3 excluding individuals who were assigned a zero interest rate and column 4 including the whole sample, but assuming all participants faced a 1% interest rate.
  - Each column shows the OLS estimates of equation (2), controlling for the baseline average of the dependant variable (ANCOVA), age, sex, daily piece rates for each incentive sessions, interest rate, maximum payment from cognitive tasks, dummy for risk and social activity day, the randomized piece rate for the present bias task, surveyor fixed effects, and the amount defaulted for savings.
  - Standard errors are clustered at the participant level.
- Panel B: Present-bias
  - The dependent variable throughout Panel B is one of our two preferred measures of present bias parameter: the structurally estimated β and the percentage decrease in effort chosen on "work-days" (OLS). In all columns, we present the treatment effect of the night sleep and the nap interventions on the present-bias parameter, controlling for the baseline value of the dependent variable (one of the measures of present bias) and other relevant controls. In Columns 2, 4, and 6 we also present results restricting the sample to the participants which engaged in the new version of the present-bias task (See Appendix B.6 for more details).
  - The dependent variable in columns 1 and 2 is our preferred structurally-estimated present bias parameter. Column 2 restricts the sample to participants with an updated version of the task where the salience differences were made more stark. We exclude individuals for whom the structural estimator did not converge.
  - The dependent variable in columns 3 and 4 is the OLS present bias parameter. Column 4 is similarly restricted to the updated version of the task. In this specification, we also exclude the participants for whom the structural estimator did not converge.
  - The dependent variable in columns 5 and 6 is the OLS present bias parameter. Column 6 is similarly restricted to the updated version of the task. This sample includes all participants who completed the present bias task successfully at least once in the treatment period.
- Corrected p-values that control for the Family-Wise Error Rates are included in brackets in Column 1 in both panels. A full description of our approach to multiple hypothesis corrections can be found in Appendix E.

A Online Only Supplementary Tables and Figures

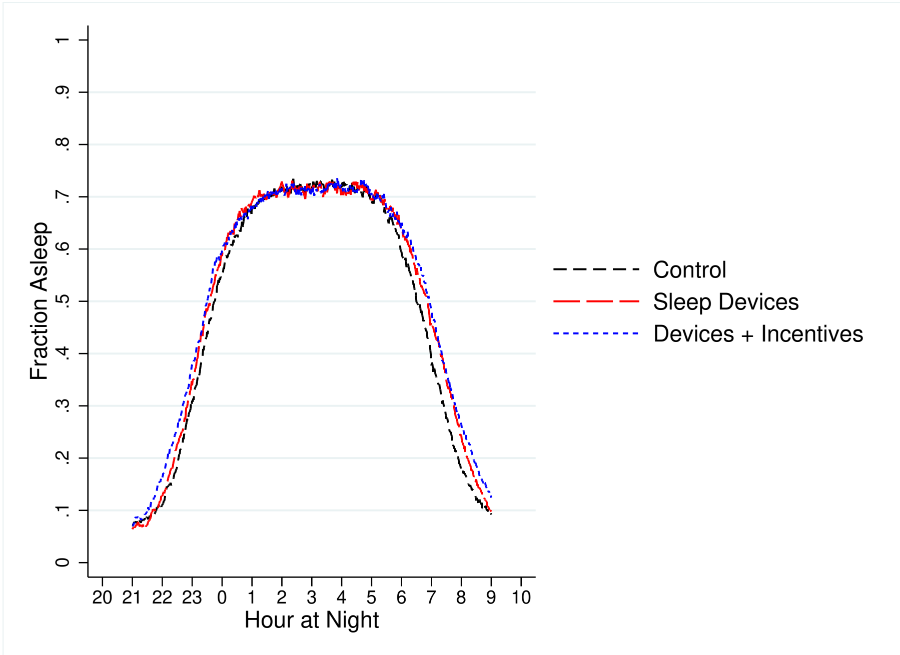
FIGURE A.I: Correlation Between Self-Reports and Actigraph Night Sleep Data



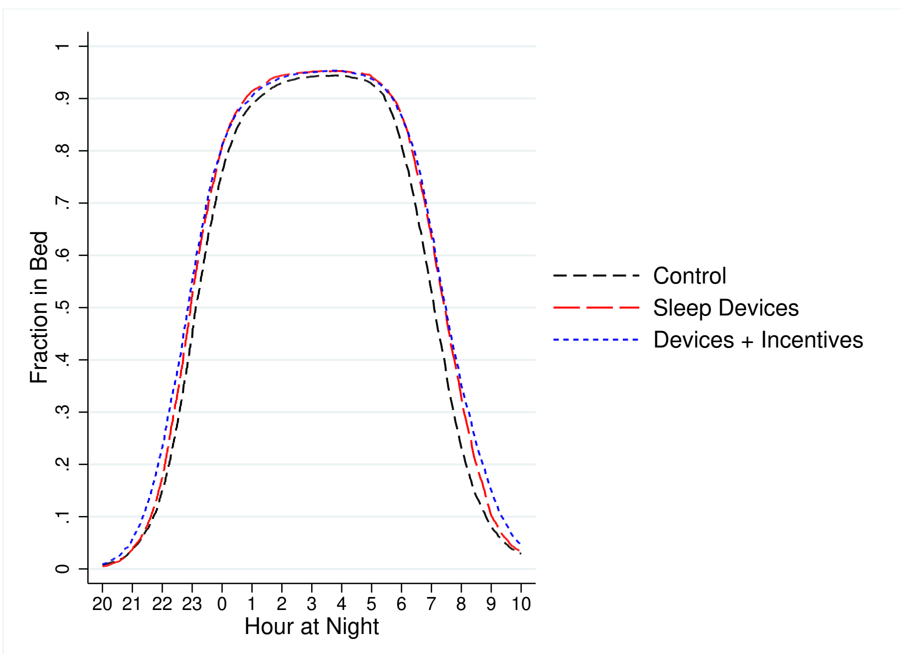
Notes: These figures show the correlations between self-reported and actigraph-measured time asleep (Panel A) and time in bed (Panel B) across all participants during the full study.

FIGURE A.II: Fraction of Individuals in Bed and Asleep by Hour of Night and by Treatment Group

(a) Fraction of Asleep by Hour of the Night

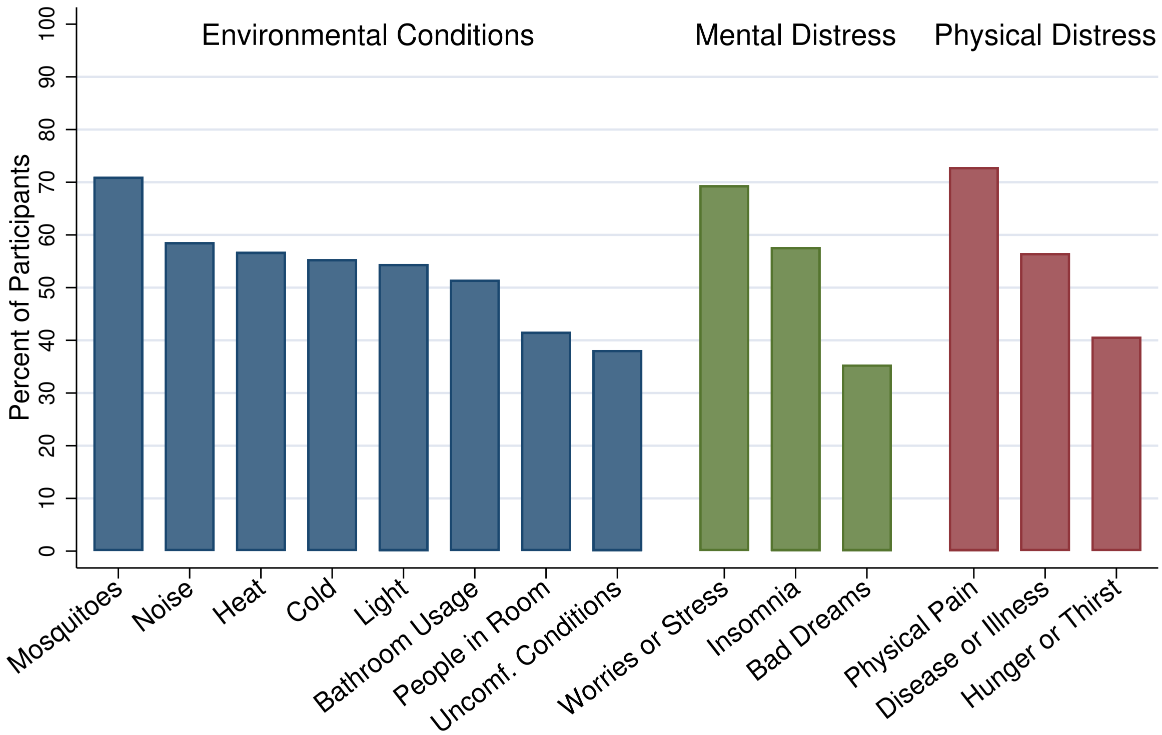


(b) Fraction in Bed by Hour of the Night



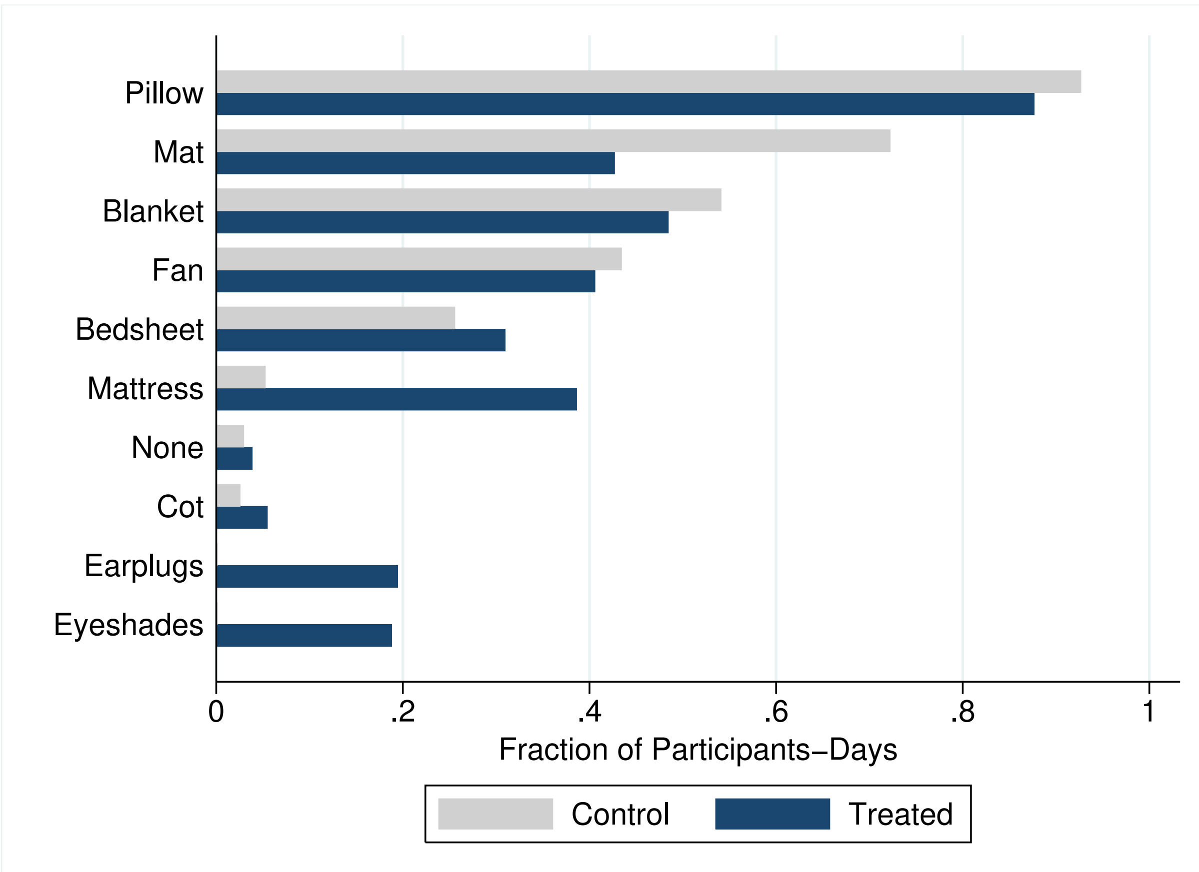
Notes: This figure shows the average fraction of participants asleep and in bed over the course of the night. In panel (a), the lines show the fraction of participants in each night sleep intervention group that are asleep at any time during the night, as measured by the actigraph. In panel (b) the lines show the fraction of participants in each night sleep intervention group that are in bed at any given time during the night, as measured by the actigraph.

FIGURE A.III: Factors Interfering with Study Participants' Sleep



Notes: This figure shows the fraction of participants who reported various factors impacting their sleep, including environmental conditions, mental distress, and physical distress. A participant is considered to have been affected by a awakening if they ever reported the factor bothering them.

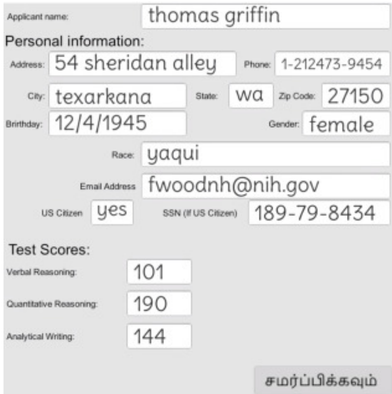
FIGURE A.IV: Sleep Aid Usage



Notes: This figure shows the fraction of participants who reported using each sleep aid provided in the study, divided between the Control group and the Treated group (which pools both night sleep interventions). Control group participants have positive values as some participants had these devices in their homes before entering the study.

FIGURE A.V: Data-Entry Interface with Salient and Non-Salient Piece Rates

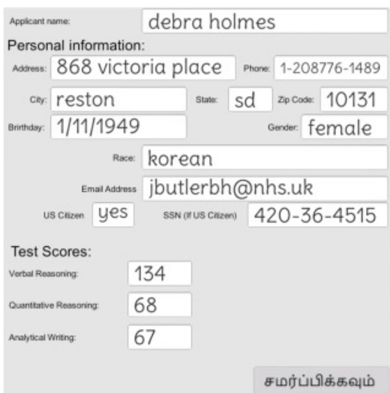
(a) Left side



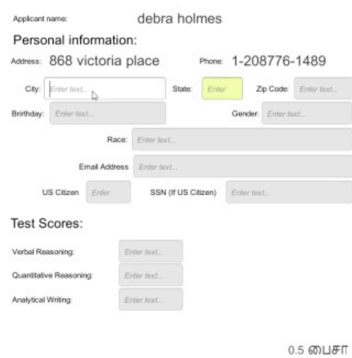
(b) Right side (salient, low)

(c) Right side (salient, high)

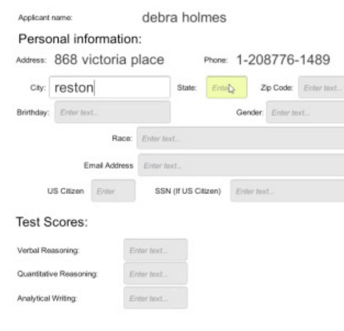
(d) Left side



(e) Right side (non-salient, beginning)



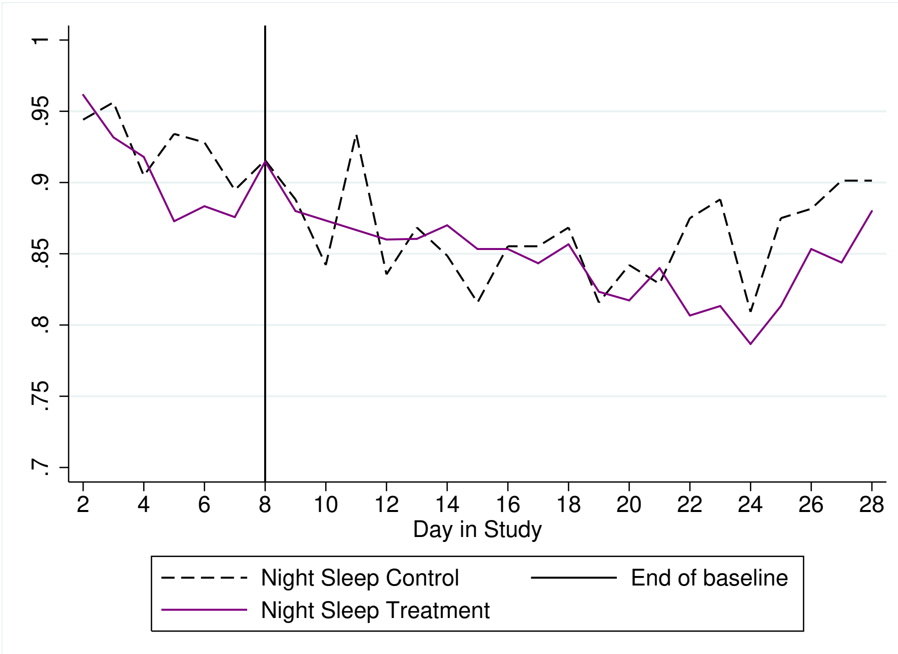
(f) Right side (non-salient, remainder)



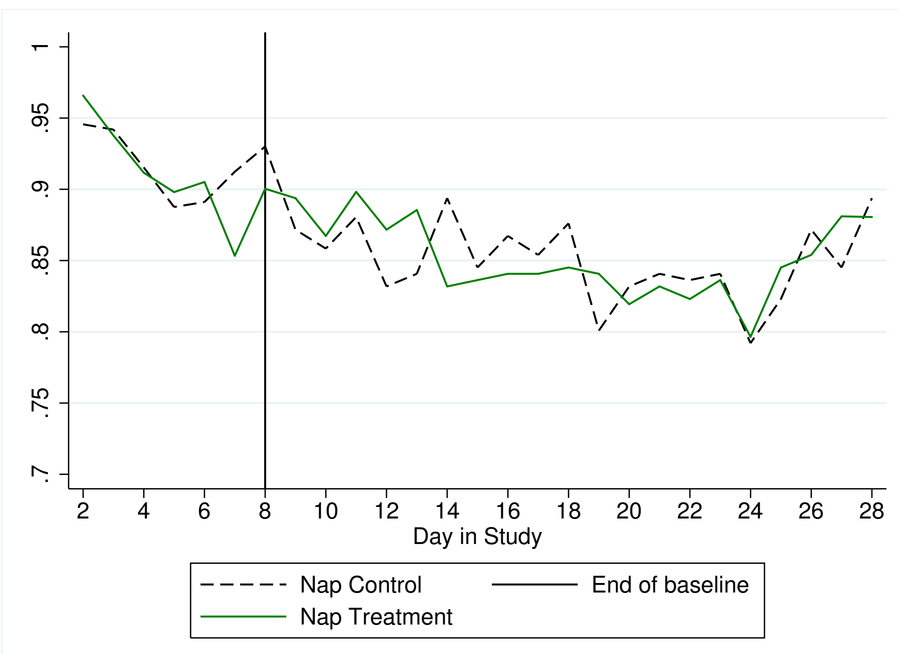
Notes: This figure shows screen shots of the data-entry task interface used by participants. Panels (a) and (d) show the left side of the screen, which contains the data to be transcribed by individuals. The remaining panels show versions of the right side of the screen, where the data is to be entered. Panels (b) and (c) show right side of the screen under salient incentives, once for low incentives (panel (b)) and once for high incentives (panel (c)). Panels (e) and (f) show the right side of the screen under non-salient incentives. Panel (e) is taken from the very beginning of a 30-minute period when individuals can see the (non-colored) piece rate for 15 seconds. Panel (f) is taken from the remaining part of the 30-minute period when the piece rate is no longer visible.

