THE ECONOMIC CONSEQUENCES OF INCREASING SLEEP AMONG THE URBAN POOR

Pedro Bessone
Gautam Rao
Frank Schilbach
Heather Schofield
Mattie Toma

Working Paper 26746
http://www.nber.org/papers/w26746

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2020, Revised November 2020

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 or the National Bureau of Economic Research. We thank numerous seminar audiences and colleagues for helpful comments and generous feedback. We thank Alosias A, Vivian Aluoch, Srinivas Balasubramanian, Dillon Bowen, Phoebe Cai, Stephanie Chan, Fiona Chen, Thomas ESCANDE, Juliette Finetti, Gabriel Jardanovski, Kannan Kumar, Gunjita Gupta, Erik Hausen, Yihong Huang, Dexin Li, Andrew Locke, Maddie McKelway, Adrien Pawlik, João Pugliese, Jane von Rabeau, Sangeetha Ramanathan, Cory Rand, Maya Roy, Krishna Prasad Srinivasan, Nikkil Sudharsanan, Sifan Xue, 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.

The Economic Consequences of Increasing Sleep Among the Urban Poor
Pedro Bessone, Gautam Rao, Frank Schilbach, Heather Schofield, and Mattie Toma
NBER Working Paper No. 26746
February 2020, Revised November 2020
JEL No. C93,D9,I1,I12,I15,O1,O12,O18

ABSTRACT

Using state-of-the-art technology, we document that adults in Chennai sleep only 5.5 hours per night on average despite spending 8 hours in bed. Their sleep is highly interrupted, with sleep efficiency—sleep per time in bed—comparable to those with disorders such as sleep apnea or insomnia. A randomized three-week treatment providing information, encouragement, and improvements to home sleep environments increased sleep duration by 27 minutes per night but came at the cost of more time in bed. Contrary to expert predictions, increased night sleep had no detectable effects on cognition, productivity, decision-making or well-being, and led to small decreases in labor supply. Yet, increased sleep can have benefits in this setting: short afternoon naps at the workplace improved an overall index of outcomes by 0.12 standard deviations, with significant increases in productivity, psychological well-being, and cognitive function, but less time available for work.

Pedro Bessone
MIT
50 Memorial Drive
Cambridge, MA 02142
tepedino@mit.edu

Gautam Rao
Department of Economics
Harvard University
Littauer M-30
Cambridge, MA 02138
and NBER
grao@fas.harvard.edu

Frank Schilbach
MIT Department of Economics
The Morris and Sophie Chang Building
50 Memorial Drive, E52-560
Cambridge, MA 02142
and NBER
fschilb@mit.edu

Heather Schofield
University of Pennsylvania
Perelman School of Medicine and
Wharton Business School
Blockley Hall 11th floor
423 Guardian Drive
Philadelphia, PA 19104
heather.schofield@gmail.com

Mattie Toma
Department of Economics
Harvard University
Littauer
Cambridge, MA 02138
mattietoma@g.harvard.edu

A randomized controlled trials registry entry is available at https://www.socialscienceregistry.org/trials/2494

1 Introduction

Understanding the lives of the poor is central to modern development economics. Economists have studied many deprivations associated with poverty, such as lack of access to nutrition, education, health care, and clean air and water. In contrast, the most time-consuming activity of our lives—sleep—remains largely unexamined in poverty research and policy. The urban poor in low-income countries face many barriers to a good night's sleep, such as heat, noise, crowding, physical discomfort, and psychological distress. Sleep could be a crucial input to their productivity, well-being, and cognitive function. Yet we know little about how much and how well people in low-income countries sleep, or the returns to policies that seek to increase sleep.

Using state-of-the-art technology to measure sleep objectively, we uncover widespread sleep deprivation in two samples of low-income adults in Chennai, India. People in our samples sleep only 5.5 hours per night on average, far below the minimum level recommended by sleep experts (Hirshkowitz et al., 2015; Watson et al., 2015). This is not due to a lack of trying. They spend about 8 hours per night in bed, but their sleep is highly disrupted, with 32 awakenings in a typical night. The implied sleep efficiency—time asleep per time in bed—of 70% is much lower than objective measures from general US populations, and similar to those suffering from disorders such as sleep apnea (Hedner et al., 2004) or insomnia (Trauer et al., 2015) in high-income countries.

An enormous body of research, mostly conducted in sleep labs in rich countries, documents severe negative impacts of sleep deprivation on a range of outcomes from attention and memory to immune function and mood (Lim and Dinges, 2010; Banks and Dinges, 2007). While experimental evidence on the impact of increasing sleep in field settings is scarce,[1] there is a widely-held belief among researchers and the public that reducing sleep deprivation would lead to improvements in economic outcomes (Walker, 2017). To document these priors, we surveyed 118 experts from sleep science and economics. The experts predicted sizable economic benefits, including a 7% increase in work output, of increasing sleep by half an hour per night from the low levels observed in our setting.

To measure the economic impacts of increasing sleep in the field, we conducted an RCT with 452 low-income adults in Chennai (our main sample). We employed participants in a data-entry job with flexible hours for one month, allowing us to precisely measure the productivity and labor-supply effects of increasing sleep in a real-world environment. Our first randomized treatment consisted of two interventions to increase night sleep in people's natural home environments. These interventions included (i) daily information about their night sleep and practical tips to increase it, (ii) items to improve people's home-sleep environments, and (iii) verbal and/or financial encouragement to increase night sleep.[2]

The night-sleep interventions increased sleep by an average of 27 minutes per night for three weeks

[1] Notable exceptions include Avery et al. (2019), who evaluate commitment contracts to increase sleep among college students in the United States and United Kingdom, and Barnes et al. (2017), who study the effects of cognitive behavioral therapy for insomnia on job satisfaction and related outcomes.

[2] 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.

over a base of about 5.6 hours per night.[3] This gain is larger than many alternative interventions such as sleeping pills (Riemann and Perlis, 2009), cognitive behavioral therapy for insomnia (Trauer et al., 2015), or commitment contracts (Avery et al., 2019). The increase in sleep duration was entirely driven by greater time spent in bed—on average 38 additional minutes per night—rather than improved sleep efficiency. The effects on sleep did not vary systematically with baseline sleep quantity, efficiency, or other observable characteristics such as gender and age. Our results show that people do have substantial ability to adjust their time asleep through changes in time in bed, but improvements to sleep efficiency appear to be costly. Given the low sleep efficiency, increasing sleep costs a great deal of time, with potentially high opportunity costs.

In contrast to expert predictions, we find no significant positive impacts of increased night sleep on average across our outcomes or even on any individual outcome. The night-sleep treatment group was not significantly more productive at the data-entry job, despite working on a relatively cognitively-demanding task intended to be sensitive to sleep deprivation. Instead, increased sleep came at the cost of lowering labor supply by nine minutes per day, leading to a small (but not statistically significant) decrease in earnings. We clearly reject the median expert prediction of a 7% increase in output ($p < 0.001$). Similarly, we find no significant effects of increased nighttime sleep on detailed measures of physical and psychological well-being, a battery of cognitive tests, and standard measures of social, risk, and time preferences.

Why does increased night sleep not have benefits in our setting, contrary to expert predictions and a large body of lab studies? One possibility is that the large effects from lab experiments, which typically drastically reduce sleep for up to a few nights, do not generalize to marginal, policy-relevant increases in sleep in the field. Another possibility is that the low quality of sleep observed in our setting—as proxied by low efficiency and frequent awakenings—explains the lack of benefits. It could be that returns to increased sleep would indeed be high in typical rich-country settings. We cannot adjudicate these reasons, but our results highlight the importance of studying sleep in the field, where outcomes have real stakes and sleep is a choice variable with opportunity costs. They also caution against extrapolating sleep-science findings across diverse contexts.

However, even in our context, we find that sleep *can* matter. Evaluating a cross-randomized treatment that offered participants the opportunity for a daily half-hour afternoon nap at the office, we find evidence of improvements across a number of outcomes. Naps increased work productivity by 0.04 SD (2.3%), boosted cognitive function by 0.1 SD, and psychological well-being by 0.12 SD. We also find suggestive evidence of naps increasing 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. However, naps still entail significant opportunity costs. Napping reduced labor supply and thus total output and earnings compared to simply working through the afternoon, while increasing earnings compared to taking an enforced break of the same length.

Given the large number of outcomes, we pursue three approaches to correct for multiple hypothesis tests. First, we create a combined index variable which averages across all (standardized)

[3] In most of our analysis we pool across the two night-sleep treatments.

outcomes following Anderson (2008). Naps had a treatment effect of +0.12 SD (s.e.=0.04, p<0.01) on this aggregate index, significantly different than the small negative effect of −0.01 SD (s.e.=0.04) achieved by the night-sleep treatments. Second, we create indices or pick summary variables at the level of the main families of outcomes, and correct for the existence of multiple families. This reveals significant positive effects of naps on earnings ($p = .06$, compared to breaks), well-being ($p = .03$), and cognition ($p = .09$).[4] Finally, to test for effects on individual outcomes, we correct for multiple outcomes within each family. Under this approach, naps have significant *positive* effects on productivity, psychological well-being, and lab measures of cognitive function; on output and earnings when compared to breaks; and a significant *negative* effect on labor supply, output, and earnings when compared to working through the nap period.

The different impacts of additional night sleep and nap sleep are striking, especially considering that naps add less than a third as many minutes of additional sleep. One possible reason for this difference is simply that the timing of sleep may matter. Contrary to hypotheses and some evidence in sleep science (e.g. Nicholson et al., 1985; Mollicone et al., 2007, 2008), naps and night sleep may not be close substitutes. An alternative explanation is that sleep quality may play a role, since naps in our study occurred in a more comfortable office environment. We cannot separate these explanations, but hope that future work in similar settings may help answer this question.

Our paper makes the following contributions. First, it contributes to a better understanding of the living conditions faced by the poor in developing countries, by providing objective measures of sleep. We discover surprisingly low levels of sleep duration and efficiency among the urban poor in Chennai. These findings are consistent with two recent papers measuring sleep objectively in smaller samples in Sri Lanka and Haiti (Schokman et al., 2018; Castro et al., 2013) and contrast with self-reported measures of sleep, which may fail to capture the low sleep efficiency and its impact on total sleep (Stranges et al., 2012; Gildner et al., 2014; Simonelli et al., 2018).

Second, we build on a recent literature that estimates the causal impact of sleep outside of sleep laboratories. The lack of impacts of night sleep we find using a field experiment contrasts with an economics literature which uses natural experiments in rich countries to demonstrate that sleep can have sizable effects on wages (Gibson and Shrader, 2018; Giuntella and Mazzonna, 2019), hospitalizations (Jin and Ziebarth, 2020), accidents (Smith, 2016), and civic behaviors (Holbein et al., 2019).[5] We speculate that the stark difference in sleep efficiency in our setting compared to rich-country populations may explain this difference. Studying the economic impact of increasing sleep efficiency (or more generally sleep quality) is an important question for future research.

Third, we show that afternoon naps in a comfortable office environment have positive effects on a range of outcomes, including productivity, well-being, and cognition. Naps are a common feature of

[4]For the work outcomes, earnings serves as the summary variable, since it naturally combines labor supply and productivity. We compare the nap treatment separately with the work and enforced-break control conditions for work outcomes, since these different conditions have large mechanical differences in labor supply, making the pooled control treatment effect difficult to interpret. For other outcomes, we simply compare naps with the pooled control group.

[5]Our findings also contrast with Jagnani (2018), who exploits variation in sunset times in India to show that less time in bed is associated with worse educational outcomes for children. It could be that children are farther away from optimal sleep levels, or that sleep quantity may matter more for children and for learning outcomes.

life around the world, and are particularly prevalent in tropical countries (Dinges, 1992). High rates of self-employment arguably make naps even more relevant in developing countries, since working adults often have substantial control over their own schedules. Naps have been studied in sleep labs, but we have little causal evidence on the impacts of naps on worker productivity and other real-world outcomes (Lovato and Lack, 2010; Ficca et al., 2010). Our work takes a step towards filling this gap.

Finally, recent research in behavioral and development economics argues that people in developing countries often under-invest in high-return investments such as preventive health, agricultural inputs, or capital investments (Kremer et al., 2019). At first glance, the low levels of sleep we discovered appear to tell the same story, and experts predicted substantial impacts of increased sleep. Instead, our evidence suggests that people do not under-invest in sleep duration given the environmental constraints that they face. The returns to increasing night sleep in their home environments are low and possibly even negative. To paraphrase Schultz (1964), our evidence suggests that low-income people in Chennai are poor, but efficiently tired.

2 Measuring Sleep in Chennai

2.1 Measuring Sleep Outside the Lab

The gold standard for objectively measuring sleep in labs 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 that may themselves interfere with sleep. These constraints have meant that measuring sleep outside of sleep laboratories has historically been challenging. Self-reported measures of sleep are notoriously unreliable and usually correlate only moderately with objective sleep measures. Individuals asked to report their sleep instead tend to report the hours spent in bed, leading to over-reporting of sleep duration (Lauderdale et al., 2008; Schokman et al., 2018).