FIGURE A.VI: Attendance by Day of Study and Treatment Group

(a) Fraction of Participants Present by Night Sleep Intervention Groups



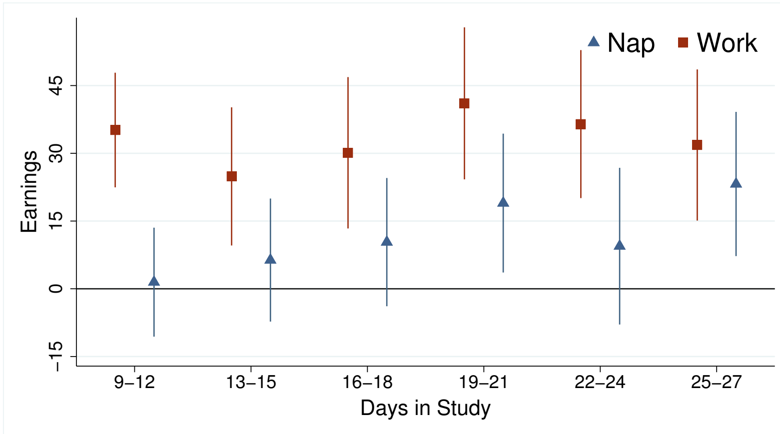
(b) Fraction of Participants Present by Nap Intervention Groups



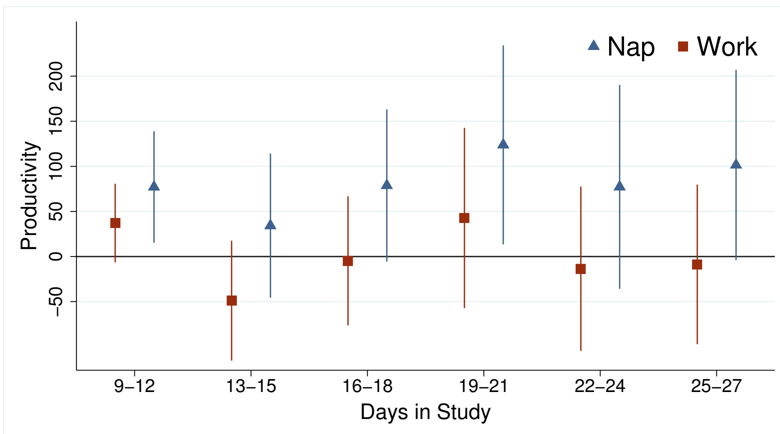
Notes: This figure shows the fraction of participants present during each day in the study by treatment group. In panel (a), the solid purple line shows the fraction of participants in the pooled night sleep intervention group who were present in the study office each day, while the dashed line does the same for night sleep control participants. In panel (b) the solid green line shows the fraction of participants in the nap intervention group who were present in the study office each day, while the dashed line does the same for nap control participants.

FIGURE A.VII: Comparison between Nap and Work Group Over Days

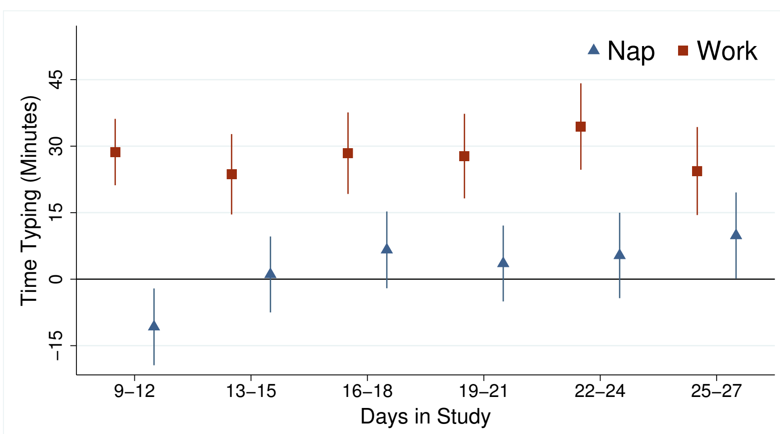
(a) Earnings



(b) Productivity

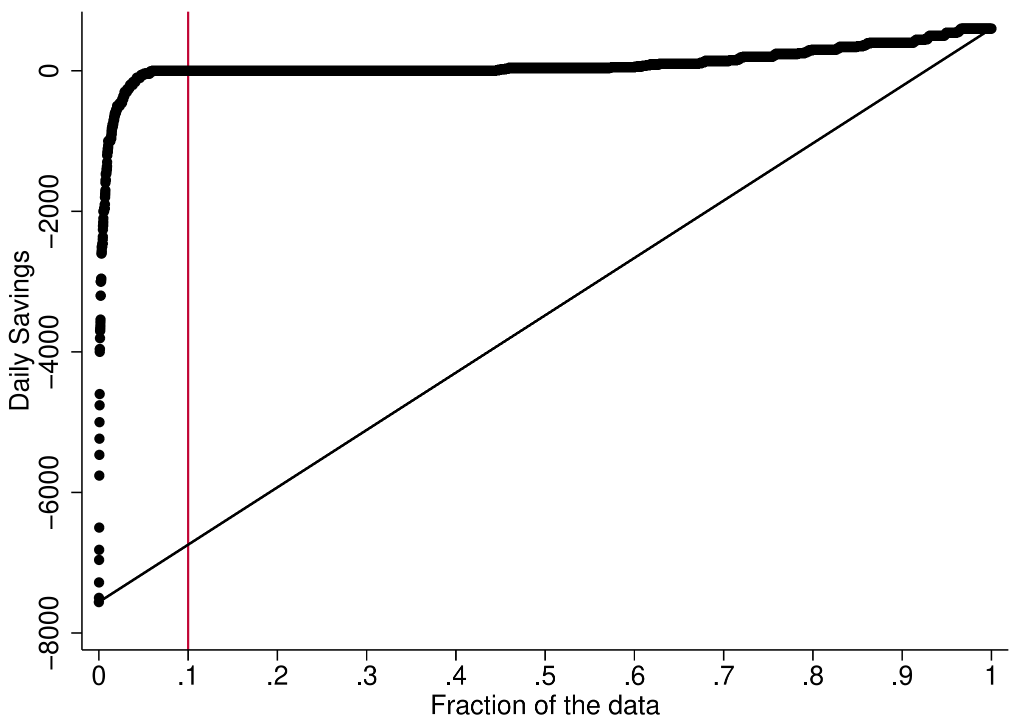


(c) Hours Typing



Notes: These figures plot regression coefficients of an outcome variable (productivity, hours typing and earnings, respectively) on the indicators of nap and work groups following specification (2). In this regression, the post-treatment period is grouped in 3-day bins to highlight the dynamics of the nap treatment.

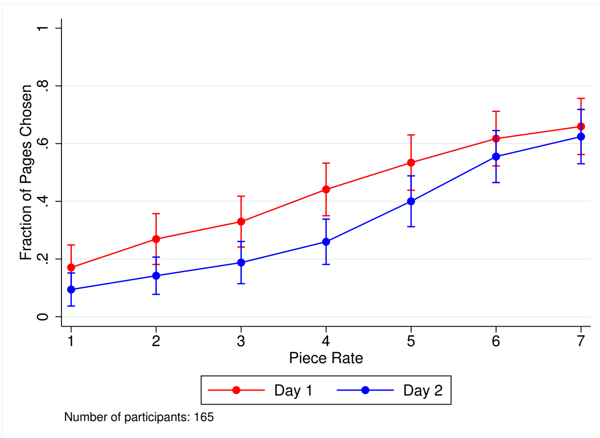
FIGURE A.VIII: Quantile Plot of Daily Net Savings



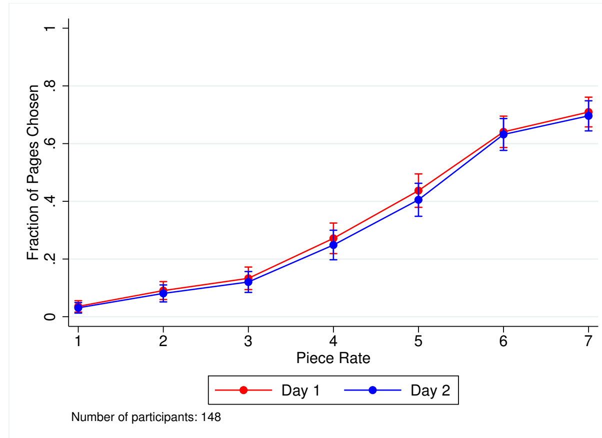
Notes: This figure shows the ordered values of daily net savings (difference between deposits and withdrawals) in black dots plotted against the quantiles of a theoretical uniform distribution, represented by the solid black line. The solid red line highlights the 5th percentile of the distribution, associated with a daily net savings of -50.

FIGURE A.IX: Present Bias Choices by Piece Rate

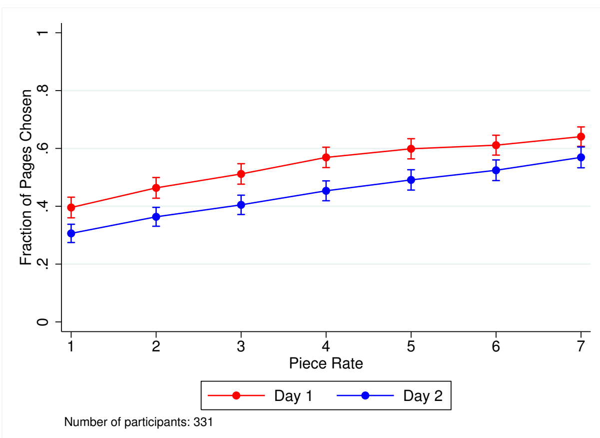
(a) Version 1 - Baseline Period



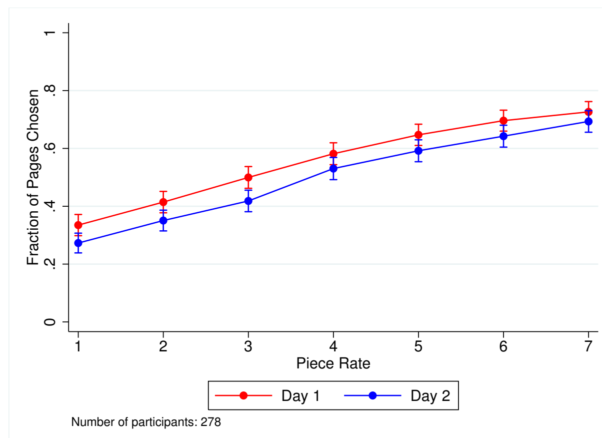
(b) Version 1 - Treatment Period



(c) Version 2 - Baseline Period



(d) Version 2 - Treatment Period



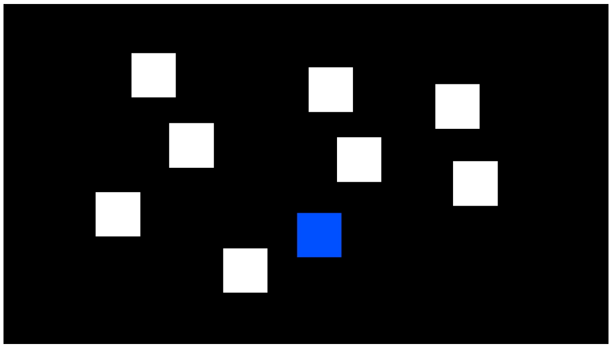
Notes: This figure shows fraction of pages chosen by participants under different piece rates. The figures in the top show the relation between average fraction of pages chosen and piece rate offered (1 is the lowest and 7 is the highest piece rate offered) in the first version of the Present Bias Experiment. The figures in the bottom row are analogous for version 2 of the task. Figures (a) and (c) show the relationship during the baseline period, while Figures (b) and (d) show the results in treatment period.

FIGURE A.X: Images for Cognitive Tests

(a) Psychomotor Vigilance Task (PVT)

(b) Hearts and Flowers

(c) Corsi Blocks



Notes: This figure shows examples of the three cognitive tests used in the study: Psychomotor Vigilance Task (PVT), Hearts and Flowers, and Corsi Blocks. In the PVT task, participants must click the screen when a number appears. In the Hearts and Flowers test, they must click the screen on the same side (for a heart) or opposite side (for a flower) when the image appears. In the Corsi test, a random sequence of blocks light up and the participant must then click on the blocks on the same order that they were highlighted. The number of blocks highlighted increases with each round.

FIGURE A.XI: Survey of Experts: Representative Prediction Page

Question 1: Work capacity - **number of correct entries per day**

The **control group** participants type on average **14,476 correct entries per day**. How many more (fewer) correct entries per day do the treated participants type on average? Please enter your prediction in the box below.

The number of correct entries *changes* by [ ] correct entries per day (You can insert either zero, positive, or negative numbers).

	maps to:	maps to:
hypothetical answer	**% Change**	**SD Change**
±145	±1.0	±0.01
±1,082	±7.5	±0.10
±1,448	±10.0	±0.13
±2,165	±15.0	±0.20
±5,412	±37.4	±0.50
±8,117	±56.1	±0.75
±10,823	±74.8	±1.00
±21,646	±149.6	±2.00

Place mouse here to see benchmark table.

Note: The values in the table below can be used as a reference point to guide your prediction. For example, the first row indicates that if the treated participants type more (less) 145 entries, this corresponds to a 1% increase (decrease) and a 0.01 standard deviation increase (decrease) in working time in comparison to the control group.

← →

Notes: This figure shows a typical page in the Expert Survey. Experts input their prediction by indicating the difference in outcome values between the treatment and the control. Survey participants could use the reference table at the bottom to guide their responses.

Table A.I: Survey of Experts: Summary Statistics

	All					General Economists					Behavioral Economists					Medical				
	Mean (1)	Median (2)	p25 (3)	p75 (4)	N (5)	Mean (6)	Median (7)	p25 (8)	p75 (9)	N (10)	Mean (11)	Median (12)	p25 (13)	p75 (14)	N (15)	Mean (16)	Median (17)	p25 (18)	p75 (19)	N (20)
Correct Entries	0.12	0.07	0.03	0.15	122	0.05	0.05	0.03	0.06	27	0.03	0.02	0.01	0.05	19	0.17	0.10	0.07	0.21	76
Hours Working	0.07	0.04	0.00	0.10	118	0.02	0.02	0.00	0.04	26	0.02	0.02	0.00	0.04	19	0.10	0.09	0.01	0.13	73
Savings	0.05	0.03	0.01	0.07	44	0.04	0.03	0.01	0.06	25	0.05	0.04	0.02	0.08	19	.	.	.	.	0
Attention (PVT)	0.12	0.08	-0.03	0.17	69	.	.	.	.	0	.	.	.	.	0	0.12	0.08	-0.03	0.17	69
Blood Pressure	-0.11	-0.08	-0.23	0.00	66	.	.	.	.	0	.	.	.	.	0	-0.11	-0.08	-0.23	0.00	66

Notes: This table describes survey responses from experts in economics and sleep science.

- Each row presents a different outcome. The values in the table are the intention-to-treat parameters predictions for each outcome divided by the control group's mean, which was provided for respondents in the survey.
- The statistics we consider are the average, median, 25^{th} percentile, 75^{th} percentile, and number of responses for each outcome.
- Correct Entries (row 1) refers to the number of daily correct characters in the data-entry task. The control mean provided in the survey form was 14,476.
- Hours Working (row 2) refer to the daily number of hours working in the typing task (excluding voluntary and scheduled pauses). The control mean provided in the survey form was 5.75 hours.
- Savings (row 3) refers to the daily amount of money (in Rupees) stored in the savings box during the experiment. The control group mean provided in the survey form was 144 Rupees.
- Attention (row 4) refers to an index pooling inverse response times (IRT) and minor lapses (ML) in the Psychomotor Vigilance Task (PVT), which differs from the rest of the paper (where false starts are also a component of PVT indices). Control group averages are 2.9 for inverse reaction time and 2.96 for minor lapses.
- Blood Pressure (row 5) refers to a variable that pools both systolic and diastolic blood pressure. The control group mean for systolic blood pressure was 15.6, while the mean for diastolic blood pressure was 11.

Table A.II: Balance Across Experimental Arms: Demographics and Baseline Sleep

	Night Sleep Treaments						Nap Treatments		
	Control	Devices	Incentives	$1 = 2$	$1 = 3$	$1 = (2 \bigcup 3)$	No Nap	Nap	$7 = 8$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Demographics									
Female	0.66	0.64	0.69	0.74	0.60	0.91	0.65	0.64	0.62
	(0.04)	(0.04)	(0.04)				(0.03)	(0.04)	
Age	35.84	35.28	33.72	0.50	0.01	0.06	34.94	35.28	0.97
	(0.62)	(0.58)	(0.56)				(0.46)	(0.58)	
Number of Children	1.42	1.35	1.29	0.54	0.29	0.34	1.30	1.35	0.29
	(0.09)	(0.08)	(0.09)				(0.07)	(0.08)	
Years of Education	10.35	10.00	10.20	0.29	0.65	0.39	10.34	10.00	0.26
	(0.23)	(0.24)	(0.23)				(0.19)	(0.24)	
Familiar with Computer	0.30	0.28	0.38	0.90	0.41	0.67	0.35	0.28	0.41
	(0.07)	(0.07)	(0.07)				(0.06)	(0.07)	
Unemployed	0.95	0.94	0.94	0.60	0.60	0.54	0.95	0.94	0.84
	(0.02)	(0.02)	(0.02)				(0.01)	(0.02)	
Panel B. Baseline Sleep									
Self-Reported Night Sleep (Hrs)	7.22	7.21	7.14	0.95	0.45	0.63	7.24	7.21	0.24
	(0.08)	(0.07)	(0.07)				(0.07)	(0.07)	
Actigraph Night Sleep (Hrs)	5.57	5.57	5.60	0.99	0.73	0.85	5.57	5.57	0.89
	(0.07)	(0.07)	(0.07)				(0.06)	(0.07)	
Actigraph Time in Bed (Hrs)	8.11	8.08	8.14	0.83	0.73	0.94	8.09	8.08	0.66
	(0.07)	(0.08)	(0.07)				(0.06)	(0.08)	
Sleep Efficiency	0.69	0.70	0.70	0.79	0.77	0.75	0.70	0.70	0.77
	(0.01)	(0.01)	(0.01)				(0.01)	(0.01)	
Number of Sleep Devices Owned	2.52	2.71	2.34	0.30	0.32	0.97	2.54	2.71	0.87
	(0.13)	(0.15)	(0.11)				(0.11)	(0.15)	
Number of Participants	152	150	150				226	226	

Notes: This table considers any underlying differences that may exist between the randomized experimental arms.

- Columns 1 to 3 show baseline means and standard errors by night sleep treatments. Columns 4 to 6 show p-values of t-tests between columns 1 vs. 2, 1 vs. 3, and 1 vs. 2 and 3.
- Columns 7 to 8 show baseline means and standard errors by nap treatment group. Column 9 shows the p-value for the t-test between no nap group and nap group.

Table A.III: Balance Across Experimental Arms: Health, Well-Being, Cognition, Work, and Savings

	Night Sleep Treaments						**Nap Treatments**		
	Control	Devices	Incentives	1 = 2	1 = 3	1 = (2 U 3)	No Nap	Nap	7 = 8
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel C. Health, Well-Being, Cognition									
Health Index	0.01	-0.01	0.00	0.66	0.89	0.74	-0.02	-0.01	0.29
	(0.04)	(0.04)	(0.04)				(0.03)	(0.04)	
Well-being	-0.01	0.04	0.02	0.42	0.64	0.46	0.02	0.04	0.90
	(0.04)	(0.05)	(0.04)				(0.04)	(0.05)	
Low Incentive PVT Pay (Rs.)	12.71	12.54	12.63	0.41	0.68	0.48	12.74	12.54	0.21
	(0.14)	(0.16)	(0.15)				(0.11)	(0.16)	
Low Incentive HF Pay (Rs.)	13.56	13.72	13.62	0.31	0.73	0.43	13.65	13.72	0.83
	(0.12)	(0.11)	(0.11)				(0.09)	(0.11)	
Low Incentive Corsi Pay (Rs.)	13.95	13.87	13.79	0.72	0.48	0.54	13.80	13.87	0.47
	(0.15)	(0.16)	(0.17)				(0.13)	(0.16)	
Panel D. Baseline Work and Savings									
Typing Time (Hrs)	4.49	4.55	4.40	0.57	0.39	0.86	4.48	4.55	0.99
	(0.05)	(0.11)	(0.05)				(0.04)	(0.11)	
Time in Office (Hrs)	7.95	7.92	7.86	0.67	0.25	0.36	7.92	7.92	0.70
	(0.06)	(0.06)	(0.06)				(0.05)	(0.06)	
Productivity	2373.51	2451.13	2468.20	0.67	0.61	0.59	2558.52	2451.13	0.09
	(127.65)	(141.07)	(122.99)				(118.05)	(141.07)	
Earnings	403.03	404.78	405.07	0.89	0.87	0.86	413.41	404.78	0.07
	(8.67)	(9.50)	(8.41)				(8.10)	(9.50)	
Attendance	0.94	0.93	0.92	0.14	0.05	0.05	0.93	0.93	0.25
	(0.01)	(0.01)	(0.01)				(0.01)	(0.01)	
Attendance 2	0.95	0.94	0.94	0.12	0.08	0.06	0.94	0.94	0.28
	(0.01)	(0.01)	(0.01)				(0.00)	(0.01)	
Savings (Rs.)	96.70	90.76	112.75	0.67	0.25	0.67	98.69	90.76	0.81
	(9.07)	(9.88)	(10.33)				(7.71)	(9.88)	
Prior Savings (Rs. 1000)	27.17	16.08	32.67	0.29	0.60	0.76	27.71	16.08	0.57
	(5.59)	(4.16)	(10.73)				(7.63)	(4.16)	
Joint Orthogonality Test				0.74	0.25	0.75		0.87	
Number of Participants	152	150	150				226	226	

Notes: This table considers any underlying differences that may exist between the experimental arms.

- Columns 1 to 3 show baseline means and standard errors by night sleep treatments. Columns 4 to 6 show p-values of t-tests between columns 1 vs. 2, 1 vs. 3, and 1 vs. 2 and 3.
- Columns 7 to 8 show baseline means and standard errors by nap treatment group. Column 9 shows the p-value for the t-test between no nap group and nap group.
- The Joint Orthogonality Test row refers to the F-test of a regression of the treatment dummy on all variables present in the balance table. This joint test provides an overall evaluation of the balance between treatments.

Table A.IV: Heterogeneous Treatment Effects on Night Sleep

	Night Sleep				
	(1)	(2)	(3)	(4)	(5)
Night Sleep Treat	0.433***	0.452***	0.467***	0.427***	0.445***
	(0.0749)	(0.0718)	(0.0731)	(0.0780)	(0.0720)
Above Median Sleep Length	-0.0533				
	(0.103)				
NS Treat × Above Median Length	0.0267				
	(0.0980)				
Above Median Sleep Quality		0.198**			
		(0.0883)			
NS Treat × Above Median Quality		-0.0188			
		(0.100)			
Below Median Awakenings			0.193**		
			(0.0869)		
NS Treat × Below Median Awakenings			-0.0350		
			(0.102)		
Above Median Efficiency				0.155*	
				(0.0852)	
NS Treat × Above Median Efficiency				0.0398	
				(0.101)	
Above Median Longest Streak					0.153*
					(0.0863)
NS Treat × Above Median Longest Streak					0.00307
					(0.0997)
Control Mean	5.096	5.152	5.176	5.183	5.168
Control SD	1.141	1.136	1.095	1.191	1.112
N	8428	8428	8428	8428	8428

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table presents heterogeneous treatment effects with respect to the effects of the treatments on quality.

- In all columns we run specification (2) with total night sleep time as the dependent variable. Controls include sex, age, and a dummy for the Nap Treatment.
- We separetely assess heterogeneous night sleep treatment effects with respect to: baseline average sleep length (column 1), the baseline sleep quality index (column 2), baseline average number of awakenings per hour (column 3), baseline average sleep efficiency (column 4), and baseline average longest sleep episode (column 5).
- The quality index is the average of standardized number of awakenings per hour, sleep efficiency and longest sleep episode.
- Standard errors clustered at the participant level.

Table A.V: Heterogeneous Night Sleep Treatment Effects on Sleep Quality

	Quality Index				
	(1)	(2)	(3)	(4)	(5)
Night Sleep Treat	0.0297	0.0684	0.0760*	0.0488	0.0303
	(0.0480)	(0.0441)	(0.0455)	(0.0460)	(0.0526)
Above Median Sleep Length	0.00570				0.00470
	(0.0564)				(0.0577)
NS Treat × Above Median Sleep Length	-0.00519				-0.00352
	(0.0642)				(0.0684)
Above Median Sleep Quality		0.0790			
		(0.0640)			
NS Treat × Above Median Quality		-0.0858			
		(0.0652)			
Below Median Awakenings			0.118*		
			(0.0614)		
NS Treat × Below Median Awakenings			-0.0961		
			(0.0654)		
Above Median Efficiency				0.0589	
				(0.0608)	
NS Treat × Above Median Efficiency				-0.0443	
				(0.0653)	
Above Median Longest Streak					-0.00349
					(0.0633)
NS Treat × Above Median Longest Streak					-0.00263
					(0.0694)
Control Mean	-0.249	-0.451	-0.434	-0.369	-0.418
Control SD	0.857	0.684	0.680	0.810	0.659
N	8392	8392	8392	8392	8392

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table examines potential heterogeneity in the treatment effects on the quality of night time sleep.

- In all columns we run specification (2) with the sleep quality standardized average index as the dependent variable. Controls include sex, age, and a dummy for the Nap Treatment.
- We separetely assess heterogeneous night sleep treatment effects with respect to: baseline average sleep length (column 1), the baseline sleep quality index (column 2), baseline average number of awakenings per hour (column 3), baseline average sleep efficiency (column 4), and baseline average longest sleep episode (column 5).
- The quality index is the average of standardized number of awakenings per hour, sleep efficiency, and longest sleep episode.
- Standard errors clustered at the participant level.

Table A.VI: Decomposing Impacts on Labor Supply and Productivity

	Labor Supply						**Productivity**		
	Minutes Typing (1)	Total Pause (2)	Voluntary Pause (3)	Minutes in Office (4)	Arrival Time (5)	Leave Time (6)	Productivity (7)	Speed (8)	Accuracy (9)
Night Sleep Treat vs. Control	-10.29***	0.97	1.65**	-8.94***	0.11***	-0.04	44.32	35.47	0.08**
	(3.05)	(1.13)	(0.87)	(2.66)	(0.03)	(0.03)	(39.29)	(40.47)	(0.04)
Nap Treat vs. Break	1.55	3.01***	1.53**	3.41	-0.01	0.05	82.29**	80.10**	0.07*
	(3.06)	(1.05)	(0.77)	(2.66)	(0.03)	(0.03)	(36.49)	(37.37)	(0.04)
Nap Treat vs. No Break	-26.10***	29.07***	-2.61***	3.26	0.01	0.06**	74.47**	74.73**	0.06
	(3.14)	(1.26)	(1.01)	(2.56)	(0.03)	(0.03)	(36.90)	(37.58)	(0.04)
Control Mean	244.04	113.48	10.90	400.32	10.52	18.34	3390.86	3587.47	99.14
Control SD	126.13	55.89	18.69	178.97	0.70	0.97	1839.11	1827.30	0.76
N	6992	6992	6992	6989	6989	6989	6992	6992	6992
Participants	451	451	451	451	451	451	451	451	451

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table considers the treatment effect of the night sleep and the nap interventions on labor supply and productivity outcomes.

- The outcome variables in columns 1 to 6 are various measures of labor supply. Column 1 considers minutes spent typing, column 2 considers total pauses, and column 3 considers voluntary pauses (excluding mandatory pause for participants randomized to stop work instead of napping). Column (4) considers total minutes in office and columns 5 and 6 consider office arrival and departure times (in hours), respectively.
- The outcome variables in columns 7, 8, and 9 are measures of productivity: productivity (output/hour), typing speed, and typing accuracy, respectively.
- Each column shows the OLS estimates of equation (2), controlling for average baseline values (ANCOVA), age, sex, "regular" study day, fraction of high piece-rate sessions, and day in study and date fixed effects. Standard errors are clustered at the participant level.

Table A.VII: Treatment Effect on Sleep Quality

	Efficiency	**Awakenings/Hour**	**Longest Sleep Episode**	**Shortened PSQI**	**Karolinska**
	(1)	(2)	(3)	(4)	(5)
Night Sleep Treat	-0.000751	-0.0256	1.640*	-0.336***	-0.0781
	(0.00421)	(0.0775)	(0.931)	(0.111)	(0.0541)
Nap Treat	0.00273	0.00795	-1.308	0.216**	-0.00469
	(0.00395)	(0.0714)	(0.904)	(0.101)	(0.0539)
Baseline	0.744***	0.780***	0.698***	0.248***	0.144***
	(0.0317)	(0.0257)	(0.0427)	(0.0789)	(0.0407)
Control Mean	0.698	5.887	54.74	2.071	-2.960
Control SD	0.113	2.206	25.74	1.535	1.698
N	8426	8426	8392	399	1800
Participants	451	451	449	399	442

Note: This table presents first-stage regressions regarding quality measures. All regressions include controls for age quartiles and female dummy along with the average baseline value of the dependent variable.

- Efficiency is defined as time asleep divided by time in bed. Both are measured by actigraph.
- Awakenings is defined as a switch from "asleep" to "awake". The actigraph records, minute-by-minute, whether there is enough wrist movement detected ("awake") or not ("asleep").
- Longest Sleep Episode is defined as the duration (in minutes) of the participant's longest episode of uninterrupted sleep in a night.
- Shortened PSQI corresponds to an adapted version of the Pittsburgh Sleep Quality Index. This measure is based on one-time questions conducted in the endline survey. We make use of questions 6, 7, and 8 of the original PSQI (see Buysse et al. (1989)) to create components 1, 6, and 7 (component 7 uses only information from question 8). We also adapt component 5. Instead of using the original disturbance questions (questions 5b-5j in Buysse et al. (1989)), we use our own, which asks whether participants have trouble sleeping because of the presence of heat, light, mosquitoes, stress, bad dreams, diseases, insomnia, physical pain, flooded sleeping areas, noise, hungriness/thirstiness, use of bathroom and/or presence of child/baby. We then code component 5 in accordance to the number of factors reported, with 3 or more factors being assigned value 3 (the maximum score). We then sum our (adapted) components 1, 5, 6, and 7 and standardized the measure according to the control mean and standard deviation.
- The Karolinska variable asks the participant to "describe your sleepiness durnig the previous five minutes" and is asked on some days during the Daily Survey. The nine possible answers range from 0 ("Extremely Alert") to 8 ("Extremely sleepy, fighting sleep"). We then standardize this variable.
- Both the Shortened PSQI and Karolinska measures are multiplied by minus 1 so that higher values indicate more desirable outcomes.
- Standard errors are clustered at the participant level.

Table A.VIII: Correlation Between Sleep Quality Measures

	Sleep Efficiency (1)	**Longest Sleep Episode** (2)	**Awakenings per Hour** (3)	**Awakenings per Hour (5-Min)** (4)
Sleep Efficiency	1.00	-	-	-
Longest Sleep Episode	0.67	1.00	-	-
Awakenings per Hour	0.82	0.78	1.00	-
Awakenings per Hour (5-Min)	0.87	0.64	0.81	1.00

Note: This table presents simple Pearson correlation measures between our sleep quality measures.

- Sleep efficiency is defined as time asleep divided by time in bed. Both are measured by actigraph.
- Awakenings per Hour is defined as a switch from "asleep" to "awake". The actigraph records, minute-by-minute, whether there is enough wrist movement detected ("awake") or not ("asleep"). We flip the sign of the variable so that a positive relationship stands for fewer awakenings (a “better" outcome).
- Awakenings per Hour (5-Min Disruptions) is defined as a switch from "asleep" to "awake" in which participants remain awake for at least 5 minutes. We also flip the sign so that a positive relationship stands for fewer awakenings (a “better" outcome).
- Longest Sleep Episode is defined as the duration (in minutes) of the participant’s longest episode of uninterrupted sleep in a night.

Table A.IX: Heterogeneous Night Sleep Treatment Effects on Productivity

	Productivity				
	(1)	(2)	(3)	(4)	(5)
Night Sleep Treat	107.3**	-36.93	-66.65	14.64	-20.03
	(50.52)	(52.31)	(51.04)	(51.80)	(53.06)
Above Median Sleep Length	93.56				
	(68.47)				
NS Treat × Above Median Sleep Length	-122.2				
	(80.19)				
Above Median Sleep Quality		-139.4**			
		(68.18)			
NS Treat × Above Median Quality		168.8**			
		(78.81)			
Below Median Awakenings			-203.7***		
			(65.67)		
NS Treat × Below Median Awakenings			219.6***		
			(77.30)		
Above Median Sleep Efficiency				21.97	
				(68.69)	
NS Treat × Above Median Efficiency				61.47	
				(81.26)	
Above Median Longest Streak					-134.7**
					(68.31)
NS Treat × Above Median Longest Streak					131.9
					(80.28)
Control Mean	3253.2	3496.5	3549.8	3431.3	3350.2
Control SD	1930.4	2053.2	2078.6	2049.1	2029.4
N	7350	7350	7350	7350	7350

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table presents heterogeneous treatment effects with respect to the effects of the treatments on productivity.

- In all columns we run specification (2) with productivity as the dependent variable. Controls include sex, age, a dummy for long days, fraction of high incentive sessions, and a dummy for the Nap Treatment.
- We separately assess heterogeneous night sleep treatment effects with respect to: baseline average sleep length (column 1), the baseline sleep quality index (column 2), baseline average number of awakenings per hour (column 3), baseline average sleep efficiency (column 4), and baseline average longest sleep episode (column 5).
- The quality index is the average of standardized number of awakenings per hour, sleep efficiency, and longest sleep episode.
- Standard errors clustered at the participant level.