Actigraphs, which resemble wristwatches and infer sleep/wake states from body movement, have recently emerged as a viable alternative for field studies. These devices allow researchers to objectively measure sleep in participants' home environments without themselves interfering with sleep, as they are portable, comfortable and unobtrusive. Validation studies show that actigraphs reliably measure sleep duration. Comparisons between actigraphy and PSG measures show high degrees of accuracy 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 duration even among individuals with sleep disorders (Kushida et al., 2001; Smith et al., 2018) and reliably capture treatment effects of various interventions on sleep (Sadeh, 2011).

Actigraphs also measure sleep efficiency, defined as time asleep divided by time in bed. This measure is available since—in addition to number of hours asleep—actigraphs also detect when an individual is in bed, but not asleep. Sleep efficiency is perhaps the most commonly-used proxy for sleep quality in sleep science (Ohayon et al., 2017). Disruptions to sleep, such as brief awakenings during the night, drive down sleep efficiency. In addition, sleep efficiency affects the opportunity cost of sleep, since it indicates the time in bed needed to achieve an hour of actual sleep.

2.2 Sleep Deprivation Around the World

While sleep scientists recommend 7 to 9 hours of sleep per night (Hirshkowitz et al., 2015; Watson et al., 2015), numerous studies show that people in high-income countries sleep less than this (Walker, 2017).[6] For instance, Lauderdale et al. (2008) measure sleep via actigraphy among a large, diverse population of healthy young adults in Chicago, and report an average sleep duration of 6.1 hours per night, well below the recommended range. In a large study seeking to understand sleep among older adults in the US, Jackson et al. (2018) observe an average weekday sleep duration of 6.4 hours per night.

In contrast, there is scant evidence on sleep patterns in developing countries. There are reasons to suspect sleep deprivation may be widespread and even more severe in the rapidly-growing cities of the developing world, where residential structures are often of low quality, and people are exposed to excessive heat, noise, crowding, and pollution – all conditions likely to hinder sleep.

Even self-reports of sleep from developing countries suggest a substantial share do not sleep enough. For example, the 4,500 rural, older Indian adults surveyed in Gildner et al. (2014) self-report 7.1 hours of sleep on average, with about 30% of these individuals reporting six or fewer hours per night (Selvamani et al., 2018). But self-reports, as discussed above, may overestimate sleep. Two recent studies in low-income countries have collected actigraph-based measurements of sleep. Both identify even larger fractions of the population as sleep deprived. In particular, Schokman et al. (2018) finds an average of only 6 hours per night among 175 adults from urban Sri Lanka. Knutson (2014) finds that 58 adults in Haiti sleep on average 7 hours per night, although this is a rural population without electricity.

2.3 Sleep(less) in Chennai

We first measured sleep in our RCT sample of 452 adults recruited to work in a full-time data-entry job for one month in Chennai, India.[7] In order to capture reliable and objective measures of sleep, all participants wore actigraphs continuously throughout the study.[8] In addition, we elicited daily self-reported measures of sleep. Below, we describe sleep during the baseline period (before treatment) in this RCT sample. We then report very similar patterns of sleep in a broader sample in Chennai, which wore actigraphs for three nights.

[6] This guideline refers to actual time asleep, not merely time in bed. However, sleep and time in bed are often similar in healthy rich-country populations, where sleep efficiency typically exceeds 85 percent (Jackson et al., 2018; Cespedes et al., 2016).

[7] Section 3 provides details on the sample and study design.

[8] Participants received a modest daily incentive of Rs. 10 to wear the actigraph, which they forfeited if they removed it. To determine whether the participants wore the devices continuously, a small breakable strap was put through the watch band and checked daily. Compliance rates were high across all experimental groups, with approximately 6% of participants removing the device on any given day.

A Typical Night in Chennai. Before providing a more systematic discussion of sleep in our sample, we provide an example to highlight key features of participants' sleep patterns. Figure Ia illustrates a typical night for a study participant, using minute-by-minute actigraph measures of sleep (gray) and wake (red) status. On such a night—which closely matches the average time in bed, time asleep, sleep efficiency and nightly awakenings seen in the sample—the participant spends about 8 hours in bed. However, despite spending 8 hours in bed, they sleep for only 5.6 hours. As can be seen from the interspersed grey and red bands, this sleep is highly fragmented and interrupted, with over 30 awakenings per night. For comparison, we show a less interrupted night of sleep with 90% efficiency in Figure Ib. While this night is unusual in Chennai – only 1% of nights in our sample have a sleep efficiency as high – it is similar to healthy adults in the US, who have a sleep efficiency of 86-89% on average (Cespedes et al., 2016; Jackson et al., 2018).

Time in Bed vs. Time Asleep. The RCT sample spends roughly 8 hours per night in bed before treatment begins, with strong congruence between actigraph measures (Figure IIa) and self-reports (Figure IIb). Time in bed in Chennai is quite similar to that found in US samples.[9]

Despite this significant time in bed, study participants only enjoy an average of 5.6 hours *asleep* per night (Table I column 1 and Figure IIc). This time asleep is significantly below both time in bed and the recommended 7 to 9 hours (Hirshkowitz et al., 2015). During the pre-treatment period, 95% of control participants slept less than 7 hours per night, and 71% slept less than 6 hours per night on average. This sleep duration is significantly lower than actigraph-measured sleep duration among similarly aged adults in high-income countries, which is typically 6 to 7 hours (Lauderdale et al., 2008; Jackson et al., 2018; Tworoger et al., 2005). In high-income countries, such low average time asleep is typical in populations with disorders such as sleep apnea (Cole et al., 1992; Kushida et al., 2001; Gershon et al., 2012).

Sleep Efficiency. RCT participants have an average baseline sleep efficiency of 70% (Figure IIe). Like sleep duration, this figure is much lower than estimates of sleep efficiency in high-income countries, which are typically 85 to 95 percent (Carrier et al., 2001; Cole et al., 1992; Walker, 2017). Sleep efficiency in our sample is instead similar to, or slightly below, US-based patients suffering from sleep disorders such as sleep apnea (Roure et al., 2008) or insomnia (Trauer et al., 2015). This low sleep efficiency is also far below recommended levels; an expert panel of sleep scientists found that a minimum of 85% is needed to indicate "high-quality" sleep (Ohayon et al., 2017). Sleep efficiency is low throughout the night, remaining around 70% between 1 and 5 am (when almost everyone is in bed), consistent with interrupted sleep throughout the night (Figure A.I). Participants experience about 32 awakenings on an average night (Table I column 1), again comparable to sleep-disordered populations such as insomniacs in the US (Lichstein et al., 2006).[10]

[9] Cespedes et al. (2016) reports an average of 7.8 hours in bed. Among older participants, Kurina et al. (2015) and Jackson et al. (2018) find 8.4 and 7.2 hours in bed, respectively.

[10] Considering longer awakenings which last for at least 5 minutes each, we still find an average of 10 such awakenings per night, compared to expert guidelines of 4 or fewer per night (Ohayon et al., 2017).

Barriers to Sleep. Why is sleep so inefficient? Survey responses highlight the importance of mental and physical distress (e.g. worries, hunger) as well as environmental factors. Over 50% of study participants indicate that cold or heat, noise, and/or light disrupt their sleep (Figure A.IV).

Napping. Naps are relatively common in this population. 73% of participants in our study reported taking at least one nap in the week before enrolling in the study. Conditional on napping, the median time reported for a nap is about one hour. The frequency and length of naps in US populations is fairly similar: Dinges (1992) find that, across a broad population of US adults, 61% report napping at least once a week with an average nap duration of 73 minutes, while in Pilcher et al. (2001), 74% of healthy adults report napping during a 7-day period.

Self-reported Sleep. Self-reports significantly overestimate time asleep, relative to the objective actigraph measures, consistent with findings that people tend to overestimate their sleep duration (Lauderdale et al., 2008; Avery et al., 2019).[11] Average baseline self-reported sleep duration in our study is 7.2 hours (Figure IId), quite similar to the average of 7.1 hours found in a representative survey of older adults in rural India described in Gildner et al. (2014). In comparison, average self-reported sleep duration in US ranges from 6.8 to 7.9 hours per night (Jackson et al., 2018; Lauderdale et al., 2008; Watson et al., 2015).

Broader Population. To investigate the representativeness of our RCT sample, we conducted a supplementary "Sleep Survey" with 3,833 individuals across randomly-sampled neighborhoods in Chennai. Details of the survey and population are described in Appendix F. A subset of 439 of the survey 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 duration and efficiency in this broader sample is quite similar to that of the participants in the RCT, with an average of 5.5 hours of sleep per night and 71% sleep efficiency (Table I column 2).

As in the RCT sample, napping is common, with 25% of participants napping on any given day in the broader population sample. Conditional on napping, the average duration of the nap as measured by the actigraph is roughly 50 minutes.

3 Experimental Design

Figure III provides an overview of the experimental design and timeline of the study. 452 participants worked for 28 days in an office in Chennai, spending most of their workdays doing paid data-entry work. Enrollment took place on a rolling basis between October 2017 and April 2019. 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.

[11]Despite the overestimation on average, self-reports are moderately correlated with actigraph measures at the individual level (r = .48). However, given that self-reported levels of sleep exceed actigraph measures, self-reports also overestimate sleep efficiency relative to actigraph measures (Figure IIf).

3.1 Interventions to Increase Sleep

For their first eight days in the study, participants remained in a control condition, allowing us to collect rich baseline data. Then, we cross-randomized participants to two night-sleep treatments and a nap treatment, stratified by baseline sleep and earnings.

Night-sleep Treatments.

Each participant was randomly assigned to one of two night-sleep treatment groups ('encouragement' or 'incentives') or to a control group in equal proportions.

1. *Devices + Encouragement:* This treatment involved a bundled intervention to increase night sleep. Individuals were offered: (i) information regarding the benefits of sleep (in particular, generic health benefits) and tips to improve their sleep (such as going to bed at the same time every day, avoiding caffeine after 4 pm and avoiding screens before bed), (ii) encouragement to increase their sleep as well as daily feedback on their night-sleep duration as measured by the actigraph, and (iii) loaned devices to improve their sleep environment. The offered devices included eye shades, earplugs, a cot, a mattress, sheets, pillows, and a fan (see Appendix Figure A.IIb).[12]

2. *Devices + Incentives:* This group received the same bundled intervention as the Devices + Encouragement group *plus* financial incentives to increase their actigraph-measured sleep during the treatment period. Each day, participants were paid Rs. 1 per minute of increased sleep for up to two hours of increased sleep (Rs. 120, about $1.70), relative to their baseline-period sleep. There was no penalty for sleeping less than in the baseline period.[13] To control for any income effects due to the sleep-incentive payments, participants in the control and encouragement groups were randomly and anonymously matched to participants in the incentives group and received the same stream of payments, *independent of their own sleep*.

3. *Control:* This group did not receive any of the above treatments. 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 to a subset of control participants. 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, and a flashlight. These goods were also returned at the end of the study.[14]

[12] Participants were permitted to take more than one of each device, as piloting had suggested that the devices were often shared with family members. They were asked to return—and penalized for not returning—the devices at the end of the study; virtually all complied.

[13] One concern is that participants could game the incentives by strategically reducing their baseline sleep. This is unlikely because participants did not know their treatment status during the baseline period. Also consistent with a lack of gaming, control group participants did not increase their sleep after treatment assignment and, as described in Section 2.3, baseline sleep is very similar to levels seen in the broader Sleep Survey.

[14] The use of placebo item offers to the control group was not randomized, and instead began about halfway through the experiment, after which all control group participants were offered these items. We find no detectable difference in treatment effects based on whether the control group had been offered these placebo goods, and thus pool all control participants in the analysis.

Given the difficulty of increasing sleep in the field, we took a bundled approach in designing our treatments, working to increase sleep through multiple channels. Participants could respond to the encouragement and financial incentives by spending more time in bed or by taking steps to increase their sleep efficiency. The tips to improve sleep, such as avoiding caffeine in the evening, turning off the television and putting away ones cellphone at night could also plausibly increase sleep efficiency. Finally, the loaned devices could plausibly increase both sleep efficiency and time in bed, if the devices made it easier to fall asleep or reduced awakenings at night, or if they made time spent in bed more enjoyable.

Nap Treatment. Motivated by existing lab evidence that naps can be effective in boosting cognitive function (Lovato and Lack, 2010) and can make up for limited night sleep (Mollicone et al., 2008), we cross-randomized the night-sleep treatments with a nap intervention. 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. Located in a quiet and gender-separated part of the study office, the 25 private nap spaces each included a bed, blanket, pillow, table fan, ear plugs, and eye shades (see Appendix Figure A.IIc). The actigraphs show that roughly 90% of study participants did indeed sleep during their allotted nap time. Those who did not want to nap were asked to sit quietly or rest in their nap area; they did not have the option to work during this time.

The remaining (non-nap) participants were randomized each day with equal probability to either a work group, in which we allowed them to work through the 'nap period', or a break group, in which we enforced a half-hour break from data entry during the same period. Break group participants were allowed to engage in any leisure activity they chose, including sitting in a comfortable office break room. By comparing the nap and break groups, we isolate the effect of a nap relative to a break of the same length. By instead comparing the nap and work groups, we can estimate the net effect of naps on work output, including the lost work time.[15]

3.2 Outcome Measures

Sleep and work are the two key sets of outcomes of this study. We measured each of them daily using actigraphs and the data-entry platform, respectively. Study participants also completed a series of short surveys and experimental tasks throughout the study (see Appendix Table A.III). Described in greater detail in Appendix C, these measures fall into three broad categories: 1) physical and mental well-being, 2) cognition, and 3) preferences.

Measures of Sleep. We measure night and nap sleep—sleep duration, time in bed, efficiency and interruptions—using actigraphs, as described in Section 2.1. 94% of participants wore their actigraph on any given day, balanced across treatments. We complement these measures with daily self-reports of time in bed, time asleep, and number of awakenings during the night.

[15] Of course, nap and break participants may adjust their hours in response to the treatment conditions. Since we observe work hours, we can estimate such labor supply effects.

Work-related Outcomes. Participants were engaged each day in data-entry work. We designed a software interface which presented participants with images containing alphanumeric text, and asked them to transcribe the data by typing into text boxes (see Figure A.III). The task was designed to mimic a real-world data-entry job.[16] Participants were paid for time spent typing as well as the amount of data entered, as described below. This design allows us to precisely measure labor supply, productivity, and earnings.

Labor Supply. Our pre-registered measure of labor supply is the *active* typing time as automatically measured by the data-entry software.[17] Participants could choose their labor supply freely. Most days, the participants could arrive or depart from the office as they chose between 9:30 am and 8 pm.[18] Even when in the office, participants were free to take breaks from data entry. We can precisely measure even short breaks: if a participant spent two consecutive minutes without typing, the software automatically paused and the break period did not count towards the labor-supply measure. Thus, participants had a great deal of control over their labor supply, except for time slots set aside for surveys, experimental tasks, and the lunch break.

Earnings. Earnings in the data-entry work is our pre-registered measure of performance at work, and used as our 'summary' measure of work since it combines labor supply and productivity. It has two components: an "attendance pay" per hour of active typing (one-third of work earnings) and a "performance pay", a piece rate for each correct character and a penalty per mistake (two-thirds of data-entry earnings). Each half hour, piece rates were randomly varied between a low value (Rs. 5 per 1,000 correct characters) and high value (4 times as large) with equal probability. The penalty rate remained constant throughout at Rs. 1 per 10 mistakes. The variation in piece rates allows us to benchmark any productivity effects of the sleep treatments against monetary incentives. The participants were paid daily, just before leaving the office for the day.[19]

Productivity. Our pre-registered measure of worker productivity is output divided by active typing hours. Output is the number of correct entries minus (a weighted) number of mistakes. The weight on mistakes was defined as the ratio of the average piece-rate and the penalty rate.

[16]The data to be digitized were actually artificially generated. By generating the data, we had ready access to the correct 'answers', allowing us to measure the accuracy of the work easily. We were also able to vary the complexity of the data to be entered across fields. Study participants were unaware of the artificial nature of the data, and we believe they had no reason to not take their work seriously.

[17]We also pre-registered total time at office as a measure of labor supply. The measures are highly correlated, and we focus on active time typing because it is the measure of labor supply an employer would care more about.

[18]On a subset of days ("short days"), work hours were limited from 11 am to 5 pm, in an effort to provide clean estimates of productivity, unconfounded by potential changes in labor supply. To encourage presence during these hours, we paid a bonus of Rs. 50 to anyone present during the entire period. Since there is only a small difference in labor supply across night-sleep treatment arms, we do not separate the analysis of work outcomes between short and long days.

[19]Control group participants earned Rs. 283 ($3.80) per day on average through their typing work (not including additional payments for surveys, experimental tasks and sleep incentives). For context, GDP per capita in Chennai is approximately $9 per day. The piece-rate accounted for 57% of typing earnings, while the remaining 43% was compensation for time spent typing.

Well-being. We collected a wide range of measures of psychological and physical well-being. As pre-registered, we examine these variables both as indices and individually. The pre-registered measures of *mental well-being* are happiness, sense of life possibilities (Cantril Scale), life satisfaction, stress, and depression. The measures of *physical well-being* are performance in a stationary biking task; reported days of illness; self-reported pain; activities of daily living; and blood pressure.[20]

Cognition. Sleep scientists have documented a strong relationship between sleep and cognition in numerous laboratory studies in rich countries (Lim and Dinges, 2010; Killgore, 2010). We collected (i) laboratory measures of cognitive function borrowed from cognitive psychology and sleep medicine; (ii) a measure of attention to incentives at work embedded in the data-entry task.

Lab Measures of Cognition. Each afternoon, participants completed the Psychomotor Vigilance Task (PVT), a standard measure of alertness and attention used in sleep medicine (Basner and Dinges, 2011). Every other day, they also completed cognitive tasks measuring memory (Corsi blocks task) and inhibitory control (Hearts and Flowers task), described in detail in Dean et al. (2019) and briefly in Appendix Section C.5. All cognitive tasks were incentivized with monetary payments for performance (e.g. speed, accuracy).

Attention to Work Incentives. To test whether sleep impacts the ability to attend to important aspects of one's work environment, e.g. the incentives faced, we randomized the visual salience of piece rates across days starting on day 6 of the baseline period. In the *salient condition*, the current piece-rate was highlighted in different colors for each rate and displayed on the screen at all times. We consider this condition the "full-attention" benchmark, as in Chetty et al. (2009). In the *non-salient condition*, noticing and remembering the piece-rate was more challenging. A single muted color was used for both piece-rates and the rate was only visible for the first 15 seconds of a half-hour slot, fading out slowly.[21] Figure A.III provides screenshots of each condition described below. The participant-level attention measure is the difference in average response to piece rate incentives in the full-attention benchmark and in the non-salient condition.[22]

Preferences. Sleep may impact preferences either through its impacts on cognition or directly. For instance, sleep has been hypothesized to play a critical role in replenishing self-control (Vohs

[20]Well-being related outcomes were pre-registered in Clinicaltrials.gov, Identifier: NCT03322358.

[21]The feature of the piece-rate display disappearing over time was added halfway through the study. We show results for the entire sample, although the effect of salience and the treatment effects of sleep on this measure of attention are larger for the second half of the sample.

[22]Formally, it is given by

$$A_i = \left(\overline{Y}_i(H,S) - \overline{Y}_i(L,S)\right) - \left(\overline{Y}_i(H,NS) - \overline{Y}_i(L,NS)\right)$$

where $\overline{Y}_i(H,S) - \overline{Y}_i(L,S)$ is the average difference of output under high and low piece rates of participant i when incentives are salient, and $\overline{Y}_i(H,NS) - \overline{Y}_i(L,NS)$ is the same for non-salient incentives. We residualize output with respect to participant, day in study, and date fixed effects.

and Baumeister, 2016) and sleep deprivation has been correlated with cyberloafing at work (Wagner et al., 2012) and cheating (Barnes et al., 2011). Similarly, sleep could alter the weight placed on sure things versus gambles or on others versus the self (Anderson and Dickinson, 2010; McKenna et al., 2007; Holbein et al., 2019). To examine such effects we study time preference via financial savings outcomes and choices on a real-effort task, and risk and social preferences via standard experimental-economics measures described below.

Savings. We measured savings behavior by providing participants an opportunity to save money in a lock-box at the study office, as in Schilbach (2019). At the end of each work day, after receiving their earnings, individuals had the opportunity to deposit or withdraw money. Participants were randomly assigned to receive *daily* interest rates between of 0 and 2% for any money saved in the box.[23] For participants receiving the positive interest rate, at least, the savings account we offered was quite lucrative. The deposits made in this account constitute our main savings outcome.[24]

Effort Discounting. We measured present bias using real-effort choices, following Augenblick and Rabin (2019). Participants made decisions about how many pages to type at the end of the day on a particular date under different piece rates. Using choices elicited both in advance and on the day of the work itself, we structurally estimate an individual-level present bias parameter β_i, once each in the baseline and treatment periods. A complete description of the task is in Appendix C.6.3.

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

3.3 Expert Predictions

To quantify how our results compare with existing scientific understanding, we conducted surveys of experts in sleep science and economics to elicit their prior beliefs about the treatment effects of this study (DellaVigna et al., 2019). Participation in the survey was solicited via emails to experts in each field. The survey provided information on the design of the study, the magnitude of the increase in night sleep reported in Section 4.2 below, and the outcome measures described above.[25] Three versions of the survey were tailored to different respondents: 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. In an effort to keep

[23] The interest rates changed twice during the study. Details are provided in Appendix D.

[24] 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 Section C.6.2.

[25] The expert surveys were conducted after over half the RCT sample had been acquired, in order to provide respondents with information on the achieved gains in sleep. However, the paper had not been publicly presented or circulated with results at that time.

the survey short, we did not elicit predictions about the effects of the nap treatment. All experts made predictions on labor-supply and work-output effects. Both types of economists responded with their beliefs about savings. 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 shown in Figure IV and in Table A.IV, and discussed when presenting results. Further details are provided in Appendix C.1.

3.4 Study Population and Balance Checks

We followed two strategies to recruit our study sample. First, recruiters went to low-income neighborhoods in Chennai and spread information about the study, advertising a one-month data-entry job. Second, past participants could refer potential new participants to the study. In both cases, recruiters interviewed individuals to determine their eligibility to participate in the study.

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

Importantly, this recruitment and screening procedure does not seem to select participants on average levels of sleep quantity and efficiency. In Table I, we find very similar patterns of sleep among individuals in Chennai in the broader Sleep Survey, as described in Appendix F.

Informed Consent. All participants went through a detailed informed-consent procedure. They were informed about the work task, the additional experimental measures and surveys, the actigraphs and the randomized treatments. The specific research hypotheses were not shared with participants to avoid demand effects. The goal of the research was described in more general terms as being to understand the "difficulties underprivileged people in India face, and how these problems affect their lives."

Sample Characteristics. Table A.I shows sample characteristics. A typical study participant was about 35 years old with 1 to 2 children and 10 years of education. Two-thirds of study participants were female. While only 30% of participants had prior computer experience, participants were eager to learn and improved rapidly in their data-entry speed during the baseline period.

Balance Checks. We test for baseline imbalances in demographics and baseline measures of outcome variables across the experimental conditions in Tables A.I and A.II. Whether we separately consider each treatment cell (Table A.I) or instead compare the pooled night-sleep treatment groups with the control and the nap and no-nap groups (Table A.II), the treatment groups were well-balanced across key characteristics. For each treatment arm, a joint F-test comparing it to the control group indicated 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, participants in the night-sleep treatment groups were about a year younger than those in the control group, and baseline productivity and earnings were about 3 to 4 percent lower in the nap group than in the no-nap groups (Table A.II). We control for age and for the participant-level baseline average of the outcome variables, so these imbalances should not affect our results.[26]

4 Experimental Results

4.1 Empirical Framework

Most of our empirical analyses, including all work-related outcomes, estimate treatment effects on outcomes measured at the *participant-day level* using variants of the following equation:

$$y_{itd} = \beta_1 T_i^{NS} + \beta_2 T_i^{Nap} + \gamma_1 \bar{y}_i^B + \gamma_2' X_{it} + \delta_t + \lambda_d + \varepsilon_{itd}, \tag{1}$$

where y_{itd} is the relevant outcome for participant i on her t^{th} day in the study on calendar date d. T_i^{NS} and T_i^{Nap} indicate whether the participant was assigned to one of the two night-sleep treatments and to the nap treatment, respectively. The average treatment effect of the night-sleep and nap treatments is captured by β_1 and β_2, respectively.

Following McKenzie (2012), we control for the average baseline value of the outcome variable $\bar{y}_i^B$ in all specifications, and drop the baseline days from the regression. We also drop days in which participants were absent, since attendance was balanced across groups. X_{it} includes controls for participants' age (quartiles) and gender. Where specified, it may also include a dummy variable for whether a given non-nap participant i was assigned to work through the nap period or instead to take an enforced break on day t. This allows us to compare the nap group separately with the work and break groups.[27] Finally, we include day-in-study and calendar-date fixed effects, captured by δ_t and δ_d, respectively. All standard errors are clustered at the participant-level.

For some outcomes, such as preferences, we only have one observation in the baseline and one in the treatment period per participant. In those cases, we run *participant-level* regressions:

$$y_i = \beta_1 T_i^{NS} + \beta_2 T_i^{Nap} + \gamma_1 y_i^B + \gamma_2' X_i + \varepsilon_i \tag{2}$$

[26] Note that we have an imbalance in earnings between the nap and no-nap groups in spite of stratifying on a dummy variable indicating whether the participant's baseline earnings was above the median. For this dummy variable, we have almost perfect balance, as expected. However, we have a few outliers with very large baseline earnings who all happened to be assigned to the no-nap group.

[27] For some outcomes, X_{it} includes additional outcome-specific controls. For example, in the work-related outcomes we additionally control for the fraction of the day worked at high piece rate (which was randomized each day) and the length of the work day (i.e., long or short day). Table notes specify these additional controls where they are used.

where again the outcome variable only uses the observations from the treatment period and the variable y_i^B is the baseline observation of the outcome variable. The vector X_i includes the same gender and age controls. This specification does not include day-in-study and calendar-date fixed effects. It also does not include an indicator for whether the non-nap participants worked or took a break, since this assignment varies on a daily level.