Table A.X: Heterogeneous Night Sleep Treatment Effects on Earnings

	Earnings				
	(1)	(2)	(3)	(4)	(5)
Night Sleep Treat	1.927	-21.04*	-32.79***	-11.98	-18.14*
	(10.09)	(10.78)	(11.70)	(10.21)	(10.33)
Above Median Sleep Length	16.41				
	(12.98)				
NS Treat × Above Median Sleep Length	-29.07*				
	(15.26)				
Above Median Sleep Quality		-0.223			
		(12.92)			
NS Treat × Above Median Quality		16.04			
		(15.37)			
Below Median Awakenings			-16.71		
			(13.24)		
NS Treat × Below Median Awakenings			39.36**		
			(15.50)		
Above Median Sleep Efficiency				17.15	
				(13.05)	
NS Treat × Above Median Efficiency				-1.981	
				(15.19)	
Above Median Longest Streak					-8.527
					(13.05)
NS Treat × Above Median Longest Streak					10.65
					(15.57)
Control Mean	262.2	277.0	293.2	267.2	284.5
Control SD	194.5	214.1	221.6	219.3	206.2
N	8647	8647	8647	8647	8647

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table examines heterogeneous treatment effects on earnings.

- In all columns we run specification (2) with productivity as the dependent variable. Controls include sex, age, a dummy for long days, fraction of high incentive sessions, and a dummy for the Nap Treatment.
- We separetely assess heterogeneous night sleep treatment effects with respect to: baseline average sleep length (column 1), the baseline sleep quality index (column 2), baseline average number of awakenings per hour (column 3), baseline average sleep efficiency (column 4), and baseline average longest sleep episode (column 5).
- The quality index is the average of standardized number of awakenings per hour, sleep efficiency, and longest sleep episode.
- Standard errors clustered at the participant level.

Table A.XI: Heterogeneous Night Sleep Treatment Effects on Hours Typing

	Hours Typing				
	(1)	(2)	(3)	(4)	(5)
Night Sleep Treat	-0.123	-0.214	-0.443**	-0.162	-0.264
	(0.150)	(0.161)	(0.171)	(0.166)	(0.161)
Above Median Sleep Length	0.0579				
	(0.182)				
NS Treat × Above Median Sleep Length	-0.192				
	(0.213)				
Above Median Sleep Quality		0.0789			
		(0.189)			
NS Treat × Above Median Quality		-0.0174			
		(0.217)			
Below Median Awakenings			-0.151		
			(0.196)		
NS Treat × Below Median Awakenings			0.438*		
			(0.226)		
Above Median Sleep Efficiency				0.224	
				(0.191)	
NS Treat × Above Median Efficiency				-0.119	
				(0.219)	
Above Median Longest Streak					-0.0770
					(0.190)
NS Treat × Above Median Longest Streak					0.0875
					(0.220)
Control Mean	4.057	4.059	4.235	3.924	4.287
Control SD	2.116	2.136	2.137	2.254	2.025
N	8647	8647	8647	8647	8647

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table examines heterogeneous treatment effects on hours typing.

- In all columns we run specification (2) with productivity as the dependent variable. Controls include sex, age, a dummy for long days, fraction of high incentive sessions, and a dummy for the Nap Treatment.
- We separetely assess heterogeneous night sleep treatment effects with respect to: baseline average sleep length (column 1), the baseline sleep quality index (column 2), baseline average number of awakenings per hour (column 3), baseline average sleep efficiency (column 4), and baseline average longest sleep episode (column 5).
- The quality index is the average of standardized number of awakenings per hour, sleep efficiency, and longest sleep episode.
- Standard errors clustered at the participant level.

Table A.XII: Treatment Effects on Default Pass-Through

	Follow Default		**Default Pass-Through**	
	(1)	(2)	(3)	(4)
Default	0.24***	0.24***	0.42***	0.44**
	(0.01)	(0.03)	(0.09)	(0.19)
Night Sleep Treat × Default		-0.03		-0.03
		(0.03)		(0.21)
				[1.000]
Nap Treat × Default		0.04		0.00
		(0.03)		(0.20)
				[1.000]
Control Mean	0.01	0.01	120.44	120.44
Control SD	0.08	0.08	174.62	174.62
N	7280	7280	7280	7280
Participants	452	452	452	452

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Note: This table considers the treatment effect of the night sleep and the nap interventions on participants' propensity to follow the default and on default pass-through.

- The dependent variable in columns 1 and 2 is an indicator of whether the participant saved exactly Rs. 40, the default amount. Columns 3 and 4 capture daily deposits.
- Columns 1 and 2 show the OLS estimates of equation (5), whereas columns 3 an 4 show the OLS estimates of equation (4). At all four specifications we control for the participant's average baseline outcome (ANCOVA), age, sex, daily piece rate, interest rate, maximum payment from cognitive tasks and randomized piece rate for the present bias task.
- For columns 1 and 2, row 1 is the estimated coefficient associated with a dummy equal to one if the participant was randomized to the default condition on that day. In columns 3 and 4, row 1 shows the estimates for the default pass-through effect on daily deposits (i.e. what fraction of the amount defaulted into savings was passed through to the savings account). Rows 2 and 3 interact each of these variables with the night sleep and nap treatments, respectively.
- In columns 1 and 2, the control mean is the fraction of participants in the control group in non-default days that saved exactly Rs. 40. For columns 3 and 4, it is the average daily deposit in non-default days in the control group.
- Corrected p-values that control for the Family-Wise Error Rates are included in brackets in Column 4. A full description of our approach to multiple hypothesis corrections can be found in Appendix E.
- Standard errors are clustered at the participant level.

Table A.XIII: Treatment Effects on Inhibitory Control and Memory

	Inhibitory Control			**Memory**
	Payment	Frac. Correct	Avg. Reaction	Payment
Night Sleep Treat	0.0418	0.0751	0.0010	0.0143
	(0.0467)	(0.0654)	(0.0527)	(0.0460)
Nap Treat	0.0451	-0.0557	0.1051**	-0.0159
	(0.0430)	(0.0588)	(0.0487)	(0.0444)
Baseline	0.681***	0.604***	0.589***	0.671***
	(0.0382)	(0.0640)	(0.0404)	(0.0307)
Control Mean	14.98	0.891	551.9	14.64
Control SD	1.37	0.097	60.2	2.67
N	3554	3554	3554	3506
Participants	449	449	449	449

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table considers the treatment effect of the night sleep and nap interventions on inhibitory control and memory.

- All columns show the OLS estimates of equation (2), controlling for baseline values (ANCOVA), age, sex, whether participants faced high or low incentives for the task (which varied randomly within-participant each day), and day in study and date fixed effects.
- All variables are standardized by the control group's average and standard deviation, with signs flipped as needed such that higher outcomes indicate more desirable outcomes.
- The outcomes in columns 1-3 are all related to inhibitory control, measured by the Hearts and Flowers task. The outcome variable in Column 1 is the payment participants earn for completing the H&F task, where the payment is a weighted average of the fraction of correct entries and reaction time. Columns 2 and 3 break apart performance, respectively, by the fraction of correct entries, out of 40, and average reaction time.
- The outcome variable in column 4 is the payment participants earn for completing the Corsi blocks task, which measures working memory. Payment depends on the maximum number of blocks they can recall within the task.
- Standard errors are clustered at the participant level.

Table A.XIV: Goodness of Fit for Structural Present Bias Models

	Cost Function (1)	Censoring (2)	Alpha (3)	N (4)	N Fail (5)	NLL (6)	Avg (7)	Sd (8)	Min (9)	Max (10)	P10 (11)	P25 (12)	P50 (13)	P75 (14)	P90 (15)
All periods	Power	Yes	Yes	443	26	8593.06	18.47	363.53	0.00	7651.51	0.51	0.74	0.92	1.00	1.12
	Power	Yes	No	443	9	10354.10	3480.41	7.32E+04	0.00	1.54E+06	0.29	0.61	0.85	1.02	1.26
	Power	No	Yes	443	19	8570.95	0.91	0.31	0.16	2.82	0.56	0.78	0.93	1.00	1.14
	Power	No	No	443	5	10589.73	1.01	1.58	0.01	21.61	0.35	0.63	0.86	1.02	1.24
	Exp	Yes	Yes	443	53	49247.10	1.39	4.46	0.04	60.85	0.26	0.51	0.79	1.04	1.53
	Exp	Yes	No	443	51	50478.62	13.65	234.64	0.00	4930.92	0.27	0.52	0.78	1.13	1.60
	Exp	No	Yes	443	62	66734.72	1.05	2.95	0.04	55.37	0.24	0.48	0.78	1.00	1.29
	Exp	No	No	443	46	68750.85	1.87	16.81	0.00	346.84	0.24	0.49	0.77	1.02	1.41
Treatment Period	Power	Yes	Yes	398	46	718.16	1.20	4.97	0.00	99.66	0.62	0.84	0.98	1.03	1.17
	Power	Yes	No	398	27	3303.24	2.03	20.90	0.00	417.55	0.43	0.75	0.95	1.07	1.32
	Power	No	Yes	398	35	1406.77	0.96	0.31	0.06	3.68	0.64	0.85	0.98	1.02	1.16
	Power	No	No	398	15	4384.23	0.99	0.95	0.00	17.21	0.46	0.77	0.95	1.07	1.29
	Exp	Yes	Yes	398	79	25073.50	5.71	57.30	0.03	964.75	0.30	0.57	0.92	1.07	1.72
	Exp	Yes	No	398	65	26787.38	1940.45	3.76E+04	0.00	7.49E+05	0.36	0.61	0.93	1.17	1.79
	Exp	No	Yes	398	84	38183.22	1.39	4.13	0.03	46.57	0.30	0.57	0.91	1.03	1.41
	Exp	No	No	398	55	40020.84	1.99	19.44	0.00	386.61	0.35	0.60	0.88	1.07	1.45
Baseline Period	Power	Yes	Yes	430	74	-4685.44	35.41	426.88	0.00	7651.51	0.27	0.54	0.86	1.03	1.36
	Power	Yes	No	430	25	-148.29	1.36E+04	1.46E+05	0.00	2.01E+06	0.10	0.33	0.76	1.16	1.94
	Power	No	Yes	430	68	-5089.91	1.14	3.56	0.03	59.85	0.33	0.60	0.87	1.02	1.31
	Power	No	No	430	14	-95.16	1688.41	3.02E+04	0.00	6.19E+05	0.12	0.38	0.77	1.14	1.75
	Exp	Yes	Yes	430	98	13157.07	25.92	281.53	0.02	5117.04	0.16	0.37	0.78	1.13	2.51
	Exp	Yes	No	430	64	17277.05	6158.21	5.90E+04	0.00	9.41E+05	0.10	0.32	0.76	1.23	3.23
	Exp	No	Yes	430	81	17437.87	13.37	141.02	0.01	2504.40	0.13	0.34	0.73	1.06	1.77
	Exp	No	No	430	62	21829.83	7739.36	8.49E+04	0.00	1.27E+06	0.09	0.29	0.67	1.21	2.47

Notes: This table shows the model fit and key statistics of the distribution of present bias for different parametrizations of the participant's utility function in the structural estimation.

- The first panel pools data from the entire experiment, while the second and third panels use only data from the treatment and from the baseline period, respectively.
- The first 3 columns indicate, respectively, (1) whether the cost function is a power or an exponential function; (2) Whether we use a normal-Tobit model for censoring top and bottom observation; (iii) Whether we allow the parameter α to be different than zero.
- Columns 4 and 5 show, respectively (i) the number of participants who successfully completed at least one round of the Present Bias experiment; (ii) the number of participants for whom the structural estimation algorithm does not converge.
- Column 6 is the *Negative* Log-Likelihood (NLL) of the model summing across all participants. Smaller values indicate better fit.
- Columns 7-15 show key statistics of the distribution of individual-level present bias estimated from the model with specifications described in columns 1-3. PXX stands for percentile XX of the distribution.

Table A.XV: Present Bias - Experimental Integrity

	Night Sleep Treatments			Nap Treatments		
	Control	Night Sleep	P-Value	Control	Nap	P-Value
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Study						
Individual-level PB estimate	0.77	0.78	0.74	0.78	0.77	0.82
Make decisions but do not work	0.05	0.05	0.86	0.05	0.04	0.38
Work on a different date than assigned	0.11	0.14	0.45	0.13	0.13	0.94
Panel B: Version 1						
Individual-level PB estimate	0.88	0.93	0.36	0.90	0.92	0.64
Make decisions but do not work	0.04	0.04	0.83	0.05	0.03	0.40
Work on a different date than assigned	0.08	0.09	0.88	0.07	0.10	0.21
Panel C: Version 2						
Individual-level PB estimate	0.71	0.71	0.98	0.73	0.69	0.52
Make decisions but do not work	0.05	0.05	0.92	0.06	0.04	0.56
Work on a different date than assigned	0.13	0.16	0.47	0.16	0.14	0.69

Notes: This table provides information on the experimental integrity of the present bias task.

- Panel A shows all 452 observations, while Panel B focus on participants that did the first version of the Present Bias Experiment and Panel C on the participants who did the second version of the Present Bias Experiment.
- *Individual-level PB estimate* (row 1) is a dummy indicating whether we can estimate the individual-level present bias parameter in our preferred specification using only data from the treatment period. There are two reasons why we might not be able to estimate present bias in the individual-level: the participants did not complete the second day of the task or the structural estimator did not converge for the participant.
- *Make decisions but do not work* (row 2) is a dummy indicating whether the participant makes the work decisions but does not complete the work. Participants that did not complete either rounds of choice are excluded.
- *Work on a different date than assigned* (row 3) is a dummy indicating whether the date the participants completed the task ("work date") is different than the date they were assigned to complete when they made the first round of choices. Most of the time this happens because participants were absent from the office in the date they were supposed to complete the task.
- Columns 1, 2, 5, and 6 show the average of the variables in column 1 for control (night sleep), night sleep, control (nap), and nap groups. Column 3 and 6 show the p-value of the difference between control and treatment groups.

Table A.XVI: Relation between Present Bias (β) and Behaviors Involving Time Preferences

	Daily Deposits		Lateness		Voluntary Pauses		Night Sleep	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Beta Structural	39.66*	38.81*	-6.485	-7.360*	0.662	1.438	0.0724	0.0985
	(21.11)	(21.66)	(4.474)	(4.421)	(2.243)	(2.026)	(0.154)	(0.154)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	127.1	127.1	42.37	42.37	15.44	15.44	5.603	5.603
Control SD	121.3	121.3	25.94	25.94	11.23	11.23	0.824	0.824
Observations	351	351	351	351	351	351	351	351

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table reports correlation measures between the present bias coefficient (β) and participants' behavior.

- The independent variable of interest is the present bias measure β, estimated via the benchmark structural estimation method, which excludes participants for whom the maximization problem in the structural estimation does not converge.
- The dependent variables are: daily deposits (from the savings task), lateness (how long after the office opening the participant arrives), voluntary pauses (total daily length of voluntary pauses from the typing task), and night sleep (actigraph measured). All the dependent variables are study long averages (including the baseline period).
- Even columns shows the OLS estimates when controlling for participant's age and sex.

Table A.XVII: Treatment Effects on Risk and Social Preferences

	Risk Preferences				**Social Preferences**						
	Indices		**Components**		**Indices**		**Components**				
	Anderson	Average	Risk Aversion	Loss Aversion	Anderson	Average	Dictator Send	Ultimatum Send	Trust Send	Ultimatum Receive	Trust Send Back
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Night Sleep Treat	-0.038	-0.037	-0.111	0.007	-0.035	-0.041	-0.045	0.009	-0.135	-0.066	-0.049
	(0.081)	(0.081)	(0.099)	(0.096)	(0.056)	(0.061)	(0.102)	(0.097)	(0.102)	(0.116)	(0.125)
		[1.000]				[0.999]					
Nap Treat	0.073	0.074	-0.009	0.086	0.038	0.054	0.177*	0.005	0.055	0.063	0.121
	(0.076)	(0.076)	(0.093)	(0.091)	(0.053)	(0.057)	(0.095)	(0.092)	(0.096)	(0.113)	(0.117)
		[0.996]				[0.996]					
Amount Received										0.966***	1.290***
										(0.060)	(0.047)
Baseline	0.430***	0.461***	0.290***	0.376***	0.339***	0.418***	0.224***	0.203***	0.385***	0.492***	0.388***
	(0.050)	(0.053)	(0.051)	(0.049)	(0.060)	(0.061)	(0.050)	(0.050)	(0.057)	(0.086)	(0.079)
N	415	415	383	403	415	415	415	415	415	3465	2629
Participants	415	415	383	403	415	415	415	415	415	315	239

Notes: This table considers the treatment effect of the night sleep and nap interventions on risk and social preferences.

- All variables are standardized by the control group's average and standard deviation, with signs flipped when needed such that higher outcomes indicate lower risk preferences or more pro-social preferences.
- The dependent variables are separated into two panels: risk preferences components and social preferences components. Risk preferences components include the point at which the participant switched from the risky to safe choice in the risk aversion game (column 3) and the point at which the participant switched from the risky to safe choice in the loss aversion game (column 4). Social preferences components include the amount of money the sender sent in the dictator game (column 7), the amount of money the sender sent in the ultimatum game (column 8), the amount of money the sender sent in the trust game (column 9), whether the recipient accepted the sender's offer in the ultimatum game (column 10) and the amount of money the recipient sent back to the sender in the trust game (column 11).
- Columns 1 and 2 are weighted averages of the two standardized risk and loss aversion outcomes outcomes. Column 1 averages the outcomes optimally accounting for correlation across measures (Anderson, 2008), while Column 2 is a simple unweighted average of the standardized outcomes. Similarly, Columns 5 and 6 are the weighted averages of the five standardized social preferences outcomes. Column 5 averages the outcomes optimally accounting for correlation across measures, while Column 6 is a simple unweighted average.
- We include all observations, even those with non-monotonic decisions, when calculating the indices, while non-monotonic observations are excluded in the component regressions. We also take the average of recipients' choices across different amounts of money they received from senders when calculating indices, while we separate different recipients' choices in component regressions.
- Each column shows the OLS estimates of an equation similar to (2). In the regression, we pooled the two night sleep treatments and controlled for age, sex, and baseline measures of the dependant variable (ANCOVA) when available. When the dependent variable is an index, we control for the index at baseline. When the outcomes are recipients' choices, we control for the amount of money they received from senders, which is also standardized.
- Corrected p-values that control for the Family-Wise Error Rates are included in brackets in columns 2 and 6. A full description of our approach to multiple hypothesis corrections can be found in Appendix E.

Table A.XVIII: Summary of Decision Making Outcomes

	Savings	Default Pass-Through	Present Bias	Salience	Risk Preferences	Social Preferences
Night Sleep Treat	-3.46	-0.03	0.01	0.01	-0.04	-0.04
	(9.30)	(0.21)	(0.03)	(0.04)	(0.08)	(0.06)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[0.999]
Nap Treat	16.37	0.00	0.06	0.10	0.07	0.05
	(8.28)	(0.20)	(0.03)	(0.04)	(0.08)	(0.06)
	[0.434]	[1.000]	[0.416]	[0.130]	[0.996]	[0.996]

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table considers the treatment effect of the Night Sleep and Nap interventions on the six outcomes in the decision making family, as defined in the body of the paper: daily savings deposits, default pass-through, present bias, susceptibility to the salience of piece-rates, risk preferences, and social preferences.

- All signs are flipped as needed such that higher outcomes indicate more desirable outcomes.
- Risk preferences components include a standardized average of the point at which the participant switched from the risky to safe choice in the risk aversion game and the point at which the participant switched from the risky to safe choice in the loss aversion game. Social preferences components include a standardized average of the amount of money the sender sent in the dictator game, the amount of money the sender sent in the ultimatum game, the amount of money the sender sent in the trust game, whether the recipient accepted the sender's offer in the ultimatum game and the amount of money the recipient sent back to the sender in the trust game.
- Each column shows the OLS estimates of an equation similar to (2). In the regression, we pooled the two night-sleep treatments and controls for age, sex, and average baseline measures of the dependant variable (ANCOVA) when available. When the dependent variable is an index, we control for the index at baseline. Additional controls, such as the incentive reward rate, are included as relevant to match the primary specification in which each outcome is referenced in the body of the paper.
- Corrected p-values that control for the Family-Wise Error Rate are included in brackets. A full description of our approach to multiple hypothesis corrections can be found in Appendix E.

Table A.XIX: Timing and Frequency of the Study Tasks

	Time (1)	Day in Study (2)
Blood Pressure	Morning	Every 4 days
Weight	Morning	1, 28
Well-Being Survey	Morning	All days
Information about Sleep Treatment Assignment	10:00 - 12:30	8
Risk and Social Preferences Task	10:00 - 12:30	7, 26
BDM Task for sleep devices	10:00 - 17:00	31
Biking Task	11:00 - 20:00	28
Depositing Defaulted Savings (if applicable)	Morning	All days
Lunch	12:30 - 13:00	All days
Nap Explanation	13:00 - 13:30	9
Nap Time	13:30 - 14:00	9 - 27
Cognitive tasks - H&F, Corsi and PVT	14:20 - 16:00	2 - 27
Present Bias Task	17:00 - 20:00	4, 5, 6, 19, 20, 23
Sleep Devices Delivery	18:00	8
Savings Decision	End of the day	All days
Payment for the Day's Work	End of the day	All days

Notes: This table reports the timing of all the tasks in study.

- The time is recorded for standard days. When present bias and cognitive tasks are conducted on the first day, the time is longer than normal days for participants' understanding.
- Depositing defaulted savings happens at the end of daily survey. The savings survey is conducted at the end of the day and payment happens after saving survey. "Regular" day work ends by 20:00, while shorter day work ends by 17:00.
- Cognitive tasks have four slots daily. The total daily time each participant spends on cognitive tasks is 25 minutes, the duration of one slot. PVT is done daily (10 minutes), while Hearts and Flowers and Corsi are rotated by randomization (roughly 10 minutes each).
- See Appendix B.6 for additional details regarding the timing of the Present Bias task.

Table A.XX: Summary of Averages for Main Study (RCT) and Sleep Survey

	Survey Self-report (1)	RCT Self-report (2)	Survey Actigraph (3)	RCT Actigraph (4)
Night sleep	6.49	7.19	5.45	5.58
	(1.37)	(1.38)	(1.14)	(1.19)
Total sleep	6.92	-	5.84	-
	(1.48)	-	(1.22)	-
Night sleep efficiency	0.88	0.90	0.71	0.70
	(0.09)	(0.08)	(0.10)	(0.11)
$\mathbb{1}$(Any nap in preceding week).	0.57	0.73	0.46	-
	(0.49)	(0.44)	(0.50)	-
Naps per day	1.05	1.13	1.04	-
	(0.43)	(0.83)	(0.15)	-
Average nap duration (hr.)	1.19	1.52	0.85	0.25[1]
	(0.86)	(2.42)	(0.61)	(0.08)
Nap efficiency	-	-	0.72	0.85[1]
	-	-	(0.23)	(0.25)

Notes: This table presents summary statistics (averages) for relevant sleep variables comparing the main study (RCT) to the sleep survey (SS). Standard deviations are presetend below averages in parenthesis.

- Data from sleep survey actigraphy are cross-referenced with participants self-reports of sleep.
- Data from the main study (RCT) comes from the baseline period to ensure comparability with the sleep survey data.
- Main study (RCT) data on self-reported naps comes from the baseline survey where people answer questions about their usual sleep habits.
- Survey self-reports contain data from 3,387 participants, while survey actigraph data contains information on 440 participants.
- Nap time and average nap duration are conditional on napping.
- [1]: Conditional on napping for at least 1 minute.
- -: Data not available.

Table A.XXI: Sleep Survey Stagewise Take Up

	Percent of last stage (1)	Percent of total (2)	Frequency (3)
Census	49.93	49.93	3833
Baseline Survey	88.49	44.18	3392
Interest to hear about actigraph	39.15	17.30	1328
Willingness to wear actigraph	61.60	10.66	818
Actigraph installation	61.74	6.58	505
Endline Survey	97.43	6.41	492
Actigraph component participants (all)	89.23	5.72	439
Actigraph component participants (completed)	82.52	5.29	406
N			7677

Notes: This table presents take-up across the different stages in the sleep survey.

- N represents the total number of participants approached for the study, including refusals to the census.
- Installation of actigraphs could not be done for all willing participants for multiple reasons such as non-availability of the participant on the day of installation, refusal to wear the actigraph at the time of installation, shortage of actigraph devices at our disposal, and compliance with the upper limit of installing 20 actigraphs per locality.
- The difference between the number of participants for whom actigraphs had been installed and the number of those who participated in the actigraph component is the number of participants who dropped-out from the study after installation.
- "Actigraph component participants (all)" includes all participants who wore the actigraph for 1 night or more.
- "Actigraph component participants (completed)" includes only those participants who complied with the study's requirement of wearing the actigraph for 3 nights.

Table A.XXII: Sleep Survey Demographics

	Census (1)	Baseline (2)	Actigraph (3)
Gender (female)	0.72	0.72	0.65
Low income (by self-reported)		0.43	0.53
Middle income (by self-reported)		0.28	0.28
High income (by self-reported)		0.17	0.16
Low income (by house type)	0.11	0.12	0.18
Middle income (by house type)	0.65	0.67	0.64
High income (by house type)	0.24	0.21	0.18
Low income (by area)	0.06	0.07	0.10
Middle income (by area)	0.58	0.60	0.63
High income (by area)	0.36	0.33	0.27
Age	45.81	45.19	45.76
	(15.74)	(15.10)	(15.05)
Employed		0.39	0.43
No schooling		0.06	0.08
Highest grade attended		9.38	8.64
		(3.38)	(3.67)
College degree		0.31	0.22
N	3833	3392	439

Notes: This table presents demographics of participants who agreed to take part in the three stages of the study - census, baseline survey, and actigraph component.

- Different categories of income include - "self-reported" income data, collected through the baseline; income based on "house type" and "area", based on observation by surveyors. Income categories for "self-reported" income data are as follows: Low income - monthly household income below Rs. 20,000; middle income - monthly household income Rs. 20,000 and above, but below Rs. 40,000; high income - monthly household income above Rs. 40,000.
- The percents in the income categories based on self-reporting do not add up to 100 since the data here excludes the two categories - "Do not know" and "Do not want to disclose."
- 0 for the employment dummy denotes the following categories - unemployed, housewives, and retired without pension.

Table A.XXIII: Sleep Survey - Sleep Correlates

	Self-reported Night Sleep		**Actigraph Night Sleep**			**Actigraph Total Sleep**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	PW
Self-report Sleep Q2			0.54***			0.43***			
			(0.10)			(0.11)			
Self-report Sleep Q3			0.59***			0.56***			
			(0.10)			(0.11)			
Self-report Sleep Q4			0.88***			0.82***			
			(0.10)			(0.11)			
Middle Income	0.071	0.053		0.20	0.17		0.15	0.13	0.13
	(0.10)	(0.10)		(0.12)	(0.13)		(0.12)	(0.13)	(0.13)
Higher Income	0.056	0.0041		0.035	0.027		0.016	0.036	0.026
	(0.11)	(0.11)		(0.13)	(0.14)		(0.13)	(0.14)	(0.14)
Female	-0.039	-0.065		0.19**	0.18*		0.13	0.11	0.097
	(0.054)	(0.064)		(0.086)	(0.098)		(0.090)	(0.10)	(0.11)
Age 34 - 45	-0.54***	-0.50***		0.014	0.0046		0.026	0.0047	0.026
	(0.062)	(0.064)		(0.11)	(0.11)		(0.11)	(0.12)	(0.12)
Age 46 - 58	-0.58***	-0.52***		-0.55***	-0.57***		-0.50***	-0.55***	-0.49***
	(0.069)	(0.072)		(0.11)	(0.12)		(0.12)	(0.13)	(0.13)
Age 59 - 92	-0.45***	-0.40***		-0.058	-0.077		-0.032	-0.072	-0.021
	(0.072)	(0.075)		(0.11)	(0.11)		(0.12)	(0.12)	(0.12)
Children (numb)	-0.068***	-0.062**		-0.10**	-0.11**		-0.092*	-0.10**	-0.099**
	(0.026)	(0.026)		(0.045)	(0.045)		(0.047)	(0.047)	(0.050)
School No Col		-0.11			0.26*			0.20	0.22
		(0.11)			(0.13)			(0.13)	(0.14)
College		0.046			0.12			0.016	0.058
		(0.12)			(0.15)			(0.16)	(0.16)
Employment		-0.090			0.0030			-0.0022	-0.00025
		(0.057)			(0.088)			(0.093)	(0.096)
Constant	6.90***	6.99***	4.98***	5.42***	5.27***	5.25***	5.69***	5.61***	5.57***
	(0.12)	(0.17)	(0.070)	(0.16)	(0.20)	(0.076)	(0.16)	(0.21)	(0.22)
Mean	6.5	6.5	5.5	5.5	5.5	5.7	5.7	5.7	5.7
N	3389	3387	1367	1367	1367	1367	1367	1367	1367
Participants	3387	3387	439	439	439	439	439	439	439

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table considers correlations between a series of covariates (such as age and income) and self-reported night sleep, day-level actigraph-measured night sleep and total sleep.

- Actigraph total sleep is calculated from adding actigraph night sleep and naps.
- Covariates include self-reported sleep quartiles calculated from multiple survey questions and participants' demographic characteristics: age quartiles, sex, education level dummies, number of children, and whether the respondent is employed. Income levels are derived from the surveyors assessment of the income level of the participants neighborhood. Education level dummies include whether the participant attended school but stopped before college and whether they ever attended college, compared to no schooling.
- Column 9 uses probability weights, which are calculated by dividing income proportions in census data by income proportions of population from SES Mapping data.

Table A.XXIV: Sleep Survey Nap Correlates

	Whether Nap		**Acti Nap Cond. on Napping**	
	(1)	(2)	(3)	(4)
Middle Income	-0.10**	-0.10**	0.082	0.11
	(0.044)	(0.045)	(0.089)	(0.099)
Higher Income	-0.093**	-0.080	0.18*	0.21**
	(0.047)	(0.050)	(0.098)	(0.11)
Female	0.025	0.0079	-0.32***	-0.28***
	(0.025)	(0.028)	(0.084)	(0.096)
Age 34 - 45	0.028	0.020	-0.019	-0.0080
	(0.033)	(0.034)	(0.12)	(0.12)
Age 46 - 58	0.067**	0.049	-0.031	-0.056
	(0.033)	(0.035)	(0.12)	(0.12)
Age 59 - 92	0.062*	0.041	-0.096	-0.11
	(0.033)	(0.035)	(0.12)	(0.12)
Children (numb)	0.0034	0.00096	0.016	0.019
	(0.015)	(0.015)	(0.043)	(0.042)
School No Col		0.031		-0.35**
		(0.046)		(0.14)
College		-0.035		-0.29*
		(0.051)		(0.16)
Employment		-0.019		0.085
		(0.027)		(0.085)
Constant	0.29***	0.31***	1.03***	1.26***
	(0.051)	(0.066)	(0.15)	(0.22)
Mean	.26	.26	.89	.89
N	1367	1367	345	345
Participants	439	439	203	203

*** *Significant at the 1 percent level*, ** *Significant at the 5 percent level*, * *Significant at the 10 percent level*

Notes: This table considers the correlation between demographic controls and whether someone naps as measured by the actigraph, and actigraph nap time conditional on napping.

- Dependent variables are recorded on the participant-day level.
- Controls include income levels by the respondent's neighborhood and participants' characteristics: age quartiles, sex, education level dummies, number of children, and whether the respondent is employed. Education level dummies include whether the participant attended school but stopped before college and whether they ever attended college, compared to no schooling.
- Standard errors are clustered at the participant level.

Table A.XXV: Work Outcomes - Sensitivity to Measurement and Coding

	Labor Supply		**Earnings**			
	Code NA	Input Zero	Code NA		Input Zero	
Comparison	Levels	Levels	IHS	Levels	IHS	Levels
	(1)	(2)	(3)	(4)	(5)	(6)
Night Sleep vs. Control	-0.17***	-0.22**	-0.01	-7.55	-0.10	-12.88*
	(0.05)	(0.11)	(0.02)	(5.12)	(0.12)	(7.66)
Nap vs. Break	0.03	0.00	0.03*	11.11**	0.05	11.49
	(0.05)	(0.11)	(0.02)	(4.88)	(0.12)	(7.48)
Nap vs. No Break	-0.44***	-0.34***	-0.08***	-22.67***	0.02	-14.84*
	(0.05)	(0.11)	(0.02)	(5.00)	(0.12)	(7.81)
Control Mean	4.10	4.10	5.36	285.09	5.36	285.09
Control SD	2.20	2.20	2.37	224.08	2.37	224.08
N	6992	8288	6992	6992	8288	8288
Participants	451	452	451	451	452	452

Notes: This table considers the sensitivity of our results with respect to differences in coding and measurement of the earnings and labor supply variables.

- Row 1 shows treatment effects of the two Night-Sleep interventions (pooled) in comparison to the Control Group. Row 2 shows treatment effect of the Nap intervention in comparison to participants not randomized to naps who took a break during the nap time. Row 3 is analogous, but the comparison group consists of participants who worked during the nap period.
- The dependent variables are: labor supply (capturing hours actively typing) and total earnings from data-entry work.
- Each column shows the OLS estimates of equation (2), controlling for average baseline values of the dependant variable (ANCOVA), age, sex, "regular" study day, fraction of high piece-rate sessions, and day in study and date fixed effects. Standard errors are clustered at the participant level.
- In columns 1, 3 and 4 we drop participant-day observations in which the participant was absent to the office, whereas in columns 2, 5, and 6 we use the entire sample and input zero for all outcomes when the participant was absent. In columns 3 and 5 the dependent variable is the IHS transformation in earnings level, whereas in the remaining columns the dependent variable is in levels.

B Detailed Description of Tasks and Surveys

B.1 Survey of Experts

B.1.1 Data collection

Three versions of the expert survey were used in order to ensure that respondents were all well-informed regarding the questions asked and that the survey could be conducted in language familiar to the respondents (e.g. the statistical methods used). The three different surveys have similar introductory and concluding sections and all surveys asked for predictions on the impact of night sleep in the data entry task. Both economists surveys also elicited predictions on savings, while the behavioral economists survey additionally elicited predictions on present bias. For the sleep experts, we elicited predictions on cognitive and health outcomes, asking about outcomes in the Psychomotor Vigilance Task (PVT) and blood pressure. In order to make all predictions comparable despite survey administration over an extended period, all the statistics we provide to the respondents only used data from study participants that had finished the study by the time we sent the first survey wave.

The survey can be divided in three main parts. In the first part, we introduce important information necessary to be able to take the survey. Each key information was separated in an individual page, and participants could move back and forth between these different pages using buttons in the bottom of the page. The introductory pages had the following information: (i) explanation of the survey's goal, who was it directed for, and informed consent; (ii) overview of the study, explaining the night sleep intervention and how we measured sleep, (iii) average and variance of night sleep in the control and treatment groups; (iv) explanation of the data-entry task; (v) a benchmark treatment, in which we provided the treatment effect of quadrupling the piece rate on the number of correct entries in the data-entry task and, in a some versions of the survey, also the predictive effect of an additional year of education on the same outcome. In versions of the survey that were distributed via email list, we also asked the participants to identify themselves with their name, degree, and institution they worked for. Even though the participants could not move before entering this information, nothing impeded them of entering false information.