Pooling Treatments. In equations 1 and 2, we pool the two night-sleep treatment arms and do not include interactions between the night-sleep and nap treatments. The estimated treatment effects should thus be interpreted as weighted averages of treatment effects within the relevant cells. For instance, β_1 is the average effect of being assigned to one of the two night-sleep treatments (with equal probability), in a population which either receives naps or does not (with equal probabilities). The main reason we make these restrictions is to increase statistical power. We were not well powered to make comparisons across each individual treatment cell, with only about 75 participants per cell. Pooling the treatments also increases ease of exposition and interpretation, given the large number of outcomes and possible comparisons. In Appendix Table A.VII, we report estimates separating all treatments and interactions. We find little evidence of systematic differences in effects across the two different night-sleep treatments, or of interactions between the night-sleep and nap treatments.[28]

Combining Outcomes into Indices. Given the large number of outcomes, we divide them into four major families: work, well-being, cognition, and preferences. We then construct a single 'summary' outcome for each family. The work outcomes are naturally summarized by (standardized) earnings in the data-entry task, which combines productivity and labor supply into a single quantity. For the other families, we create standardized index variables. To do so, we residualize each constituent outcome with respect to day in study and calendar date, standardize by the control group's mean and standard deviation, and then take a weighted average to form the index. Following Anderson (2008), the weights are the inverse of the covariance matrix of the (residualized, standardized) outcomes. This ensures that outcomes which are highly correlated receive less weight than outcomes that capture new information. Signs of outcome variables are flipped when necessary so positive treatment effects imply an improvement in the outcome.[29] We also report treatment effects on an 'overall' index, which combines the four family-level summary outcomes into a single variable. We use the same procedure as above to create the overall index.

Multiple Hypothesis Testing. We report three approaches to dealing with the multiple hypothesis testing issues caused by observing many outcomes. Our simplest approach is to examine

[28] Muralidharan et al. (2019) point out that when researchers test the model with full interactions and then decide to pool treatments if the interaction terms are not significant, then standard inference does not apply anymore. However, the restrictions we imposed were based on lack of power. In fact, we only estimated effects separately by each treatment cell at the request of a referee. The inference in Tables III and IV therefore remains valid.

[29] This requires taking a normative stance on each variable. Some classifications are relatively uncontroversial: higher productivity and earnings, lower blood pressure and self-reported illness, higher cognitive function and more happiness are all classified as better. We also take the (more arguable) stances that greater patience (lower present bias), higher savings, higher labor supply and more prosocial behavior in lab experiments all constitute improvements.

a single overall index variable which combines *all* outcomes, as described above. Our intermediate approach is to consider outcomes at the level of the four families described above, using one index variable for each family, while applying multiple hypothesis corrections *across* the families of outcomes. Finally, at the level of the individual outcomes, we report p-values for each outcome that correct for the existence of multiple outcomes *within* each family. All corrections are calculated using simulations to control for the family-wise error rate. Details on our approach can be found in Appendix E. The corrected p-values are displayed along with our main results in Tables III and IV.

Pre-Analysis Plan (PAP). This study was pre-registered on both the AEA Registry and ClinicalTrials.gov, including a pre-analysis plan (PAP).[30] We deviate from from the PAP in some instances when further reflection, suggestions from editors, referees and other expert readers, new insights from recent research, or empirical results made us realize our pre-specified analyses were inappropriate or irrelevant. The main deviations (in our view) are the following. First, we pre-specified a regression model which included all interactions of treatments. We soon came to realize we were not well-powered for this analysis, and that it would lead to a large number of coefficients and comparisons which would be difficult to present and interpret. We therefore chose to pool the two night-sleep treatments and omit the interactions between nap and night-sleep treatments. This decision was not made based on empirical results, and we did not estimate the fully-separated model (presented in Table A.VII) until requested by a referee. Second, we had not fully specified our approach to multiple-hypothesis testing and made some changes after receiving comments and discussing with experts. We added the 'overall' index variable to parsimoniously aggregate all outcomes. We also redefined the four families of outcomes (work, well-being, cognition and preferences rather than work and decision-making) and created a summary variable for each family. Other smaller deviations are detailed in Appendix Section D.

4.2 Impacts on Sleep

Overview. It is possible to substantially increase night sleep in our highly sleep-deprived study population through the encouragement and incentive interventions. Offering short afternoon naps is effective in increasing daytime sleep.

Night-sleep Treatments. Both night-sleep treatments increased sleep markedly and immediately, as measured by the actigraphs (Figure V and Table II). On average, individuals in the *Devices+Encouragement* and *Devices+Incentives* treatment groups increased their time asleep at night by 20 and 33 minutes compared to the control group, respectively (Table II, Column 1). Pooling these two treatments, the night-sleep treatments on average increased night sleep by 27 minutes (s.e.=3 minutes). These increases are larger than the gains typically achieved by sleeping pills and cognitive behavioral therapy for insomnia (Riemann and Perlis, 2009; Trauer et al., 2015).

The increase in sleep was driven entirely by additional time in bed rather than improved sleep

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

efficiency. Both night-sleep treatment groups increased their time in bed significantly throughout the treatment period—31 minutes for the encouragement group and 46 minutes for the incentives group (Figure Va and Table II, Column 2).[31] We find no significant changes in sleep efficiency compared to the control group (Figure Vc and Table II, Column 3), even in the middle of the night when all participants are likely to be in bed (Appendix Figure A.Ia).

Increasing sleep duration is feasible for our study participants, simply through spending more time in bed. Participants also faced substantial incentives to improve their sleep efficiency: a participant who improved their sleep efficiency from 70% to 80% would earn on average Rs. 48 more each night in sleep incentives (holding time in bed fixed), which is about 20% of average typing earnings. Yet we see no effects on sleep efficiency. This suggests that changing sleep efficiency is relatively costly for participants, even with the aid of the loaned devices and tips surrounding sleep hygiene. Increasing sleep efficiency may require different and possibly more substantial or multi-faceted interventions than those we tested. The night-sleep treatments thus did increase sleep substantially, but did so at significant opportunity costs of time, since they required large increases in time in bed.

Nap Treatment. The nap intervention was effective at increasing participants' daytime sleep (Figure Vd). Nearly all participants in the nap treatment (92%) reported falling asleep during their allotted nap time. These reports are consistent with actigraph data which recorded that participants fell asleep in 93% of all nap sessions. The mean actigraph-recorded (unconditional) time asleep during the nap period was 14 minutes, and the median duration was 16 minutes (Figure A.VI).

We have suggestive evidence that the "quality" of nap sleep in the office is higher than that of night sleep and naps at home. For instance, sleep efficiency during naps in the office (85%) is higher than efficiency in night sleep (66%) and in naps at home (72%, similar to night sleep), if one excludes in all cases the time taken to first fall asleep.[32] The average number of awakenings per minute of sleep achieved is also lower for the office naps. Better sleep quality during naps in the office—compared to both naps and night sleep at home—is consistent with a more comfortable sleep environment in the office.

Interactions and Heterogeneity. We find only modest interactions between the night-sleep and nap treatments in terms of their effects on sleep. Participants randomized to the night-sleep treatments did not nap any less when offered a nap (Table II, Column 4). Those treated with naps did spend ten minutes less in bed at night and slept, on average, five minutes less per night (Table II, Columns 1 and 2). Given this modest crowd-out, both treatments increased total time asleep in 24 hours, although naps had a substantially smaller impact on total sleep (Table II, Column 5). Finally, the impact of the night-sleep treatments on sleep quantity and efficiency did not differ

[31] On average, night-sleep treatment participants went to bed 17 minutes earlier at night and got out of bed 25 minutes later in the morning (Table A.VI).

[32] To make these measures as comparable as possible, we calculate time in bed for naps (both in the office and at home) as beginning with the minute the participant is first detected to fall asleep. To obtain a comparable number for night sleep, we examine sleep efficiency during the first 15 minutes of night sleep, which is the approximate length of the office naps.

significantly by baseline sleep quantity or efficiency, nor by characteristics such as participants' sex, age, or baseline earnings (Table A.V). Nor did these baseline factors predict meaningful differences in nap duration for the nap treatment group. The treatments thus seem to have been equally effective at increasing sleep (and leaving efficiency unchanged) for different categories of participants.

4.3 Impact of Night-Sleep Treatments

Overview. Experts from sleep science and economics predicted that increased night sleep due to our treatments would result in increased work output, labor supply, financial savings, cognitive function, and health (Figure IV). In contrast to these predictions and an influential literature in sleep science, we find no effect of the night-sleep treatments on our overall outcome index, or on any of the four summary variables corresponding to the work, well-being, cognition, and preferences outcome families. Considering each individual outcome separately, we find no significant positive effects on *any* outcomes (Figure VI and Table III). Instead, increases in night sleep come at the cost of significantly reduced labor supply and therefore a marginally significant reduction in work output. Below, we describe these findings in more detail.

Work Outcomes. The night-sleep treatments did not cause significant improvements in productivity, labor supply, output, or earnings (Table III). While the night-sleep treatment groups were 1.3% (0.02 SD, s.e.=0.02) more productive than the control group (column 3), this difference is not statistically significant even without multiple hypothesis testing corrections.

The night-sleep treatments *reduced* labor supply by approximately 10 minutes per day (0.08 SD with s.e.=0.02, column 3) with no change in the number of days worked (Figure A.VIII). This effect on labor supply remains significant at the 1% level when correcting for multiple outcomes within the work family. Given the additional time in bed induced by the night-sleep treatments, participants have less time available for work and leisure, which comes at the cost of reduced labor supply. Specifically, participants arrive at work 6 minutes later in the morning, take 2 minutes more of breaks at work, and leave for home 3 minutes earlier on average (Table A.VI). While obvious ex post, the opportunity costs of increasing sleep are typically neglected in the sleep literature. Indeed, the mean expert prediction was an *increase* in labor supply of 7%, which is strongly rejected by the data ($p<0.0001$).

The small increase in productivity was not enough to outweigh the reduction in labor supply, leading to a small and marginally significant decrease in earnings and output, respectively (each 0.04 SD with s.e.=0.02). This finding is again in contrast to the mean (median) expert prediction of a 12% (7%) increase in output. The discrepancy can in part be explained by experts overestimating the productivity impacts of increased sleep, and in part by their mispredicting that more sleep would increase the time allotted to work. 83% of experts made point predictions outside of the 95% confidence interval of our estimate of the effect on output.

Our estimates also contrast with those from natural experiments studying the economic consequences of sleep. Gibson and Shrader (2018) exploit variation in sunset times in the United States and estimate that 8.5 minutes of additional sleep per night increases earnings by 1.1% in the short run. Giuntella and Mazzonna (2019) use time-zone border discontinuities in the United States and find that 19 fewer minutes of sleep are associated with 3% lower earnings. Extrapolating these estimates linearly to our experiment would predict 3.5% and 4.3% increases in earnings, respectively, which we firmly reject ($p<0.01$). We return to discussing these differences in Section 5.

Well-being. Increased night sleep did not improve physical or mental well-being (Table III, columns 6 to 8). We find no significant impacts of increased night sleep on an index of physical well-being (Table III, Column 7) which combines objective and self-reported measures of health status. We do find positive (but not significant) point estimates for some of the underlying components such as performance in a cycling task and self-reported illness, pain, and daily activity (Table A.IX). Of course, three weeks is a short time for effects on physical health and behaviors to emerge. It could be that a longer intervention would generate health improvements in line with the observational literature (Strine and Chapman, 2005; Gottlieb et al., 2006; Cappuccio et al., 2008; Grandner et al., 2010; Giuntella and Mazzonna, 2019).

Similarly, we find no positive impact on the index variable which combines the various measures of psychological well-being, in contrast to a largely observational literature that finds associations between self-reported sleep duration or quality and psychological well-being (Kahneman and Krueger, 2006; Hamilton et al., 2007; Zhang et al., 2017). In fact, the point estimate is negative (-0.05 SD with s.e.=0.06, Table III, column 8). The individual estimates of this measure again show no significant positive effects on any component (Table A.VIII).

Cognition. In stark contrast to a large number of laboratory experiments in sleep science, we find no evidence of increased night sleep affecting cognition (Table III, columns 9 to 11). There is no significant effect on an index variable comprising of laboratory measures of simple attention, memory, and inhibitory control which closely mimic outcomes used in sleep laboratory studies (Table III, Column 10). The individual components are reported in Table A.X. Lab evidence shows that inducing sleep deprivation—typically by keeping participants up all night—substantially worsens performance on these tasks (see Lim and Dinges (2010) and Killgore (2010) for reviews). The more modest but sustained and policy-relevant increases in sleep we achieve do not have a corresponding positive effect.

We also find no evidence of impacts on attention measured in a more economic domain: how much people react to salient versus non-salient incentives (Table III, column 11). Consistent with limited attention, participants in the control group reacted 16% *less* to high incentives when piece-rates were *non*-salient (Table A.XI, column 1). Increased night sleep did not close this gap between responses to salient and non-salient incentives.

Preferences. Consistent with the lack of positive impacts of increased night sleep on any of the above outcomes, we find no evidence of the night-sleep treatments affecting any of the measures of time, risk, or social preferences or on the index which combines these outcomes (Table III, columns 12 to 15).