In the second part of the survey, we elicited the experts' predictions. In each page we elicited a single prediction, and participants could move back and forth between the pages to check and modify their answers. The typical page is represented in Figure A.XI. There are a few elements worth highlighting:

1. The predictions were supposed to be entered numerically in the white box in the middle of the screen.

Participants were asked to input their prediction as the difference in levels of the outcome variable between treatment and control groups.

2. In each question, we inform the respondent about the level of the outcome variable during the treatment period for the RCT's control group participants. We also provided a table at the bottom of the screen which mapped participant's answers to percentage and standard deviation changes for ease of interpretation. Finally, if the respondent passed the mouse over the red text below the table, a pop-up message appeared with the benchmark effects we had showed them in the last page of the introductory section.

3. Sleep medicine experts were asked to predict the intention-to-treat (ITT), i.e., the effect of being randomized to one of the night sleep interventions on outcomes. In both economists' surveys, we asked the respondent to predict the 2SLS coefficient associated with an hour increase in sleep. These differences are in accordance to common practice within each field.

4. For both groups, we informed then that the treatment effect of the pooled night sleep treatments on sleep was 32 minutes. So to make the predictions comparable across experts and with our own ITT results, we divide the economists' answers by 2 and the sleep science experts by 32/30.

5. To guarantee that the respondents had answered correctly, Qualtrics prompted the participants, just after they submitted their prediction, with a page repeating their answer, saying what it meant in percentage and s.d. terms, and asking if they wanted to continue with the survey or go back to modify their answer.

6. We allowed the respondents to skip any of the questions. They would be prompted with a pop-up asking whether they were sure they wanted to move on without answering the question, but they could just dismiss it, skipping to the next section.

The third part of the survey was a check-out page, thanking the respondents for their time and inviting them to add their email address in case they wished to receive information about the final results of the study and if they had any comments about the survey.

B.1.2 Data collection

We piloted the survey with PhD students from Harvard and MIT. In this phase we made constant alterations to the survey based on the comments and answers of the respondents.

In the first wave of data collection, we sent 68 personalized emails to researchers we know personally to the PIs. We classified 35 of the potential respondents as non-behavioral economists, 26 as behavioral economists, and 7 as sleep medicine experts. Of those 45 completed the entire survey and an additional 7 answered some of the questions.

We performed a second wave of data collection specifically for the sleep experts survey through an email list one of the authors in the paper was part of. This email list included young scholars studying topics related to sleep science. We obtained 2 additional completed responses from this source.

Because of the low number of answers from sleep science experts, we performed a third round of data collection with the sleep scientists. We sent an email to three email lists of professionals in the field of sleep medicine, including doctors and full-time researchers. Those emails were kindly sent by Dr. PhD Michael Perlis to email lists he was a part of. [42] In this wave, we were able to obtain 60 complete answers and an additional 10 incomplete answers.

In total, we use 107 complete surveys and 17 incomplete surveys.

A final issue important emphasizing is that our results were not available in any format publicly and we did not present them before the last wave of the survey of experts. Colleagues that were aware of early stage results through conversations with us were purposefully excluded. For the second and third waves of the sleep scientists, even though we cannot guarantee their identities, all participants included their names and we do not know any of them personally nor believe to be possible they would know the partial results of the study.

B.2 Attention in the Work Environment: Estimation Strategy

For each of the treatment groups j (i.e. night sleep, nap, and control), we estimate the (average) reaction to high piece-rate under the salient and non-salient conditions for output, productivity, and labor supply, denoted by ϵ_j^S and ϵ_j^{NS}, respectively. The attention parameter θ_j is defined as the ratio between the reaction to incentives under non-salient and salient conditions, i.e. $\frac{\epsilon_j^{NS}}{\epsilon_j^S}$. Importantly, we assume that the response to piece-rates under the salient condition is the full-attention benchmark, as in Chetty et al. (2009) and Allcott and Taubinsky (2015). We interpret θ_j as the deviation from the "full-attention benchmark" caused by inattention to non-salient incentives. Participants are fully-attentive even in the non-salient condition when $\theta_j = 1$ and completely inattentive when $\theta_j = 0$.

[42] We are extremely grateful to Michael Perlis for the help in reaching out to a vast network of sleep medicine experts. We would not have been able to reach nearly as many people without his unflagging support and generosity.

We estimate the treatment effect of the sleep interventions by comparing the attention parameter θ in each treatment group to the control group's θ. We first estimate the average reaction to incentives for each group j during the full salience and non-salient periods, using the OLS regression

$$y_{iwtd} = \sum_{j} \mathbb{1}_{\text{Treat}_i=j} \cdot \left(\beta_1^j \text{High}_{iwt} + \beta_2^j \text{Sal}_{it} + \beta_3^j \text{High}_{iwt} \cdot \text{Sal}_{it}\right) + \delta_i + \delta_t + \delta_d + \nu_{iwtd}, \tag{3}$$

where $\mathbb{1}_{\text{Treat}_i=j}$ captures whether participant i was in treatment group j, High_{iwt} captures whether the participant faced a high piece-rate during the 30-minute incentive window w (as described above), and Sal_{it} whether participant i was randomized to the salient condition on day t.

This equation differs from the benchmark reduced-form regression (2) in two ways. First, rather than using an ANCOVA specification as with other outcomes, we used participant-level fixed effects given the within-person variation in salience *during* the treatment period. Second, the unit of observation is the 30-minute window rather than the day given the frequency of potential incentive changes. We use the OLS estimates from equation (3) to recover $\hat{\epsilon}_j^{NS} = \hat{\beta}_1^j$ and $\hat{\epsilon}_j^{S} = \hat{\beta}_1^j + \hat{\beta}_3^j$. Finally, we estimate the attention parameter for each group by $\hat{\theta}_j = \frac{\hat{\epsilon}_j^{NS}}{\hat{\epsilon}_j^{S}}$. Standard errors in equation (3) are clustered at the participant level, while standard errors for $\hat{\theta}_j$ are estimated using the Delta Method.

B.3 Savings Task

Construction of counterfactual interest accrued variable. We define that savings at day 9 was zero, $s_9 = 0$, and take the participant's actual savings flow at date t, x_t, as given. Then for any day $t > 9$ we set counterfactual savings as $s_t = \max\{0, 1.01 \cdot (s_{t-1} + x_{t-1})\}$. It is necessary to introduce the maximum operator since because we set $s_9 = 0$ it is now possible to have negative balance sheets - that would be the case for participants deciding to withdraw quantities at day 10, $x_{10} < 0$, for example. Interest accrued at t is defined as $y_t = 0.01 \cdot (s_t + x_t)$ for $t \geq 9$.

For our ANCOVA specification, we repeat the same procedure but for the baseline period, setting $s_1 = 0$. We then control for the total interest accrued during baseline.

B.4 Default Task

Overview.

We implemented an experiment to measure the propensity to override default options in savings decisions. Each day, participants were randomized to have their survey completion fee deposited in their savings account or to be paid out along with their other payments at the end of the day. They could choose to override the default allocation each day when making their daily savings decision. The intention of this design was to identify possible effects of increased sleep on the strenght of default effects. We speculated that increased sleep could boost attention and memory or change the cognitive costs of making active decisions and thus reduce the stregth of default effects. Ultimately, the outcome measure ended up being severely under-powered, and thus we do not report it in the main text of the paper. Details are presented below.

B.4.1 Task Design

Default Process. As described participants, participants responded to a daily survey that elicited details about sleep, well-being, earnings, and consumption and expenditure patterns. Starting on the second day of the study, payments for completing the daily survey (Rs. 40) were either paid out in cash at the end of the day or put into the savings box (i.e. defaulted into savings) in the morning, right after the daily survey. That is, default savings were either Rs. 0 or Rs. 40. Once participants reached the end of the day, they made their savings decisions where they could 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 was in the savings box, including any payment deposited earlier in the day (default savings) or in previous days. Accordingly, default savings were not binding in any way. Participants learned whether their survey payment would be defaulted into their savings account after completing their daily survey but would not be reminded again of their default type over the course of the day.

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

Participant Instructions. Already on their first day at the study office participants learned that every day, upon completion of the daily survey, they would receive Rs. 40. Also, they were told that whether they received the survey payment in cash or would have it added to their savings in the box would be randomly decided by a computer. From the second day on, each day they after finishing the daily survey participants were informed whether they had been selected to receive the money in cash or in savings. In the case of the latter, the money would be added to the savings box in their presence and they would sign a receipt (as was standard practice for all deposits). Throughout the entire study, participants were unaware that they had been randomized in two groups according to the frequency in which their payment would be defaulted to savings; one-third of two-thirds of the time.

B.4.2 Treatment Effect on Default Pass-Through

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. 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 the difference in pass-through: the difference between the treatment and control groups in the percentage of the defaulted payment that the participants save due to the default effect alone.

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.

To estimate the treatment effect of the night sleep and the nap interventions, we estimate the following equation by OLS:

$$y_{itd} = \beta_1 T_i^{NS} + \beta_2 T_i^{Nap} + \alpha_1 D_{itd} + \alpha_2 T_i^{NS} * D_{itd} + \alpha_3 T_i^{Nap} * D_{itd} + \gamma_1 \overline{y}_{ib} + \gamma_2' X_{it} + \delta_t + \delta_d + \delta_s + \epsilon_{itd} \quad (4)$$

The outcome variable in this regression, y_{itd}, is daily deposits and the estimation mimics our benchmark specification, except for the inclusion of the default variable D_{itd}, which is either 0 or 40, and its interactions with night sleep and nap treatments.

The coefficients of interest are α_2 and α_3, respectively the night sleep and nap treatment effects on default pass-through. The term X_{it} captures control variables, including age, sex, daily piece rate, interest rate, maximum payment from cognitive tasks and randomized piece rate for the present bias task. Additionally, we control for the baseline daily average deposits ($\overline{y}_{ib}$) and for date (δ_t), day in study (δ_d), and surveyor (δ_s) fixed effects. Because participants were not informed of their default status until completing the daily survey, we exclude from our analysis absent days.

B.4.3 Treatment Effect on Propensity to Overrule Default

In the second specification, the key outcome of interest is whether the night sleep or the nap treatments affect 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 run the following specification:

$$y_{itd} = \beta_1 T_i^{NS} + \beta_2 T_i^{Nap} + \alpha_1 \text{Type}_{itd} + \alpha_2 T_i^{NS} * \text{Type}_{itd} + \alpha_3 T_i^{Nap} * \text{Type}_{itd} + \\ + \gamma_1 \overline{y}_{ib} + \gamma_2' X_{it} + \delta_t + \delta_d + \delta_s + \epsilon_{itd} \quad (5)$$

where the outcome variable y_{itd} is an indicator of whether the participant followed the default amount. Accordingly, $\overline{y}_{ib}$ is defined as the fraction of baseline days in which participant i did not deviate from the default option. The remaining controls variables are the same as in (4).

Distinct from (4), we do not estimate the default pass-through but instead are interested in treatments interactions with Type_{itd}, a dummy variable equal to one if the daily survey payment is defaulted into savings account rather than given to participants in cash. Again, we do not include absent days.

Results. Default days were associated with a 24 percentage point increase on the probability of saving exactly Rs. 40 and an average increase of Rs. 17.6 in deposits (44% of the default amount).

Because the default pass-through was not very high relative the savings overall, we are under-powered to detect any treatment effects on the adherence to the default condition. We would only be able to capture an effect if the treatment changed the daily deposit by Rs. 17, i.e., a 42 percentage points decrease in comparison to the control group. Correspondingly, we find no evidence of night time or nap treatment effects on individuals' propensity to overrule default or changing the default follow-through (Table A.XII). However, we also do not view this result as conclusive evidence of a lack of impact given the low power for this particular outcome.

B.5 Health Outcomes

We captured a battery of different outcomes relevant to participants' health over the course of the study. These measures include:

- *Stationary biking outcomes:* On the last day of the study, participants were asked to bike on a stationary bike for 30 minutes, with incentive payments for total distance. We recorded total distance covered in the 30 minutes and the maximum speed attained. *Pre-registered*

- *Self-reported illness:* Participants are asked about any symptoms of sickness (e.x., fever, cold, headache, etc.) they have experienced in the last seven days, recorded at baseline and endline. We record the maximum number of days in a week that the participant experienced at least one symptom. *Pre-registered*

- *Pain levels.* Participants are asked to self report pain on a scale of 1 to 10, recorded at baseline and endline. *Pre-registered*

- *Daily Activity.* Participants are asked how much their health has limited them in a certain number of activities. The possible answers range from "they did not limit you at all" (0, the best outcome) to "limited you a lot" (3, the worse outcome). The final scale, which is the sum of the answers, goes from 0 for people who are not limited at all in their daily life by their health to 36 for people who are substantially limited in their daily life by their health. Questions come from the SF-36 Health survey and are recorded at baseline and endline. *Pre-registered*

- *Blood pressure:* Systolic and diastolic blood pressure are measured 5 times for each participant over their time in the study using a digital blood pressure monitor and set protocol to ensure consistency. Blood pressure is winsorized at the 5% level. *Pre-registered*

- *Daytime steps:* In addition to tracking sleep, the actigraphs also count steps. We tracked daytime steps (defined as steps between 9am and 8pm) as a measure of physical activity.

B.6 Present Bias

B.6.1 Experimental Design

Real-Effort Experiment. Similarly to Augenblick and Rabin (2018), participants make decisions about how many pages to type on a fixed date (“work day", henceforth) under different piece rates. The work is very similar to the data-entry work they are used to, except that the pages are shorter to allow for a less coarse choice set for the participants. This work is completed at a fixed time after the completion of their regular working day, but before they get paid.

Choices. Participants have to make a total of 14 decisions. On each of them, the participants is offered a w^c and must choose how many pages they would like to type for that piece-rate. We impose a minimum and a maximum number of pages each participant can choose. All participants must choose to type at least 5 pages, which we impose to avoid fixed costs associated with moving from 0 pages to 1 page, as in Augenblick et al. (2015). We also impose an upper limit of pages the participants can choose. We define a participant-specific upper limit, which we determine based on the participant's typing speed up to that point in the study. We impose this limit so every participant can easily finish the task within two hours even when they choose the maximum number of pages, avoiding considerations of risk of not being able to complete the task by lack of time from the participants[43] Immediately after they make their last decision, we randomly select one of the decisions made by the participant to be the one that counts for the task they need to perform. For example, if decision c is selected, the choice's associated piece rate, w^c, and the participants choice, e^c, will be the piece rate and the output target of the participant on the work day.

Timeline. The decisions are made on two different dates: 7 on a day prior to the work day (prospective date), and 7 on the work day. The prospective date can be 1 to 5 days before the work date. The payment date is always at least one day after the work day. Moreover, the payment date is a function of the randomly selected choice: we designed it so the payment distance is fixed between the date of a given choice and the payment if that choice is selected. The difference in choices between the two decision dates is the base of the identification of time preferences.

Participants completed the present bias experiment once in baseline and 1 to 3 times during the treatment period.

Earnings from the Task. Earnings from the tasks consist of a lump-sum plus $w^s \cdot e^s$, where w^s is the piece rate and e^s is the number of pages in the selected choice. The participant only gets paid if they complete all the committed to within 2 hours, otherwise they receive nothing from the present bias task.

Changes during the Study. The major change in design during the experiment was that in the first design we repeated the same 7 piece rates on the two decision dates. However, the debriefing of participants revealed that they sometimes deliberately tried to choose the same number of pages across dates for any given piece rate for what we understood as a preference for consistency. In order to avoid that issue, we changed the piece rates so there were 7 pairs of pieces that were randomized in blocks between day 1 and day 2 of the task. All the 14 piece-rates would therefore be different, which might make it harder for choices to be guided by participants' demand for consistency. Also, to allow more time to elapse between the two

[43]We do that as a priori sleep could impact risk-aversion, which would then affect this trade-off. We do not find however any treatment effect in risk preferences.

decisions, we reduced the number of times participants completed the present bias experiment from 3 to 1 in the treatment period. That was necessary since participants only have 3 weeks during the treatment period.

Exclusion Criteria. Of the 452 participants in the study, we cannot estimate a present bias parameter for 54 of those. These 54 are broken down as follows: (i) 24 participants never completed a single date of the present bias experiment in the treatment period; (ii) 11 participants completed date 1 at least once but no date 2 in the treatment period; (iii) 19 participants always choose the maximum or always choose the minimum number of pages during the treatment period. Since we cannot identify time preferences for those, we exclude them. Of the 398 participants that do not meet any of these criteria, we cannot estimate the structural β for 46 because the algorithm does not converge. In our preferred specification we also exclude those, leaving us with a sample of 352 participants.

Experimental Integrity. Table A.XV present evidence that the present bias tasks was well implemented and that compliance with the experiment was similar across treatment arms. In the first row, we show that we can estimate individual-level β for 77%-78% of participants. This falls in the second version because we only attempt to measure present bias once per participant. Of those, only 5% of the participants make the first round choices but then never completed the work they had committed to do. Importantly, not only this is a very high compliance rate, but there are no differences across experimental arms. It is also worth noticing that when participants were absent on the "work date", we allowed them to complete the second day of the Present Bias Experiment on the next day they came to office. We find that this is only the case for 11%-14% of the participants, with no distinguishable differences across treatment arms (third row).

B.6.2 Structural Estimation of Present Bias

We estimate individual-level short-term discounting parameters β assuming participants choose the number of pages they would like to type by maximizing the utility function[44]

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

where T is the date of payment, t is the date of the work, k is the date of the choice, and $D_{k,t}$ is an indicator of whether $k = t$.

The first part of the utility function captures the cost of effort from the extra work. Following Augen-

[44] We estimate one model per participant

blick and Rabin (2016) (AR, henceforth), we assume the cost function has a power form in our benchmark specification , i.e.

$$c(e) = \frac{1}{\gamma} e^{\gamma} \tag{7}$$

In robustness checks we impose an exponential cost function, like in DellaVigna and Pope (2016), of the form:

$$c(e) = \frac{1}{\gamma} \exp(\gamma \cdot e) \tag{8}$$

The second part of the utility function captures the utility from choosing effort e under piece rate r, parameterized as

$$U_m(e \cdot w) = \phi \cdot w \cdot e + \alpha \cdot e \tag{9}$$

The first term of this function captures the utility of money. We found that some participants also appear to have an intrinsic motivation in working, which based on participants' debriefings is often linked to either reputation building (although we are explicit that we just want to know their preferences) or gift-exchange. We capture this effect with the term $\alpha \cdot e$ above. In practice adding this term improved our fit considerably. In our benchmark specification, optimal effort is given by

$$e^* \equiv e^*(k, t, T, w) = \left(\phi \cdot w \frac{\delta^{T-t}}{\cdot \beta^{\{t>k\}}} \right)^{\frac{1}{\gamma - 1}} \tag{FOC}$$

Finally, we assume that we observe the data with noise and with censoring at 5 and $\max_i > 5$. Thus, for choices interior to the participant's choice set, we assume we observe $\tilde{e} = e^*(k, t, T, w) \cdot \tilde{\varepsilon}$, where $\tilde{\varepsilon}$ is a log-normal error term independent across observations and from the covariates. When accounting for the possibility of censoring, we assume that the pages we observe being chosen is determined by

$$e_i = \begin{cases} 5 & \text{if } \tilde{e}_i < 5 \\ \tilde{e}_i & \text{if } 5 \leq \tilde{e}_i \leq \max_i \\ \max_i & \text{if } \tilde{e}_i > \max_i \end{cases}$$

In our benchmark specification, we estimate the utility parameters in 12 using a 2-sided Tobit model, with cost function 7 and return to effort 9. We also impose that $\delta = 1$ in our preferred specification, which improved the quality of estimation of our key parameter of interest, β.

Finally, we estimate one model per participant that completed the present bias experiment at least once during the treatment period only using data from the treatment period. We do the same for the baseline period, so we end up with one baseline and one treatment period estimate of present bias per participant. The structural estimation does not converge for 47 participants in the treatment period in our preferred specification, so we drop those from the sample. The structural estimation also does not converge to 10 participants in the baseline period. We replace those missing values with the average value across participants during baseline, since we only use this variable as a control.

B.6.3 Treatment Effect on Present Bias

To estimate the treatment effect of the night sleep and the nap interventions, we estimate the following equation by OLS:

$$y_i = \theta_{NS} D_i^{NS} + \theta_{Nap} D_i^{Nap} + \omega X_i + \varepsilon_i \tag{10}$$

The outcome variable in this regression is an individual-level estimate of present bias, measured by the structurally estimated β in our benchmark specification described above. We also show results when y_i is the OLS estimate $\hat{\beta}_i^{raw}$ from the following regression

$$\log e_{cit} = \beta_i^{raw} \text{Now}_{cit} + \gamma_i^0 + \gamma_i^1 \log w_{cit} + \epsilon_{cit} \tag{11}$$

where Now_{cit} is an indicator of whether t is the work date, $\log e_{cit}$ is the log of pages chosen and $\log w_{cit}$ is the piece-rate in choice c.

The coefficients of interest are θ_{NS} and θ_{Nap}, the intention-to-treat estimates for the night sleep and the nap treatments, respectively. The term X_i captures control variables, including the baseline present bias parameter, estimated in the same way as the dependent variable, except using only observations from the baseline, rather than the treatment, period. We also control for participant's sex and age, as specified in our benchmark reduced form equation in the PAP.

C A Model of Sleep

To clarify our argument regarding the roles of sleep quantity and quality, and to help interpret our findings, we present a simple time-allocation model, building on Biddle and Hamermesh (1990) and extending their setup to incorporate sleep quality as well as quantity. In this model, in addition to work and leisure, the agent has a third time-consuming activity: sleep. In our model, sleep is both a consumption good generating utility directly and an investment good affecting the agent's productivity.

C.1 Setup

The agent's preferences are described by the utility function

$$u(c, l, s) = c + \alpha z(s) + v(l) \tag{12}$$

where $\alpha z(s)$ captures the consumption value of sleep quantity s, and $v(l)$ captures the utility of leisure time, l. We assume that $z(\cdot)$ is an increasing and concave function such that $z'(0) \to \infty$, and we simplify the problem by assuming that $v(l) = -\frac{k}{2}(T-l)^2$.[45]

Time budget. The agent chooses the time spent in bed, b, which in turn produces a quantity of sleep $s \leq b$. We assume that sleep quantity is given by $s = b \cdot \chi$, where χ represents sleep efficiency, our main empirical measure of sleep quality.[46] Highlighting the fact that people with higher sleep quality have a smaller opportunity cost of sleep, the agent's time budget constraint is given by

$$T = h + l + \frac{s}{\chi} \tag{13}$$

where h represents time working.

Financial Budget. We assume that the agent earns $\lambda(s, \chi) \cdot h$, where $\lambda(s, \chi)$ is the productivity of an agent who sleeps s hours a day with quality level χ. This specification allows us to capture important features of the link between sleep and productivity. First, we assume that λ_s, the partial derivative of productivity with respect to sleep quantity s, is weakly positive. Holding quality fixed, we thus assume that more sleep cannot decrease productivity.[47] Additionally, we assume that gains in productivity from sleep come at diminishing rates, i.e. $\lambda_{ss} < 0$. Second, we assume that $\lambda_\chi \geq 0$. Holding quantity fixed, improved sleep quality will weakly increase productivity. This is consistent with evidence from sleep science finding that more interrupted sleep leads to worse daytime alertness, even controlling for total sleep time (Stepanski, 2002). The cross-partial derivative $\lambda_{s\chi}$ captures whether sleep quantity and quality are complements (> 0) or substitutes (< 0) for the agent's productivity. Finally, for the problem to be well-defined, we assume that $\lambda_{s^2} < 0$, i.e., there are

[45] This leisure function is chosen to simplify the calculations for comparative statics, but the results hold more generally for any increasing and strictly concave function v.

[46] More generally, the results are substantively the same if we assume that $s = b \cdot f(\chi)$, where f is a twice-differentiable function satisfying $f' > 0$ and $f'' \leq 0$.

[47] Some correlational studies indicate that sleeping more than 10 hours per night is associated with worse outcomes for adults. However, increases in sleep in the range typically achieved by our participants has been shown to increase productivity (Gibson and Shrader, 2018; Cappuccio et al., 2011).

decreasing marginal returns to sleep on productivity.

In this model, all of the agent's wage is used for consumption, implying that

$$c = \lambda(s, \chi) \cdot h \tag{14}$$

Agent's Problem. The agent's problem is to decide how much to consume and how much time to allocate to work, leisure, and sleep, while respecting their time and financial budget constraints. Formally, the agent solves:

$$\arg\max_{c,h,l,s} c + \alpha z(s) + v(l) \tag{15}$$

$$\text{s.t.} \quad T = h + l + \frac{s}{\chi} \tag{TC}$$

$$c = \lambda(s, \chi) \cdot h. \tag{BC}$$

Substituting in the time and budget constraints and using the assumed functional form of v, the problem becomes

$$\arg\max_{h,s} \lambda(s, \chi) \cdot h + \alpha z(s) - \frac{k}{2}\left(h + \frac{s}{\chi}\right)^2 \tag{16}$$

We denote this utility function by $U(h, s; \alpha, \chi)$.

C.2 Propositions and Discussion

The first proposition addresses the impact of increasing the consumption value of sleep, including due to our randomly-assigned financial incentives to sleep more. Such incentives to sleep can be modeled as an increase in α, which affects productivity $\lambda^*(\alpha, \chi) \equiv \lambda(s^*, \chi)$ via changes in sleep quantity s^* and work hours h^*. How sleep hours and work hours respond to changes in any parameter depends on two forces. First, the direct effect of the parameter on the marginal utility of sleep or work hours. These effects are straightforward. Second, and less obvious, whether sleep and work act as complements or substitutes in equilibrium. This relationship is an equilibrium condition which we denote by

$$U_{sh} = \lambda_s(s^*, \chi) - \frac{k}{\chi} \tag{17}$$

If $U_{sh} > 0$, then work and sleep are complements in equilibrium, meaning that an increase in sleep quantity will act as a force to increase work and vice versa. If $U_{sh} < 0$ sleep and work are substitutes, such that an increase in sleep will decrease work. There are two opposing forces determining the sign of U_{sh}. First, $\lambda_s(s^*, \chi)$ captures the effect of increased sleep on productivity, which we assume to be (weakly) positive. Second, $\frac{k}{\chi} < 0$, captures the fact that increasing sleep reduces leisure, driving an increase in the shadow price of work. The relative strength of these two forces will determine whether work and sleep are complements or substitutes.

Proposition 1. *Increasing the consumption value of sleep (or external incentive to sleep), α, will:*

i. Strictly increase sleep quantity s^*.

ii. Alter hours worked, h^*, by $h^*_\alpha = \dfrac{s^*_\alpha}{k} \cdot U_{sh}$, which is positive if and only if $U_{sh} > 0$, i.e., work and sleep are complements in equilibrium.

iii. Weakly increase productivity by $\lambda^*_\alpha = s^*_\alpha \cdot \lambda_s \geq 0$

The consumption value of sleep parameter α directly alters the marginal utility of sleep, but does not directly impact either work or leisure. Therefore, sleep increases with α, crowding out either work or leisure. As discussed above, the sign of $U_{sh} = \lambda_s - \dfrac{k}{\chi}$, the substitution pattern between sleep and work, determines which of these activities is crowded out. The sign of this term is ambiguous, hence, the treatments impact on labor supply must be determined empirically.

Moving from time allocation to productivity, Proposition 1 implies that the impact of the consumption value of sleep α on productivity should be non-negative. This follows from the fact that the impact of α on productivity is mediated by the increase in sleep. Because we assume that sleep quantity has a non-negative impact on productivity ($\lambda_s > 0$), increasing the consumption value of sleep will also (weakly) increase productivity.

The next proposition considers the sleep quality parameter, χ, and derives similar comparative statics for sleep, work, and productivity.

Proposition 2. *An increase in sleep quality χ has the following implications:*

i. Increases sleep s^* iff[48]

$$s^*_\chi \propto \left[\lambda_{s\chi} \cdot h + \frac{1}{\chi^2} \cdot \left(v'(l) + \frac{k}{\chi} \cdot s\right)\right] + U_{sh} \cdot h^*_\chi > 0 \tag{18}$$

[48]The term $\propto$ denotes that the terms have the same sign. We use this notation to suppress additional positive terms which do not alter the results or intuition. The complete derivation is presented in the proof of the proposition

ii. Increases work h^* iff

$$h^*_\chi \propto \left[\lambda_\chi + \frac{k}{\chi^2}\right] + U_{sh} \cdot s^*_\chi > 0$$

iii. Increases productivity λ^* iff

$$\lambda^*_\chi = \lambda_\chi + s^*_\chi \cdot \lambda_s > 0 \tag{19}$$

In addition, a sufficient condition for $\lambda^*_\chi > 0$ is that $s^*_\chi \geq 0$.

There are two key results in Proposition 2. The first concerns time use. Items 1 and 2 decompose the effect of sleep quality χ on time allotted to sleep and work. Both expressions have a similar structure. First, sleep quality affects the variables directly, by modifying the marginal utility of each activity keeping the other constant (terms in brackets). In addition, when sleep and work are complements in equilibrium ($U_{sh} > 0$), there will be a reinforcing effect as an increase in work increases the marginal return to sleep and vice-versa. If, however, sleep and work are substitutes in equilibrium ($U_{sh} < 0$), the opposite holds, and the effect of χ on both variables may be mitigated.

In more detail, Equation 18 highlights the three channels through which sleep quality χ affects sleep quantity. There are two elements of the direct effect. $\lambda_{s\chi} \cdot h$ captures the fact that increased sleep quality may change the marginal effect of sleep quantity on productivity. If quantity and quality are complements for work productivity, $\lambda_{s\chi} > 0$, whereas the opposite holds when they are substitutes. A second element of the direct effect, $\frac{1}{\chi^2} \cdot \left(v'(l) + \frac{k}{\chi} \cdot s\right) > 0$, captures the fact that an increase in sleep quality reduces the shadow cost of sleeping. This happens because as sleep quality improves, each minute in bed translates into more time asleep. This effectively reduces the price of sleep, causing the agent to consume more of it. Finally, as discussed above, there is also an indirect effect. Sleep quantity will also be affected by how hours worked react to sleep quality (χ), which is captured by h^*_χ in the end of equation 18.

Sleep quality similarly affects work via three channels. First, λ_χ captures the effect sleep quality has on productivity holding sleep quantity fixed. Second, similarly to the the effect on sleep quantity, increasing χ also reduces the shadow price of work. The reason is the same: for a given quantity of sleep, there is greater slack in the time constraint, allowing work to increase. The third channel, as before, captures the complementarity between work and sleep.

An important corollary to items 1 and 2 of Proposition 2 is that if sleep and hours worked are complements in equilibrium $u_{sh} > 0$ and if sleep quantity and quality are also complements for work productivity, an increase in sleep quality χ must increase both sleep quantity and work.

The second key result – in item (3) – details how sleep quality χ affects productivity. Equation 19 highlights the two channels through which sleep quality affects productivity. First, $\lambda_\chi \geq 0$ captures the fact that sleep quality may have a direct and weakly positive impact on productivity holding sleep quantity constant. Second, an increase in χ affects the optimal sleep quantity s^*. If the effect is positive, this feedback effect should further improve productivity, implying that χ increases productivity. However, it is possible that the net impact of increased quality could be negative if an increase in χ reduces sleep by enough to offset any direct productivity gain from χ.

C.3 Proof of Propositions

Lemma 1 provides sufficient conditions such that there is a unique optimum to problem 15 and that this optimum corresponds to the solution to the first order conditions (FOCs).

Lemma 1. *The following conditions are sufficient to guarantee that the solution to problem 15 is (i) unique, (ii) determined by its FOCs, (iii) and that the optimal choices of work, sleep, and leisure are positive.*

1. $k > \dfrac{\lambda_s^2(s,\chi)}{\alpha|z''(s)|}, \qquad \forall s \in [0,T]$
2. $k > \dfrac{\lambda(T,\chi)}{T}$
3. *Define* $\beta(s) = \frac{s}{z'(s)}$. *Then*

$$\lambda\left(\beta^{-1}\left(\frac{\alpha\chi^2}{k}\right),\chi\right) - \frac{k}{\chi}\cdot\beta^{-1}\left(\frac{\alpha\chi^2}{k}\right) > 0$$

Proof. To prove (i) and (ii) we argue that, under minor restrictions on k, the utility function $U(h,s)$ is strictly concave. Together with the interiority conditions (iii), we can conclude that there is an unique pair (s^*,h^*) that solves the problem and that the solution is completely characterized by the FOCs.

Because $U(h,s)$ is twice differentiable we can establish strict concavity by analyzing the definiteness of its Hessian matrix. The first order derivatives are given by

$$U_h = \lambda(s,\chi) - v'(T-h-sg(\chi)) \tag{20}$$

$$U_s = \lambda_s(s,\chi)\cdot h + \alpha z'(s) - g(\chi)v'(T-h-sg(\chi)) \tag{21}$$

where $g(\chi) \equiv \frac{1}{\chi}$. Using the functional form we imposed to $v(\cdot)$, we have $v''(\cdot) = -k$. Then, the first row of the hessian of $U(s,h)$ is given by

$$U_{hh} = -k \tag{22}$$

$$U_{sh} = \lambda_s - g(\chi)k \tag{23}$$

The second row of the hessian is given by

$$U_{sh} = \lambda_s - g(\chi)k \tag{24}$$

$$U_{ss} = \lambda_{ss} \cdot h + \alpha z''(s) - g(\chi)^2 \cdot k \tag{25}$$

The Hessian is then given by

$$H = \begin{bmatrix} -k & \lambda_s - g(\chi)k \\ \lambda_s - g(\chi)k & \lambda_{ss} \cdot h + \alpha z''(s) + g(\chi)^2 \cdot k \end{bmatrix} \tag{26}$$

The problem is strictly concave if

$$D \equiv \det(H) = -k[\lambda_{ss} \cdot h + \alpha z''(s) - g(\chi)^2 \cdot k] - [\lambda_s - g(\chi)k]^2 > 0 \tag{27}$$

for all s, h given α and χ.

Note that for any value of s, h we have $|\lambda_{ss}|k + 2\lambda_s g(\chi)k > 0$. This implies that for any value of s, h, the determinant $D > \tilde{D}(s) \equiv k\alpha|z''(s)| - \lambda_s^2$. Thus, a sufficient condition for problem 16 to be strictly concave is that $\tilde{D}(s) > 0$ for any $s \in [0, T]$. This condition is equivalent to

$$k > \frac{\lambda_s^2(s,\chi)}{\alpha|z''(s)|}, \qquad \forall s \in [0, T] \tag{28}$$

We know separately present sufficient conditions we ought to impose on the problem so that optimal choices of s, l, and h are positive.

Positive sleep. It is easy to impose that optimal sleep is positive. We just need to impose that $z'(0) \to \infty$. In that way, the marginal benefit of the first minute of sleep is large enough to hinder a zero sleep equilibrium.