We detect no significant effect on an index variable which combines two measures of time preferences: savings and present bias (Table III, Column 13). The night-sleep treatments did not meaningfully affect savings behavior, leaving deposits and accumulated interest unaffected (Table A.XII, Panel A). Similarly, we find no evidence of impacts on the measure of present bias estimated from effort choices. We do detect significant present bias in the control group ($\beta = 0.92$, Table A.XII, Panel B, column 1). But increased sleep does not shift this parameter, in contrast to the view that sleep replenishes self-control (Vohs and Baumeister, 2016).

Similarly, we find no evidence of altered risk aversion, loss aversion, or social preferences in standard experimental tasks (Table A.XIII), in contrast to the findings of McKenna et al. (2007), Dickinson and McElroy (2017), Anderson and Dickinson (2010) and Holbein et al. (2019). While the results are not precise enough to detect small effects, we are able to rule out changes greater than 0.16 SD at the 95% level for each of these outcomes.

Overall Index. As described in Section 4.1, we combine all the outcomes described above into a single, standardized index variable following Anderson (2008). Table III, Column 1 reports that the night-sleep treatments have no effect on this overall index (-0.01 SD, s.e.=0.04).

Heterogeneity and Interactions. We do not find significant evidence of heterogeneity in the effects of the night-sleep treatments. Table A.XIV considers effects on the overall index variable, and tests for heterogeneity by a number of baseline covariates. Baseline night-sleep duration, sleep efficiency or propensity to nap prior to the study do not interact significantly with the night-sleep treatments. Nor do demographics such as gender and age. Note that this does not provide strong evidence that sleep efficiency and baseline sleep duration are irrelevant for the marginal benefits of increased sleep. We have limited statistical power for heterogeneity analysis, and very few study participants have levels of sleep efficiency typically observed in high-income countries.

Table A.VII separates the treatment effects for the two types of night-sleep treatments (incentives and encouragement) and their interactions with the nap treatment. Neither of the night-sleep treatments have significant effects on the overall index by themselves or in combination with naps, nor are there clear differences in their effects (Column 1).

4.4 Impact of Naps

Overview. In contrast to night sleep, the nap treatment improved outcomes across a range of domains (Figure IV and Table IV). Given the lack of evidence on the impacts of naps on economically meaningful outcomes in real-world settings, this is an important result in itself. In addition, these results serve as a proof of concept that sleep *can* significantly alter many of the outcomes we study within a short time frame.

Overall Index. We begin by considering effects on the overall index, which parsimoniously aggregates all our outcomes. Table IV, column 1 shows that the nap treatment had a positive, economically meaningful and statistically significant effect on the overall index (+0.12 SD, s.e.=0.04). The effect is significant whether naps are compared to taking enforced breaks (+0.15 SD) or to working through the afternoon (+0.08 SD). Despite napping only 14 minutes on average, and increasing 24-hour sleep by only a third as much as the night-sleep treatments, naps have a significantly (p=0.02) larger effect than the night-sleep treatments on the overall index. Below, we delve into each family of the underlying outcomes.

Work Outcomes. Naps increased productivity. Participants randomized to naps were 2.3% (0.04 SD, s.e.=0.02) more productive on average across the day (Table IV, column 3). This effect remains statistically significant when correcting for multiple tests within the work family ($p = 0.07$). The effects of naps are similar when compared to the break or the work counterfactuals, suggesting that the productivity impacts are due to nap sleep rather than merely resting. The impacts are sizable, given that productivity itself is quite inelastic: quadrupling the piece rate increased productivity by only 14%. Figure A.IX shows the effects of naps over the course of the work day. The effects are large and significant in the afternoon (+2.7%, p=0.01), as might be expected. However, we also observe significant improvements in the morning before naps, although point estimates are slightly smaller (+1.9%, p=0.05). This suggests either cumulative effects over time of regular napping or that participants work harder in the morning in anticipation of the nap.[33]

By design, nap participants had a 30-minute period during which they could not work. Individuals were free to adjust their labor supply outside of this period, but we find no evidence of such adjustments (Table IV, column 4). The nap treatment group spent 26 fewer minutes (0.20 SD) entering data compared to the group that could work through that period. Similarly, labor supply in the nap treatment group was nearly identical to labor supply in the break group.

The impact of naps on output and earnings depends on the comparison group. Compared to taking a break, naps increased total output by 0.05 SD (s.e.=0.02, $p = 0.02$). Comparing to working, naps instead reduce output by 0.07 SD (s.e.=0.02, $p < 0.01$). Earnings closely track output: naps increased overall earnings by about Rs. 11 per day (0.05 SD, s.e.=0.02) compared to taking a break, a sizable increase of 4.1% (Table IV, column 2). Earnings is our summary variable for the work family, and this effect remains significant when correcting across the multiple families ($p = 0.06$). However, taking time to nap lowered earnings by Rs. 23 (8.3% or 0.10 SD, s.e.=0.02) compared to simply working through the break.

The negative effect of naps on earnings (compared to the work group) appears to diminish over time. During "regular" work days, when participants are not restricted to artificially short work hours, the nap treatment group appears to converge with the work group in terms of earnings towards the end of the study (Figure A.VIIa). This effect is partly driven by the nap treatment group increasing their work hours over time (Figure A.VIIc). Nap participants—who are prevented from

[33] Also evident is a brief dip in productivity in the half hour immediately following the nap. This is consistent with the well-documented phenomenon of temporary 'sleep inertia' after a nap (Lovato and Lack, 2010).

typing for 30 minutes in the afternoon—typed 39 minutes per day less than the work group in the first week of treatment, but only 14 minutes less by the end of the study. They also appear to become more productive over time, although these estimates are noisier (Figure A.VIIb).

Well-being. Naps significantly improved the well-being index by 0.08 SD (s.e.=0.03, $p = 0.03$), as shown in Table IV, column 6. This effect is driven by substantial positive effects on mental well-being. The nap treatment increased the index of mental well-being by 0.12 SD (s.e.=0.05, $p = 0.04$; Table IV, column 8). The estimates for most individual components of mental well-being are positive, with a significant effect on happiness and sense of life possibility (Table A.VIII).

In contrast, we find little evidence of naps impacting physical well-being (Table IV, column 7). This lack of impact is unsurprising, given the limited impacts of naps on overall sleep and the fact that the physical health benefits of sleep may require more time to emerge. In addition, it is worth noting that some of the physical well-being outcomes, such as the cycling task, were conducted at the end of the study on a day without naps.

Cognition. Naps boost our index of cognition by 0.10 SD (s.e.=0.05, $p = 0.09$; Table IV, column 9). Delving further, we find that naps boost the lab measures of cognitive function by 0.08 SD (s.e.=0.04, $p = 0.08$; column 10). The cognitive benefits of naps appear to be concentrated on attention (Table A.X). We find no significant evidence of impacts on memory or inhibitory control, in contrast to the sleep literature that tends to find broad impacts of sleep on many aspects of cognition (Killgore, 2010; Lim and Dinges, 2010).[34]

Consistent with the impacts on laboratory measures of attention, we have suggestive evidence that naps also increased participants' attention to work incentives (Table IV, column 11). The nap treatment group was nearly fully attentive to non-salient incentives, reacting to them about as much as they reacted to salient incentives (Table A.XI). This illustrates the improved attentional resources provided by naps in a real-world work environment.

Preferences. Naps have a positive but not significant effect (+0.07 SD, s.e.=0.04) on the index variable corresponding to the preferences family (Table IV, column 12). Within this family, we do find suggestive evidence that naps increase an index of patience by +0.13 SD (s.e.=0.06, $p = 0.12$, Table IV, column 13). This index combines two real-stakes measures of time preferences: savings at the study office and present bias in an effort-discounting task. The nap treatment caused a 14% increase in daily deposits and a 13% increase in daily net savings (deposits minus withdrawals, columns 1 and 2 of Table A.XII, Panel A). These effects are sizable: for comparison, randomly providing a 1 percentage point higher daily interest rate increased deposits by 31% (Table A.XII). Participants in the nap treatment group therefore earned 19% more interest over the course of the study (Table A.XII, Panel A, column 3), although this effect is imprecisely estimated.[35]

[34] We find suggestive evidence that naps may have improved reaction times in the inhibitory control task (0.14 SD). However, this change was not large enough to impact overall payments for performance on the task.

[35] We pre-registered daily net savings as our main variable of interest. However, this measure suffers from an unan-

Naps also reduce present bias in a real-effort task (Table A.XII, Panel B). We estimate an average present bias parameter of $\beta = 0.92$ in the control group.[36] The nap treatment significantly reduces present bias to $\beta = 0.98$, and time preferences in the nap group are statistically indistinguishable from exponential discounting (i.e. $\beta = 1$).

Naps have no significant effects on social and risk preferences (Table IV, columns 14 and 15). The nap group is not more willing to accept risk or probabilistic losses (Table A.XIII). The same tables reports that nap participants send 0.16 SD more (s.e.=0.10) in dictator games, but not in ultimatum or trust games. Nor do naps affect the choices of receivers in ultimatum or trust games.

Heterogeneity and Interactions. Table A.XIV tests for heterogeneous treatment effects of naps on the overall index. As with the night-sleep treatments, we do not detect statistically significant heterogeneity in the effects of naps by baseline sleep characteristics—including self-reported propensity to nap—or by demographic characteristics such as gender and age.

Table A.VII includes the interactions of the nap treatment with the two night-sleep treatments. The effect on the overall index appears to be driven by those who received the nap alone, rather than those who also received a night-sleep treatment. However, these differences are relatively imprecise and not statistically significant, and we urge caution in their interpretation.

5 Discussion and Conclusion

Using state-of-the-art objective measurement devices, this paper documents a novel fact about sleep in India: low-income adults in Chennai are severely sleep-deprived by usual standards. The strikingly low duration and efficiency of sleep in our sample could be a widespread but underappreciated feature of the lives of the urban poor in developing countries. More systematic research on sleep in developing countries is needed to add to the handful of existing studies from such contexts.

In our setting, substantial increases in sleep duration were achievable through more time in bed, a change which at least in principle lies within people's choice sets. Increasing sleep efficiency, however, appears to be more difficult. Providing people with tips regarding good sleep hygiene and devices to make their sleep environment more comfortable did not increase efficiency; nor did incentives to achieve more actual sleep (which substantially reward higher sleep efficiency). As a result, increasing sleep duration entailed significant opportunity costs for our study sample.

We find no positive impacts of increased night sleep on any of the outcomes we study. This result contrasts with predictions made by sleep scientists and economists. It also contrasts with many lab experiments in sleep medicine (e.g. Van Dongen et al., 2003; Lim and Dinges, 2010; Killgore, 2010), and a much smaller body of recent work in economics which uses natural experiments (such as variation in sunset time) to study the effects of sleep (Gibson and Shrader, 2018; Giuntella and

ticipated design issue: participants make large one-time withdrawals right before the study ends, which mechanically drives down net savings. We believe deposits more accurately reflect differences in savings behavior, and the accrued interest captures the benefit of savings.

[36] 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.XV).

Mazzonna, 2019; Jagnani, 2018). It is more consistent with some recent evidence from the field: Avery et al. (2019) find only small gains in academic achievement from inducing college students to increase nighttime sleep.

What explains this unexpected finding? One plausible explanation is the much lower sleep quality—as proxied by sleep efficiency—in our setting compared to those previously studied in rich countries. The low-quality sleep we discovered in Chennai may not offer the same marginal benefits as the sleep typically available in higher-income settings. A more provocative possibility is that the findings from lab studies may not generalize to the field, even in rich countries. The lab experiments used in sleep science induce severe (often total) sleep deprivation (e.g. Lim and Dinges, 2010) and typically lack steep incentives to perform well on tasks. We study the impacts of a more modest and arguably policy-relevant *increase* in sleep on highly incentivized outcomes.

As one of the first studies about the causal impacts of sleep in low-income settings, our experiment is not designed to adjudicate these different possible explanations. It does, however, point to the value of an economic perspective on sleep, which considers sleep as a choice variable, and measures both the benefits and the opportunity costs of sleeping more. Our findings regarding night sleep are consistent with optimizing economic actors: for a given sleep efficiency, people can choose their time in bed while trading off the costs and benefits, and should not leave high-return sleep unexploited. Indeed, we find that they do not. The low quantity of sleep in our setting is therefore *not* another example of unexploited high returns to investment in developing countries in domains such as firm capital (De Mel et al., 2008), fertilizer use (Duflo et al., 2008), deworming (Baird et al., 2016) and water treatment (Kremer et al., 2011).

However, our results do *not* imply that more dramatic changes in sleep environments (e.g. improved housing, noise regulations) or in psychological factors hindering sleep (such as stress) could not have large effects. Improving sleep quality could potentially generate both more sleep (due to higher sleep efficiency) and higher benefits from each minute of sleep. Identifying interventions to improve sleep efficiency in contexts like ours, and testing whether increased sleep efficiency unlocks the benefits found in sleep research in rich countries would also be valuable. It could also be that the benefits of increased night sleep manifest over longer time horizons, perhaps especially for physical health. Consistent with the hypothesis that increased sleep *can* have meaningful effects, we find the nap treatment has a significant positive effect on an overall index of outcomes, with positive effects on productivity, well-being and cognition.

The positive impact of naps is an important finding in itself. Naps are a common feature of life around the world, and even more so in tropical settings, where the opportunity cost of foregone work may be lower in the afternoon (Dinges, 1992). Yet we know very little about the impact of naps on economic outcomes in real-world settings (Ficca et al., 2010). Our findings suggest that naps have numerous benefits for workers. However, given the foregone work time and logistical costs of offering naps, an employer's decision to provide naps would depend on how much they value their workers' psychological well-being and how important attention is in the particular work setting. Of course, a majority of workers in countries like India are themselves self-employed, giving them some flexibility to incorporate naps into their schedules.

An obvious question raised by our results is why naps were effective when a larger increase in night sleep had no effect. One difference is that naps occurred in a more pleasant office environment, and may therefore have been higher quality than night sleep. Alternatively, naps may simply enter the 'production function' of our outcomes differently than marginal increases in night sleep, due to their timing. Naps were 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; Lovato and Lack, 2010). Finally, some outcomes such as the cognitive tests were measured in the afternoon, and were therefore closer in time to the nap sleep.[37]

Perhaps the broadest question our research raises is to what extent the sizable impacts of sleep from lab studies generalize to field settings. Measuring sleep in the field—made possible by wearable devices—will allow more such experiments linking sleep to real-world economic outcomes. It will also be important to understand the importance of sleep in a wider range of environments and tasks. We evaluated labor-market effects of sleep in the particular context of data-entry work. Does sleep matter differently in less cognitively-demanding or instead more physical jobs? In contexts where work is independent or collaborative? For children, as in Jagnani (2018)? Are naps in usual environments equally effective? Systematic measurement and tests across a wider variety of environments would facilitate the estimation of the overall impacts of sleep on the economy.

[37]However, we do not find effects of the night-sleep treatments on work productivity even in the mornings. And we find suggestive effects of the nap treatment on outcomes such as productivity even in the morning, before naps.

References

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.

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

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel, "Worms at work: Long-run impacts of a child health investment," *The quarterly journal of economics*, 2016, *131* (4), 1637–1680.

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

Barnes, Christopher M, Jared A Miller, and Sophie Bostock, "Helping employees sleep well: effects of cognitive behavioral therapy for insomnia on work outcomes.," *Journal of Applied Psychology*, 2017, *102* (1), 104.

_ , John Schaubroeck, Megan Huth, and Sonia Ghumman, "Lack of sleep and unethical conduct," *Organizational Behavior and Human Decision Processes*, 2011, *115* (2), 169–180.

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

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

Cappuccio, Francesco P, Frances M Taggart, Ngianga-Bakwin Kandala, Andrew Currie, Ed Peile, Saverio Stranges, and Michelle A Miller, "Meta-analysis of short sleep duration and obesity in children and adults," *Sleep*, 2008, *31* (5), 619–626.

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.

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 Christopher B. Barrett, Michael R. Carter, and Jean-Paul Chavas, eds., *The Economics of Poverty Traps*, NBER/University of Chicago Press, 2019, chapter 2, pp. 57–118.

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.

Duflo, Esther, Michael Kremer, and Jonathan Robinson, "How high are rates of return to fertilizer? Evidence from field experiments in Kenya," *American economic review*, 2008, *98* (2), 482–88.

Ficca, Gianluca, John Axelsson, Daniel Mollicone, Vincezo Muto, and Michael Vitiello, "Naps, Cognition, and Performance," *Sleep Medicine Reviews*, 2010, *14*, 240–258.

Gershon, Anda, Wesley K Thompson, Polina Eidelman, Eleanor L McGlinchey, Katherine A Kaplan, and Allison G Harvey, "Restless pillow, ruffled mind: Sleep and affect coupling in interepisode bipolar disorder.," *Journal of abnormal psychology*, 2012, *121* (4), 863.

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.

Giuntella, Osea and Fabrizio Mazzonna, "Sunset time and the economic effects of social jetlag: evidence from US time zone borders," *Journal of health economics*, 2019, *65*, 210–226.

Gottlieb, Daniel J, Susan Redline, F Javier Nieto, Carol M Baldwin, Anne B Newman, Helaine E Resnick, and Naresh M Punjabi, "Association of usual sleep duration with hypertension: the Sleep Heart Health Study," *Sleep*, 2006, *29* (8), 1009–1014.

Grandner, Michael A, Lauren Hale, Melisa Moore, and Nirav P Patel, "Mortality associated with short sleep duration: the evidence, the possible mechanisms, and the future," *Sleep medicine reviews*, 2010, *14* (3), 191–203.

Hamilton, Nancy A, CA Nelson, N Stevens, and Heather Kitzman, "Sleep and psychological well-being," *Social Indicators Research*, 2007, *82* (1), 147–163.

Hedner, Jan, Giora Pillar, Stephen D Pittman, Ding Zou, Ludger Grote, and David P White, "A novel adaptive wrist actigraphy algorithm for sleep-wake assessment in sleep apnea patients," *Sleep*, 2004, *27* (8), 1560–1566.

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.

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.*

_ **, Jessica Leino, Edward Miguel, and Alix Peterson Zwane**, "Spring cleaning: Rural water impacts, valuation, and property rights institutions," *The Quarterly Journal of Economics*, 2011, *126* (1), 145–205.

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.

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.

McKenna, Benjamin S, David L Dickinson, Henry J Orff, and Sean PA 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.

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "Returns to capital in microenterprises: evidence from a field experiment," *The quarterly journal of Economics*, 2008, *123* (4), 1329–1372.

Mollicone, D, H Van Dongen, and D Dinges, "Optimizing Sleep/Wake Schedules in Space: Sleep During Chronic Nocturnal Sleep Restriction with and without Diurnal Naps," *Acta Astronautica*, 2007, *60*, 354–361.

_ , _ , N Rogers, and D Dinges, "Response Surface Mapping of Neurobehavioral Performance: Testing the Feasibility of Split-Sleep Schedules for Space Operations," *Acta Astronautica*, 2008, *63*, 833–840.

Muralidharan, Karthik, Mauricio Romero, and Kaspar Wüthrich, "Factorial designs, model selection, and (incorrect) inference in randomized experiments," Technical Report, National Bureau of Economic Research 2019.

Nicholson, A, P Pascoe, T Roehrs, T Roth, M Spencer, and M Stone, "Sustained Performance with Short Evening and Morning Sleeps," *Aviation, Space, and Environmental Medicine*, 1985, *56*, 105–114.

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.

Pilcher, J., K. Michalowski, and R. Carrigan, "The Prevalence of Daytime Napping and its Relationship to Nighttime Sleep," *Behavioral Medicine*, 2001, *27* (2), 71–76.

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.

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.

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.

Schultz, Theodore William, "Transforming traditional agriculture.," *Transforming traditional agriculture.*, 1964.

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.

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.

Trauer, James M, Mary Y Qian, Joseph S Doyle, Shantha MW Rajaratnam, and David Cunnington, "Cognitive behavioral therapy for chronic insomnia: a systematic review and meta-analysis," *Annals of internal medicine*, 2015, *163* (3), 191–204.

Tworoger, Shelley S, Scott Davis, Michael V Vitiello, Martha J Lentz, and Anne McTiernan, "Factors associated with objective (actigraphic) and subjective sleep quality in young adult women," *Journal of psychosomatic research*, 2005, *59* (1), 11–19.

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.

Vohs, Kathleen D and Roy F Baumeister, *Handbook of self-regulation: Research, theory, and applications*, Guilford Publications, 2016.

Wagner, David T, Christopher M Barnes, Vivien KG Lim, and D Lance Ferris, "Lost sleep and cyberloafing: Evidence from the laboratory and a daylight saving time quasi-experiment.," *Journal of Applied Psychology*, 2012, *97* (5), 1068.

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

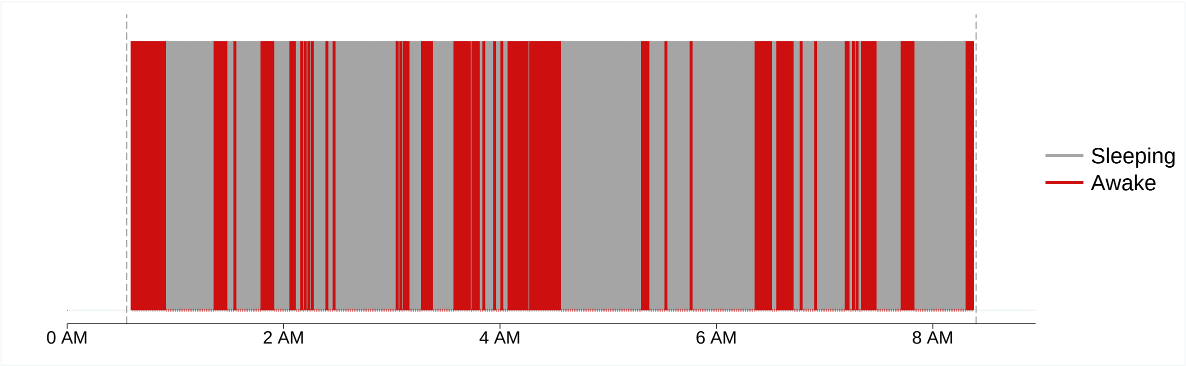
Watson, Nathaniel F., Safwan M. Badr, Gregory Belenky, Donald L. Bliwise, Orfeu M. Buxton, Daniel Buysse, David F. Dinges, James Gangwisch, Michael A. Grandner, Clete Kushida, Raman K. Malhotra, Jennifer L. Martin, Sanjay R. Patel, Stuart F. Quan, and Esra Tasali, "Joint Consensus Statement of the American 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.

Zhang, Jihui, Diana Paksarian, Femke Lamers, Ian B Hickie, Jianping He, and Kathleen Ries Merikangas, "Sleep patterns and mental health correlates in US adolescents," *The Journal of Pediatrics*, 2017, *182*, 137–143.