Positive leisure. If leisure is zero, then the time constraint implies that $h + sg(\chi) = T$. Substituting this on U_h and using the functional form we impose on $v(\cdot)$ we have that $U_h = \lambda(s,\chi) - kT$. Suppose we had $U_h = \lambda(s,\chi) - kT < 0$. Agents would then be better off reducing their amount of time allotted to h. A sufficient condition for $l > 0$ in equilibrium is then that $U_h < 0$ when s attains its upper bound, $s = T$. This would imply that no matter what combination we had of s, h, to increase s in response to h would never lead to $U_h = 0$ as long as $l = 0$, since $\lambda_s \geq 0$. Hence, we must increase l. The condition can be re-written as

$$k > \frac{\lambda(T, \chi)}{T} \tag{29}$$

Positive hours working. Assume that $h = 0$. Then, the optimal s, denoted by $s(0)$, solves $U_s(s(0), h = 0)$, or

$$U_s = \alpha z'(s(0)) - \frac{k}{\chi^2} \cdot s(0) = 0$$

Rearranging this condition we have that

$$\beta(s(0)) \equiv \frac{s(0)}{z'(s(0))} = \frac{\alpha\chi^2}{k}$$

Note that $\beta(s)$ is an increasing function, so we can rewrite this equation as

$$s(0) = \beta^{-1}\left(\frac{\alpha\chi^2}{k}\right)$$

Then, we would like to find the condition that guarantees that $(s(0), h)$ is not a solution to the agent's problem. This will be the case when $U_h(s(0), 0) > 0$. Imposing that we get

$$\lambda\left(\beta^{-1}\left(\frac{\alpha\chi^2}{k}\right), \chi\right) - \frac{k}{\chi} \cdot \beta^{-1}\left(\frac{\alpha\chi^2}{k}\right) > 0 \tag{30}$$

□

Proof of Proposition 1.

In order to derive s^*_α and h^*_α we must first get the partial derivatives with respect to α. They are

$$U_{h\alpha} = 0 \tag{31}$$

$$U_{s\alpha} = z'(s) \tag{32}$$

Expanding the FOCs equations with respect to h, s, and any parameter θ and solving for ds and dh yields:

$$ds/d\theta = \frac{U_{s\theta}}{|U_{ss}|} + \frac{U_{sh}}{|U_{ss}|} dh/d\theta \tag{33}$$

$$dh/d\theta = \frac{U_{h\theta}}{|U_{hh}|} + \frac{U_{sh}}{|U_{hh}|} ds/d\theta \tag{34}$$

Isolating the endogenous and exogenous variables yields

$$ds/d\theta = |U_{hh}| \cdot D^{-1} \cdot \left[U_{s\theta} + \frac{U_{sh}}{|U_{hh}|} U_{h\theta}\right] \tag{35}$$
$$dh/d\theta = |U_{ss}| \cdot D^{-1} \cdot \left[U_{h\theta} + \frac{U_{sh}}{|U_{ss}|} U_{s\theta}\right] \tag{36}$$

where D is the determinant of the Hessian above.

To arrive at items i. and ii. we set $\theta = \alpha$:

$$\frac{\partial s^*}{\partial \alpha} = D^{-1} \cdot kz'(s) > 0 \tag{37}$$
$$\frac{\partial h^*}{\partial \alpha} = \frac{s^*_\alpha}{k} \cdot U_{sh} \tag{38}$$

Finally, to show iii. we derive the marginal gain in productivity from altering α:

$$\lambda^*_\alpha = \lambda_s \cdot s^*_\alpha \tag{39}$$

This is weakly positive because $s^*_\alpha > 0$ and, by assumption, $\lambda_s \geq 0$. □

Proof of Proposition 2.

The partial derivatives for χ are given by

$$U_{h\chi} = \lambda_\chi - sg'(\chi) \cdot k \tag{40}$$

$$U_{s\chi} = \lambda_{s\chi} \cdot h - g'(\chi)[v'(l) + sg(\chi)k] \tag{41}$$

We then set $\theta = \chi$ and use equation (35) to derive item i. This yields:

$$\begin{aligned} s^*_\chi &= |U_{hh}|D^{-1} \left(U_{s\chi} + \frac{U_{sh}}{|U_{hh}|} \cdot U_{h\chi}\right) \\ &= |U_{hh}|D^{-1} \left[\lambda_{s\chi} \cdot h - g'(\chi)[v'(l) + sg(\chi)k] + \frac{(\lambda_s - g(\chi)k)}{|U_{hh}|} \cdot (\lambda_\chi - sg'(\chi) \cdot k)\right] \end{aligned}$$

Because $|U_{hh}|D^{-1}$ and $|U_{hh}|$ is positive we may write, after substituting $g(\chi)$ for its functional form

$$s^*_\chi \propto \left[\lambda_{s\chi} \cdot h + \frac{1}{\chi^2} \cdot \left(v'(l) + \frac{k}{\chi} \cdot s\right)\right] + U_{sh} \cdot h^*_\chi$$

We use equation (36) to derive ii. We then have

$$\begin{aligned} h^*_\chi &= |U_{ss}|D^{-1} \left(U_{h\chi} + \frac{U_{sh}}{|U_{ss}|} \cdot U_{s\chi}\right) \\ &= |U_{ss}|D^{-1} \left[(\lambda_\chi - sg'(\chi) \cdot k) + \frac{(\lambda_s - g(\chi) \cdot k)}{|U_{ss}|} \cdot (\lambda_{s\chi} \cdot h - g'(\chi)[v'(l) + sg(\chi)k])\right] \end{aligned}$$

The same reasoning as before implies that

$$h_\chi^* \propto \left[\lambda_\chi + \frac{k}{\chi^2} \cdot s\right] + U_{sh} \cdot s_\chi^*$$

Finally, we apply the chain rule to derive λ_χ^*:

$$\lambda_\chi^* = \lambda_\chi + \lambda_s \cdot s_\chi^* \tag{42}$$

and get that $\lambda_\chi^* > 0$ if $s_\chi^* > 0$, since λ_χ is weakly positive and $\lambda_s > 0$, by assumption.

□

D Additional results

D.1 Present Bias

Model Selection. We test 8 different variations on the estimation procedure. We tried two different cost functions 7 and 8 – power and exponential, respectively – with and without the α parameters capturing non-pecuniary reasons to exert effort - and with or without censoring. In Table A.XIV we show that the specification with power cost function and allowing for $\alpha \neq 0$ is far superior than the other options, regardless of censoring. First, we show in column 5 that that the number of failures in convergence of the non-linear estimation algorithm is 2 to 3 times larger for the exponential than for the power cost function. Second, in column 6 we show that the model's fit, as measured by the Negative Log-Likelihood (NNL) of the estimated parameter, is orders of magnitude better for the power cost function. Estimating the model allowing for $\alpha \neq 0$ also increases the fit substantially as measured by NNL.

These results underpin our choice to use the specification with power cost function and non-pecuniary $\alpha \neq 0$ as our benchmark estimates.

D.1.1 Evidence of Present Bias

Figure A.IX shows two important features of the choices participants are making in the present bias experiment. First, we show that participants make, on average, monotonic choices, increasing the amount of work they choose for higher piece-rates. Second, there is a clear shift in the "labor supply" curve when decisions are made in the first date of the present bias experiment. This shift is at the root of our strategy to identify present bias, as it suggests participants demand a premium to exert effort immediately in comparison to only a few days afterwards.

There are two additional noteworthy features in Figure A.IX. First, participants seem to be more present bias in baseline than in the treatment period, especially in the first version of the task. Second, there is very little evidence of present bias in the first version of the task during the treatment period. As we discussed in detail in Section B.6, we identified and corrected a couple of shortcomings with our original design. Reassuringly, in our second design – the one used to estimate present bias for two-thirds of our sample – there is a clear shift in labor supply in the treatment period, indicating the presence of present bias in the sample.

Shifting to the structural estimation of present bias, in Table A.XIV we also present key moments of the present bias distribution in the population for each of the estimating methods. In our preferred specification (first row of each panel), there is little evidence of present bias *in the median* ($\beta_{med} = 0.98$)[49] of the distribution. However, this masks substantial evidence of present bias on lower parts of the distribution. For example, in the treatment period (Panel B), the participant in percentile 25 has $\beta_{25} = 0.84$.

One concern is that the results in the tail are simply driven by estimating noise. If that is true, the percentiles of the present bias distribution should be roughly symmetric around the median.[50] For example, the participant in percentile 75 should have β_{75} close to $1 + (1 - \beta_{25}) = 1.16$. However, $\beta_{75} = 1.03$, which is very close to the median. Moreover, the estimated distribution has a large mass on values close to the median, much larger than it should for a normal distribution considering the observed kurtosis, and is also skewed to the left. Those facts are consistent with a population where many participants (perhaps more than 50%) are time consistent (i.e., $\beta = 1$) and that the rest of the participants are time-inconsistent, most being present, rather than future, biased.

D.1.2 Present Bias Correlates

In this section, we provide more supporting evidence that the measure of β we estimate do capture present bias in the population. We complement these findings with outcomes in our study.

In Table A.XVI, we regress a series of behaviors that should be affected by time preferences on our preferred, structurally estimated measure of present-bias. [51]

[52]

[49] We focus on the median as a few very large outliers drive the average up.

[50] Assuming the noise has a symmetric distribution which we, of course, cannot guarantee in small samples

[51] Incentivized lab experiments in the spirit we use to elicit present bias have been shown to correlate with real world outcomes expected to be affected by present bias. (Meier and Sprenger, 2010) for credit card debt and (Martinez et al., 2017) for procrastination in tax filing.

[52] We average the dependent variable on the participant level throughout all the study (including baseline period).

In column 2, we show that an increase of 0.1 in β is associated with an increase of Rs. 3.78 in daily deposits (approx. 3% increase). We also show that participants are more often late to office when β is smaller. An increase of 0.1 in β is associated with a 2% decrease in lateness (column 4). This could be especially costly for participants in "short days", when arriving late makes them lose a relatively large financial bonus (Rs. 50). This could be explained by present bias if individuals underweight the future benefits of punctuality relatively to leaving earlier from home, for example, where the cost is immediate. We also test whether present bias correlates with length of voluntary pauses (columns 5 and 6) and with night sleep duration (columns 7 and 8), but find no positive association with those. Our results, albeit small, are in line with the existing evidence in the literature which shows a robust small correlation betwen experimentally elicited measures of present bias and real-world behavior.(Cohen et al., 2019)

Our present bias measure only uses the treatment period. The results are similar when we pool across baseline and treatment period, but less significant.

E Multiple Hypothesis Corrections

We pre-registered adjusting p-values for multiple hypothesis corrections within our two primary families of outcomes: (i) work-related outcomes (productivity, hours worked, and earnings) and (ii) decision-making (savings, default effects, time preference, inattention, risk preferences, and social preferences). The corrected p-values for the work-related outcomes are reported in Table III. For the decision-making outcomes, no tests of significance survive our corrections. A summary of these outcomes along with the corrected p-values can be found in A.XVIII.

To apply these corrections we ran simulations to control the Family-Wise Error Rate. We took this approach rather than applying a formulaic correction (e.g. Holm or Bonferroni) so that we could capture correlations across outcomes in our data.

More specifically, our simulations followed the steps described below:

1. Select one of the primary families of outcomes, defined above.
2. Run 5000 iterations according to the following sub-steps:
   - Re-randomize the treatment assignments (night sleep and nap). When randomizing, follow the same stratification procedures as in the RCT.
   - Run the core regressions relevant to the family in question. For instance, for the work-related outcomes run the main productivity, labor supply, and earnings specifications.
   - Save the z-scores computed for each regression coefficient, so the result is 5000 z-scores multiplied by the number of outcomes in our family.
3. For each test of interest (for instance, the impact of night sleep on labor supply), examine the distribution of z-scores arrived at through the simulation. Identify the percentage of iterations for which at least one of the tests within the family would have been rejected at the critical value actually observed for the test in question in the RCT data. Because the treatment assignments underlying these tests were re-randomized, all observed rejections are rejections of a true null. As such, the observed percentage of iterations for which at least one test is rejected corresponds to the corrected p-value and is what is reported in Table III for the work-related outcomes.
4. Repeat for each family of outcomes.

F Sleep Survey

To explore the external validity of our RCT sample and deepen our understanding of sleep characteristics among different segments of the population - in particular the relationship between sleep and income - we conducted a larger-scale survey supplemented by actigraph data across a more representative sample of the adult Chennai population.

Recruitment. Neighborhoods were randomly selected from a stratified sample of geo-locations across Chennai. Lower-income households were more likely to participate in our study, so we over-sampled individuals from higher-income neighborhoods. In total, 7,677 participants were approached, 3,833 agreed to participate in at least the first stage of the survey, and 439 completed three nights of actigraph measurements.

Survey Stages. The survey consisted of three key stages: (i) a Census and Baseline survey, in which individuals were asked a set of questions about their personal and self-reported sleep characteristics; (ii) an Actigraph study, where participants wore an actigraph for three nights; and (iii) and Endline survey, where participants who undertook the Actigraph study were asked to self-report their sleep patterns over the previous four days. The portion of participants who agreed to participate at each stage (and sub-stage) of the study can be found in Appendix Table A.XXI, and the demographic characteristics across the first two stages can be found in Appendix Table A.XXII.

Findings. The first key takeaway is that this boarder sample of Chennai is severely sleep deprived, sleeping just 5.5 hours on average per night according to the Actigraph.[53] This result is nearly identical to the 5.6 hours of sleep found among RCT participants. Similarly, the individuals in the sleep survey also have similar sleep quality to RCT participants, with 71% sleep efficiency.

Sleep characteristics do not vary substantially by household income, education, or employment status.[54] Women, meanwhile, do sleep more than men; the more children someone has the less they sleep; and middle-aged individuals sleep less than younger or older adults (Appendix Table A.XXIII). The survey also revealed

[53]Although only a fraction of participants agreed to wear the actigraphs, based on self-reports, those individuals do not appear to be selected on sleep duration.

[54]It is important to note however, that given the income distribution of the city, very few participants in the survey would be considered "middle class" or "wealthy" by international standards.

that daytime naps are common in this population - 37% of individuals report napping on any given day. Higher-income individuals are less likely to nap, but conditional on napping spend more time asleep. Older participants are also more likely to nap on any given day (Appendix Table A.XXIV).

G Deviations from the Pre-Analysis Plan and Original Study Design

This study was pre-registered with AEA (ID: AEARCTR-0002494) under the title "Sleepless in Chennai: The Consequences of Sleep Deprivation Among the Urban Poor." Pre-registration took place before endline data collection began. All changes, and rationales for the changes, are listed below. Adjustments were typically made because the pre-registered specification or variable definitions presented unforeseen conceptual issues or due to changes in study design (e.g. reduced frequency of a task). We show specifications we had pre-registered whenever possible for comparison in Appendix A.

Empirical Specification

- **IV Specification.** We pre-registered IV specifications which instrumented sleeping time with treatment status. However, reduced form regressions provided similar results that were less subject to extrapolation outside of the observed range of treatment effects (i.e. the average increase in night time sleep was roughly 30 minutes, but IV effects would be reported per additional hour of sleep for ease of interpretation).

Typing

- **Absent Days.** The PAP specifies that earnings from the typing task and the labor supply variables would be coded as zero on days when participants were absent. This plan was made to account for potential imbalances in attendance across the treatment groups. In practice, however, attendance is very well balanced across treatment groups (Table A.III) and excluding missing observations improves statistical power without changing results qualitatively (See Table A.XXV for a comparison of the main estimates under different hypothesis).
- **Typing Earnings Variable.** In the PAP we specify that we would transform earnings in Rupees using an inverse hyperbolic sine transformation (IHS). However, this transformation is not needed given that earnings are not heavily right-tailed and missing days are omitted. Hence, we report earnings in levels for ease of interpretation. In Table A.XXV, we show the results with the IHS transformation of earnings, which present qualitatively similar results.
- **Speed and Accuracy.** The definitions of the speed and accuracy variables used in Table A.VI were altered in order to be able to decompose the effects linearly and interpret which channels was driving

changes in productivity.

- **Special Days Controls.** The PAP called for control variables to be used for days which differed from a "typical" day (e.g. days with unusual activities such as the real effort task used to measure present bias). Including this control did not matter for our results so we ended up choosing to drop it.

Savings

- **Savings dependent variable.** We pre-registered daily net savings as our primary outcome variable for savings. However, we discovered during data collection that this measure was problematic as the estimation was driven by a few individuals with huge large withdrawals close to the end of the study. As discussed in Section 7.1, we believe these withdrawals were driven by the study design rather than participants underlying savings behavior. Hence, in addition to this measure, we daily deposits and interest accrued.

- **Interest Rates.** Interest rates were changed in reaction to participant understanding and to allow estimation semi-elasticities to benchmark treatment effects. Specifically, in the first 7 months of the study participants received the pre-registered *daily* interest rates of 1% and 2%. In December 2017, we switched from computing interest only on days when we administered the savings survey to computing it every day, including weekends. In May 2018, we briefly changed interest rates to 1% and 2% *weekly*. Finally, in June 2018, the interest rates were changed to 0% to 1% percent for new participants to enable us to calculate the semi-elasticity both from 1% to 2% as well from 0% to 1%. Importantly, given the rolling enrollment the allocation of treated and control participants across these changes is well balanced.

- **Cap on savings.** The limit on daily deposits was increased from Rs. 400 to Rs. 600 because participants were frequently reaching the original cap.

Preferences and Cognitive Function (Attention)

- **Present Bias.** Deviations from the PAP and our current procedures are described in Appendix B.6.

- **Attention in the Work Environment.** The contrast between the salient and non-salient versions of the incentives was increased 11 months after the study began. As a result, 47% of the participants were exposed to the old design, while the remainder was exposed to the new, more salient design.

Risk and Social Preferences

- **Level of observations.** Regressions for the Risk and Social Preferences tasks were mistakenly pre-registered in the participant-day level. However, the participants only complete the Risk and Social task twice in the study, once before and once after the treatment. Accordingly, we specify our regressions in the paper at the participant level using the first measurement as the baseline control.

Well-being

- **Outcome components.** Our pre-registrations included three measures to rely on when creating the subjective well-being index - happiness, depression, and life possibilities, as measured by a Gallup Cantril Scale. We later added questions on life satisfaction and self-reported stress. In Table V we present the results for both indices, one reflecting what was pre-registered and the other including the additional measures.