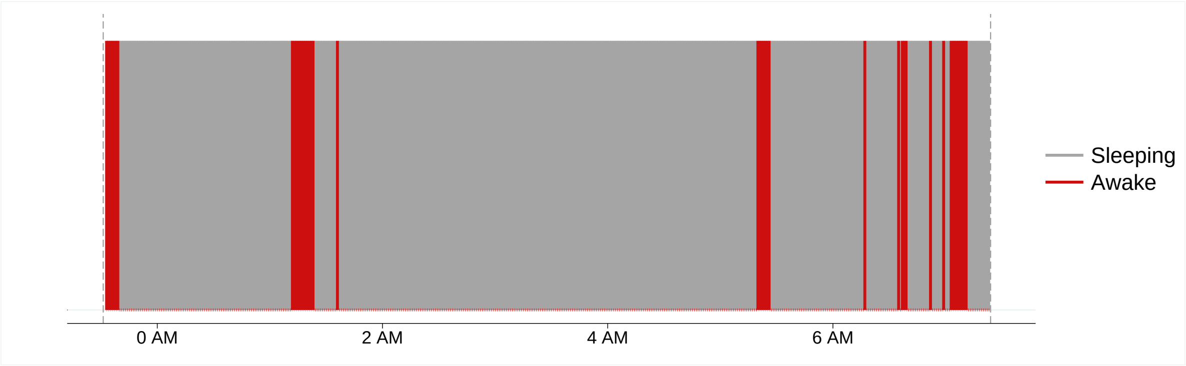
Tables and Figures

FIGURE I: Typical Sleep in Chennai

(a) A Typical Night



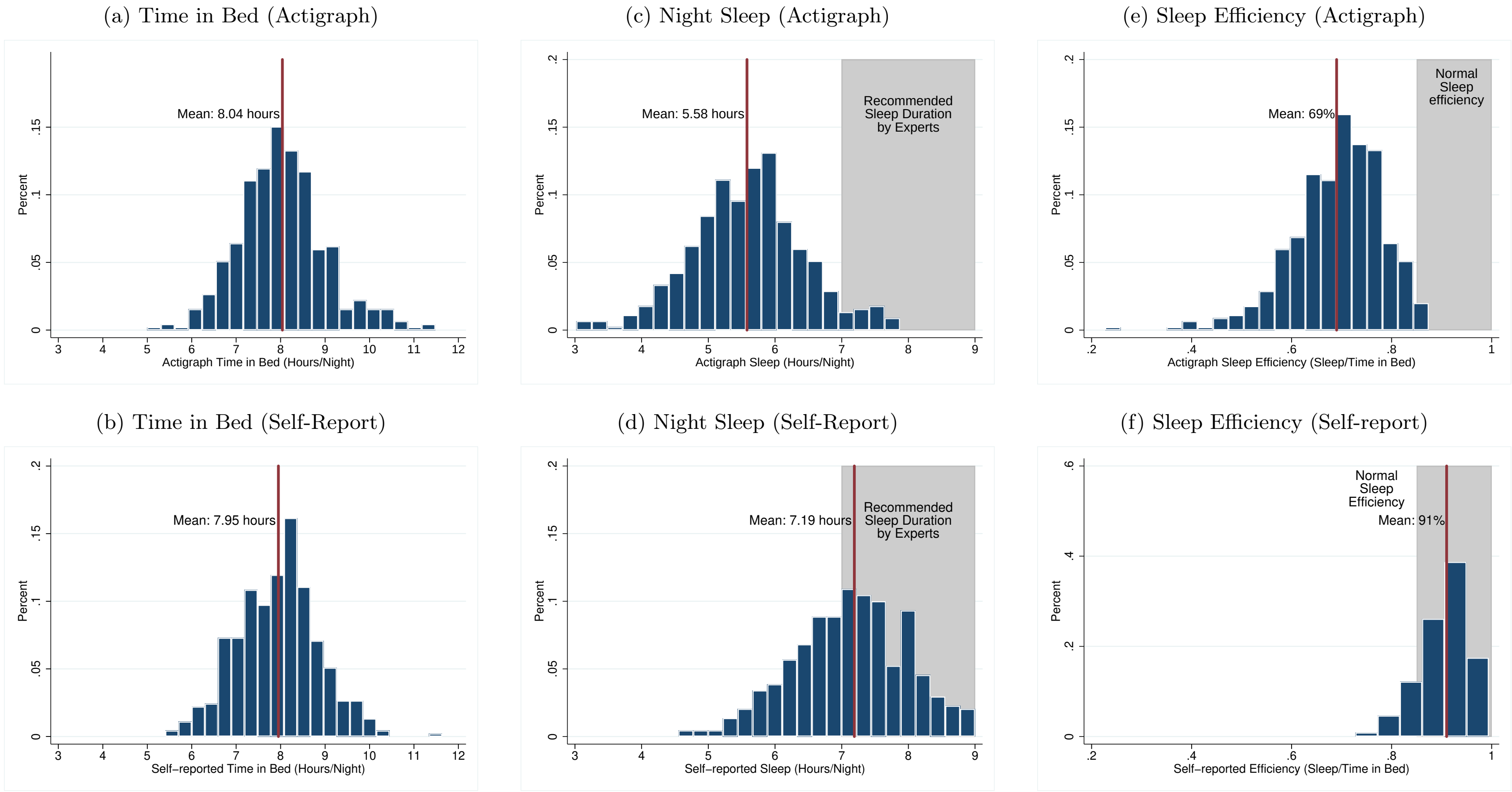
(b) A Good Night



Notes: This figure represents night sleep patterns of two participants in the study. Gray areas indicate one-minute periods in which the participant is asleep and red areas indicate periods in which the participant is awake. The light gray dashed lines indicate when the participant gets into or out of bed.

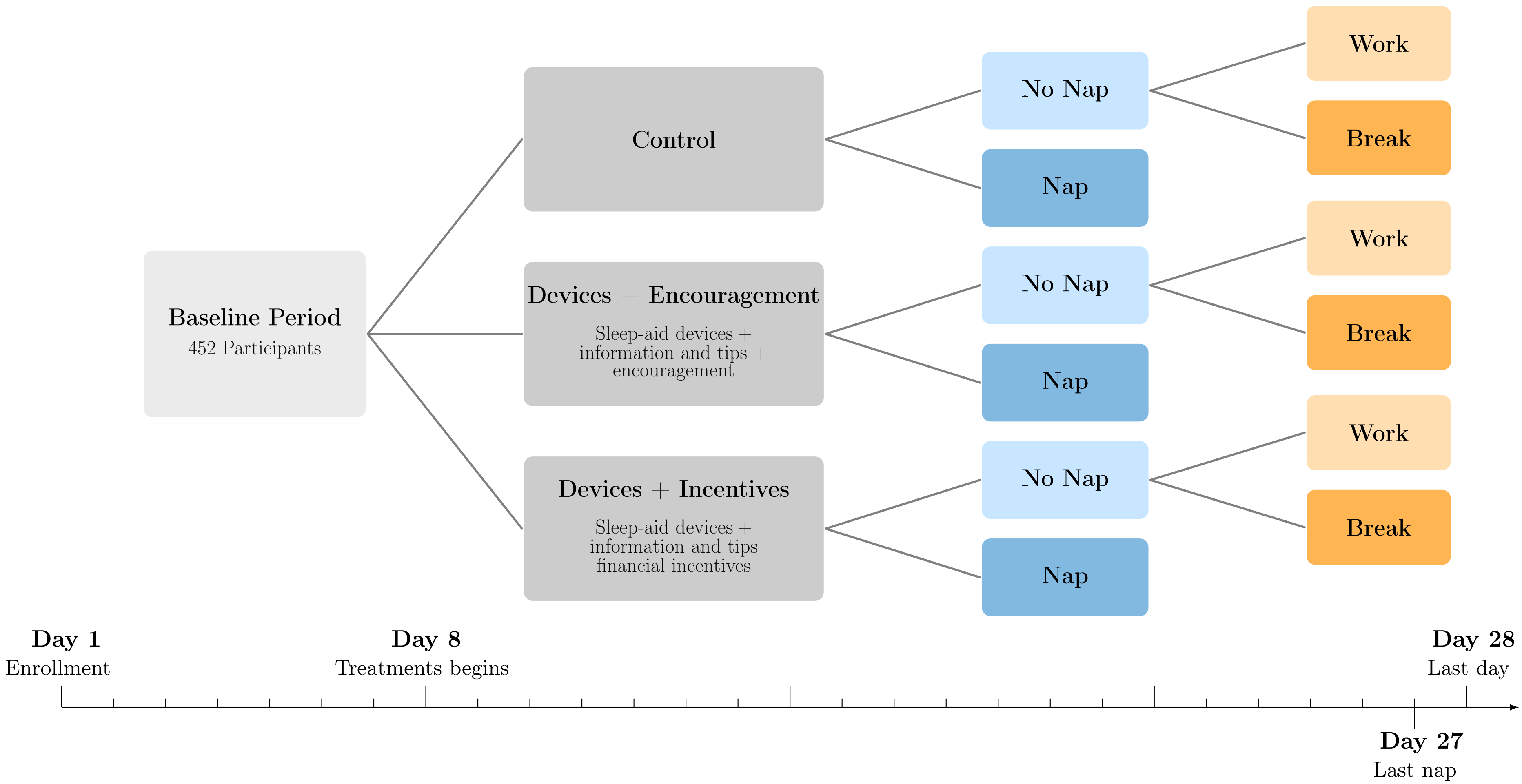
- In panel (a) we select a typical night in our sample, represented by average levels of time in bed, time asleep, and sleep efficiency. During this particular night, the participant stayed in bed for 7 hours and 45 minutes but slept for only 5 hours and 20 minutes, resulting in a sleep efficiency of 69%, corresponding to the 41st, 40th, and 43rd percentile of the control group, respectively. The participant awoke 31 times during this night, and the longest sleep episode he achieved lasted 45 minutes.
- Panel (b) depicts a good night of sleep, with sleep patterns similar to those found in the US and other rich countries: the participant stayed in bed for 7 hours and 53 minutes and slept for 7 hours and 8 minutes, resulting in a sleep efficiency of 90%, corresponding to the 46th, 91st, and 99th percentile of the control group, respectively. In this night, the participant only awoke 9 times, and the longest sleep episode lasted 202 minutes.

FIGURE II: Baseline Distributions of Sleep-Related Variables



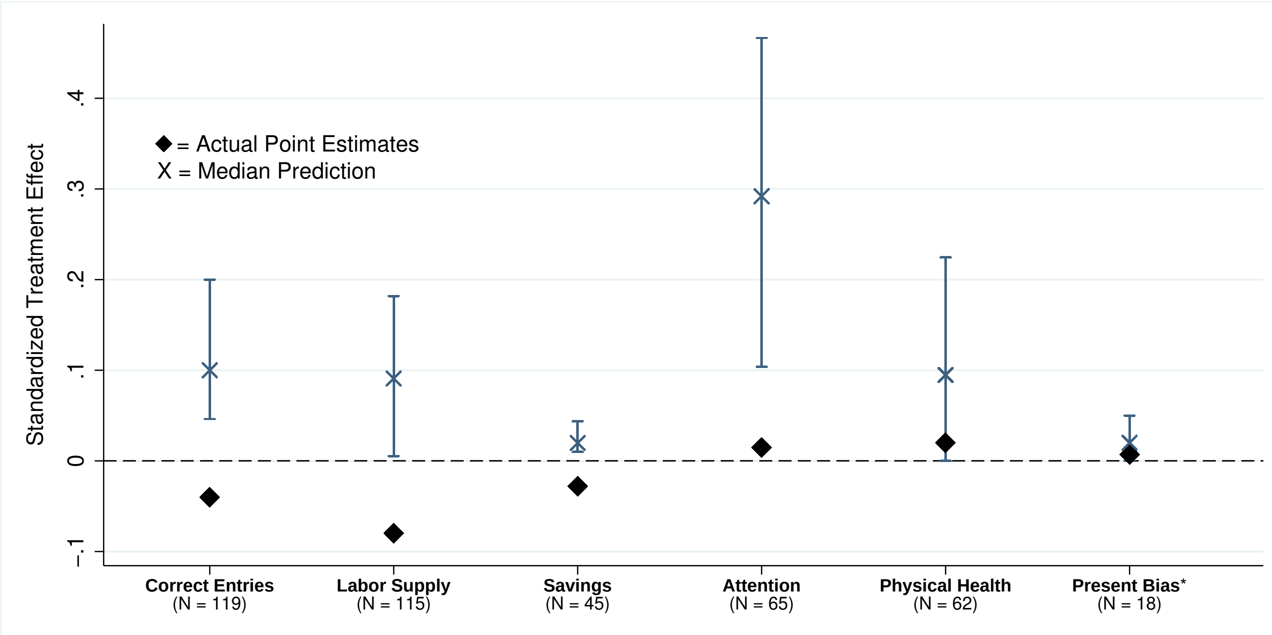
Notes: This figure shows the distribution of the sleep-related variables averaged at the participant-level over the baseline period (7 nights) in the RCT sample (N = 452). Panels (a) and (b) show hours in bed as measured by actigraphy and by self-reports, respectively. Panels (c) and (d) show night sleep duration in hours as measured by actigraphy and by self-reports, respectively. Panels (e) and (f) show sleep efficiency (night sleep duration / time in bed) as measured by actigraphy and by self-reports, respectively.

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 were first divided in 3 groups: Control, Sleep Devices + Encouragement, and Sleep Devices + 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 randomizations 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. Endline surveys occur on day 28 or shortly thereafter.

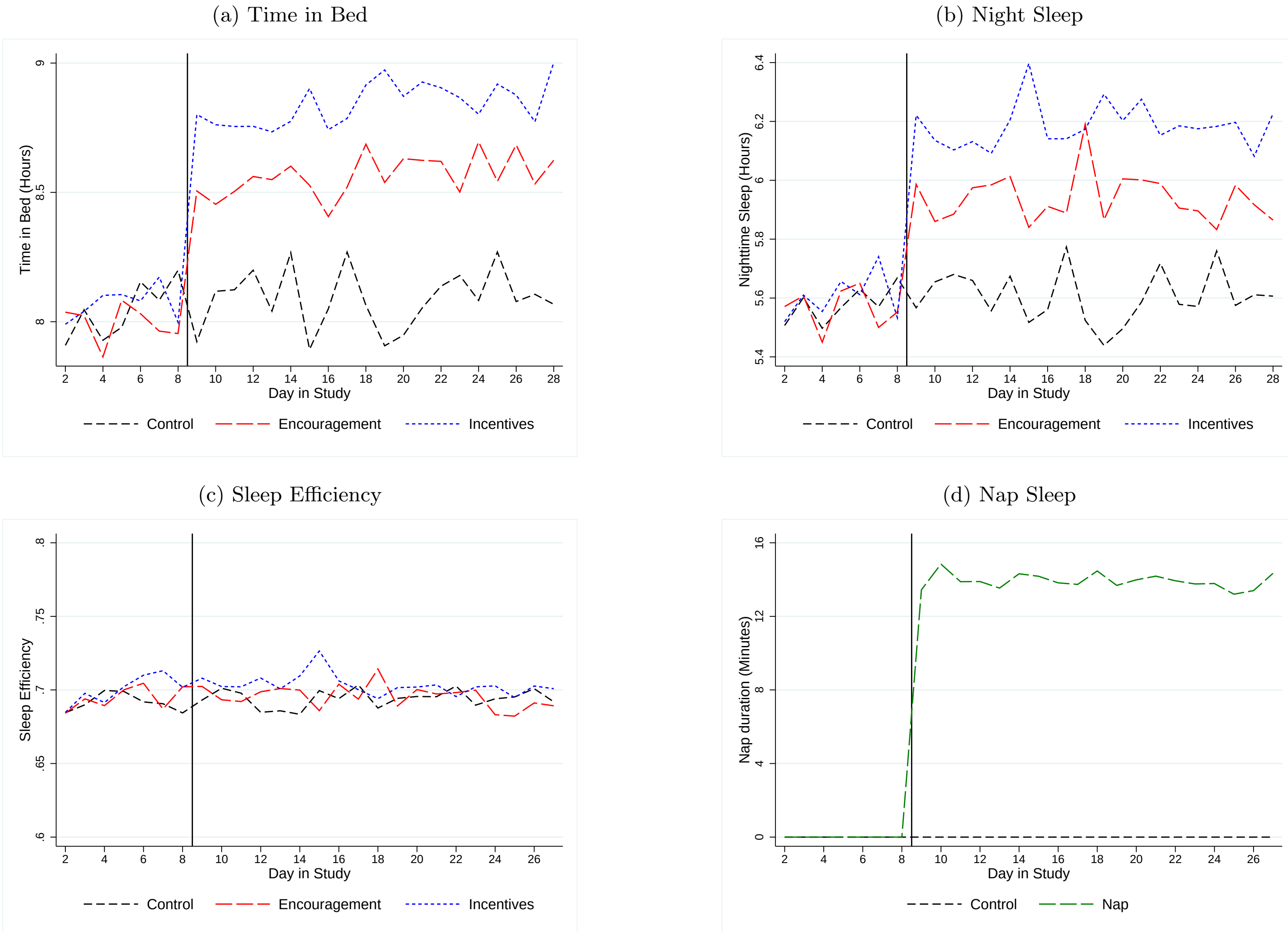
FIGURE IV: Survey of Experts



Notes: This figure summarizes the predictions about the treatment effect of our interventions from economists and sleep experts. Each expert predicted the treatment effects of our intervention. We normalize each prediction, dividing them by by the control group's standard deviation.

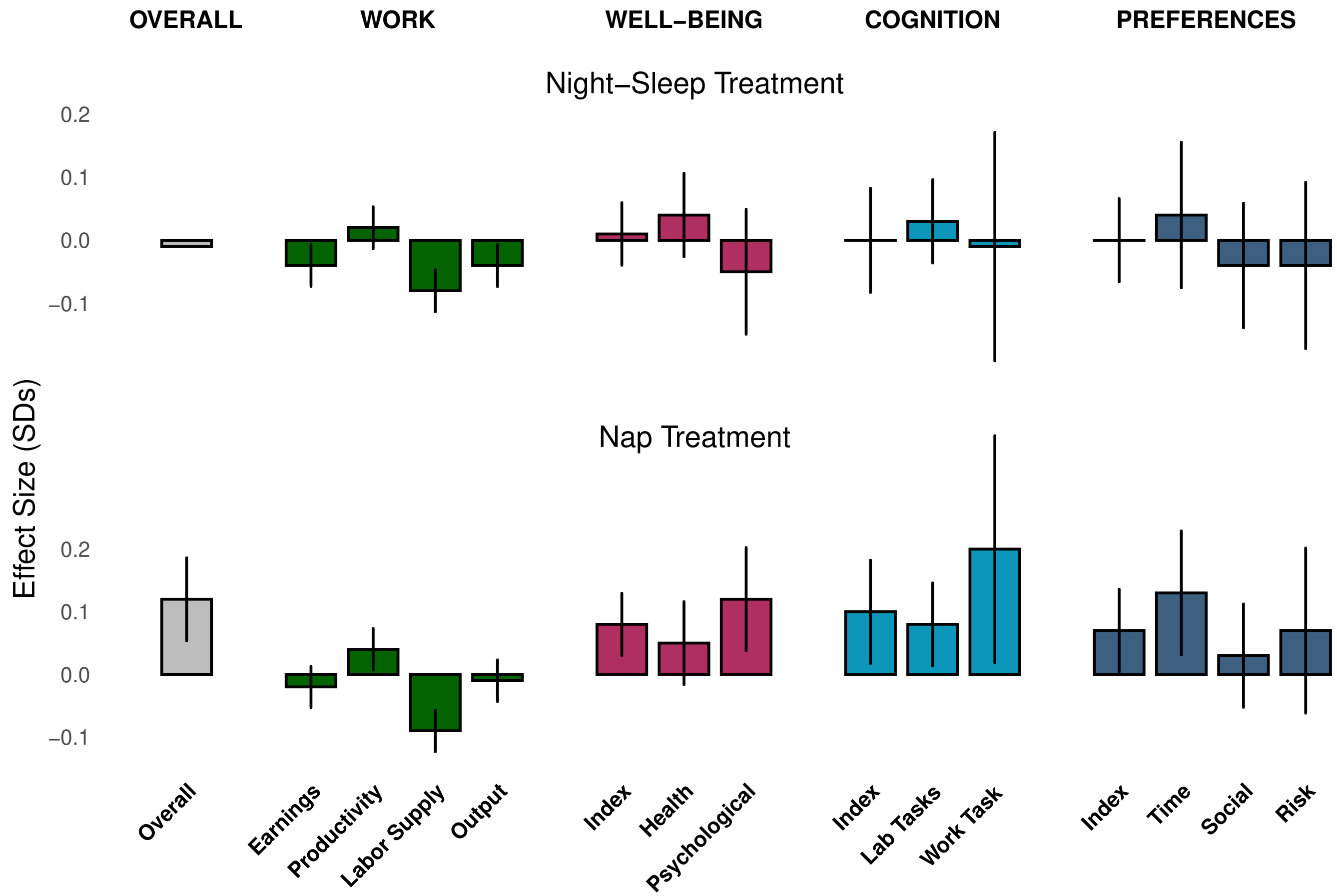
- Each column shows the inter-quartile range (25^{th} to 75^{th} percentiles) of the predictions for a given outcome variable. We also show the median prediction (X) and the actual point estimates (diamond). "N" in the x-axis refers to the number of responses for each outcome variable. This number varies by outcome because we only asked a subset of questions to some types of experts (e.g. sleep researchers).
- Correct Entries refers to the number of daily correct characters in the data-entry task.
- Labor Supply refers to the daily number of hours working in the typing task (excluding voluntary and scheduled pauses).
- Savings refers to the daily amount deposited minus the amount withdrawn in the savings box during the experiment.
- Attention refers to an index pooling inverse response times (IRT) and minor lapses (ML) in the Psychomotor Vigilance Task (PVT).
- Physical Health refers to a variable that pools both systolic and diastolic blood pressure. We flip the sign of the predictions so a positive value means an improve in health (i.e., a reduction in blood pressure).
- Present Bias refers to the β present-bias parameter. Unlike the other variables, the predictions and point-estimate refers to the level of present bias rather than a normalized outcome, for ease of interpretation.

FIGURE V: Impacts on Nighttime and Nap Sleep



Notes: This figure shows the average of different sleep-related variables for different treatment arms by day in study of the RCT. All outcomes are actigraph measures. In panels (a) and (b), we plot hours in bed at night and hours of night sleep, respectively. In panel (c), we plot the series for sleep efficiency (nighttime sleep / time in bed) as measured by the actigraph. In panel (d), we plot the duration of naps in minutes for the nap in the workplace. We only include workday nights and days in the sample. Additionally, we exclude day 28 in Panel (d), since naps were not allowed on that day.

FIGURE VI: Summary of Treatment Effects



Notes: This figure summarizes the treatment effects in our study. We plot the point estimates and 90% confidence intervals for the night-sleep interventions in Panel (a) and the nap intervention in Panel (b). All outcomes are standardized, i.e. we subtract the mean and divide by the standard deviation of the individuals receiving neither the night-sleep nor the nap interventions. The coefficients and confidence intervals are based on the estimates and standard errors in Tables III and IV. The comparison group for the nap treatment is the pooled nap control group, i.e. participants not assigned to the nap intervention. The outcomes variables, described in more detail in Section 3.2, are as follows:

- **Overall index:** aggregates across all the outcomes in the table.
- **Work:** (i) earnings from the data-entry task; (ii) productivity; (iii) active typing time; and (iv) output.
- **Well-being:** (i) composite index of the physical and mental well-being indices; (ii) physical well-being index, a composite of performance in an endline biking task, self-reported illnesses, self-reported pain, self-reported health, and blood pressure; (iii) mental well-being index, a composite of self-reported depression, happiness, life possibility, life satisfaction, and stress.
- **Cognition:** (i) composite index of a lab-based and a work-based measure of cognitive function; (ii) index of lab measures of attentiveness, memory, and inhibitory control; (iii) measure of attention to piece rates in the data-entry task
- **Preferences:** (i) composite index of time, social, and risk preference indices; (ii) index capturing time preferences, including savings and present bias; (iii) index representing social preferences; and (iv) index representing risk preferences.

Table I: Sleep Statistics in Two Samples in Chennai

	RCT Sample (pre-treatment) (1)	**Broader Sample** (2)
Night Sleep		
Hours in bed	8.03 (0.97)	7.68 (1.23)
Hours asleep	5.58 (0.87)	5.46 (1.15)
Sleep efficiency	0.70 (0.08)	0.71 (0.10)
Number of Awakenings	31.95 (7.95)	N/A
Fraction sleeping below 7 hours		
Participant-level	0.95 (0.22)	0.93 (0.26)
Participant-day-level	0.89 (0.31)	0.87 (0.33)
Fraction sleeping below 6 hours		
Participant-level	0.71 (0.45)	0.69 (0.46)
Participant-day-level	0.65 (0.48)	0.64 (0.48)
Self-reported hours asleep	7.20 (0.94)	6.42 (1.49)
Nap sleep		
Percent napping on a given day	N/A	0.25 (0.43)
Duration of naps (conditional on napping)	N/A	0.85 (0.61)
Total sleep		
Hours asleep	5.58 (0.87)	5.69 (1.15)
Participant-nights	3080	1367
Participants	452	439

Notes: This table presents sleep patterns in two samples in Chennai.

- Column 1 presents summary statistics from the RCT sample, only using data from the 7 baseline period nights. Column 2 presents summary statistics from the 3 nights in our complementary Sleep Survey across a broader population in Chennai (described in Appendix F).
- All measures use data from actigraph measurements unless indicated otherwise.
- All means and standard deviations (in parentheses) are on the participant-level unless indicated otherwise.
- The variables shown in the table are: (i) hours in bed (regardless of whether awake or asleep); (ii) hours asleep at night; (iii) sleep efficiency (hours asleep / hours in bed); (iv) number of awakenings per night; (v) proportion of participants below 7 hours of night sleep; (vi) proportion of participants below 6 hours of night sleep; (vii) self-reported hours asleep at night; (viii) proportion of participants napping on any given day; (ix) duration of naps conditional on taking a nap; and (x) total hours asleep per 24 hours (the sum of night sleep and nap sleep).

Table II: Treatment Effects on Sleep

	Night Sleep (1)	**Time in Bed** (2)	**Sleep Efficiency** (3)	**Nap Sleep** (4)	**24-Hr Sleep** (5)
Night-Sleep Treatments	0.44***	0.64***	-0.11	-0.00	0.44***
	(0.05)	(0.06)	(0.42)	(0.00)	(0.05)
Devices + Encouragement	0.33***	0.51***	-0.44	-0.00	0.33***
	(0.06)	(0.06)	(0.47)	(0.01)	(0.06)
Devices + Incentives	0.55***	0.76***	0.22	-0.00	0.55***
	(0.06)	(0.07)	(0.49)	(0.01)	(0.06)
Nap Treatment	-0.08*	-0.17***	0.27	0.24***	0.13**
	(0.05)	(0.05)	(0.40)	(0.00)	(0.05)
Control Mean	5.61	8.07	69.86	0.00	5.61
Control SD	1.20	1.37	11.28	0.00	1.20
N	8454	8454	8454	7191	8035
Participants	451	451	451	450	451

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

- Night sleep, time in bed, nap sleep and 24-hour sleep (col. 1, 2, 4 and 5) are measured in hours. Sleep efficiency (col. 3) is the ratio of night sleep and time in bed (multiplied by 100 for clarity). 24-hour sleep is the sum of nap sleep to night sleep.
- The first row (Night-Sleep Treatments) shows the pooled treatment effect for the two night-sleep treatments. The following two (indented) rows show the relative impacts, from a separate regression, of each of the night-sleep treatments: the Devices Treatment and the Devices + Incentives Treatment. The fourth row shows the treatment effect for the Nap Treatment; these coefficients come from the same regression as the Pooled Night-Sleep Treatment coefficients.
- Each column shows the OLS estimates of equation (1) 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.
- Standard errors are clustered at the participant level.

Table III: Night-Sleep Treatment Effects

	OVERALL	WORK				WELL-BEING		
	Index (1)	**Earnings** (2)	Productivity (3)	Labor Supply (4)	Output (5)	**Index** (6)	Physical (7)	Mental (8)
Night-Sleep Treatments	-0.01	-0.04*	0.02	-0.08***	-0.04*	0.01	0.04	-0.05
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.06)
		[0.29]	[0.58]	[0.00]	[0.12]	[0.99]	[0.61]	[0.62]
Participants	451	451	451	451	451	452	452	452

	COGNITION			PREFERENCES			
	Index (9)	Lab Tasks (10)	Work Task (11)	**Index** (12)	Time (13)	Social (14)	Risk (15)
Night-Sleep Treatments	-0.00	0.03	-0.01	-0.00	0.04	-0.04	-0.04
	(0.05)	(0.04)	(0.11)	(0.04)	(0.07)	(0.06)	(0.08)
	[1.00]	[0.74]	[1.00]	[1.00]	[0.93]	[0.84]	[0.96]
Participants	452	452	429	452	452	415	415

Notes: This table shows treatment effects of the pooled night-sleep interventions on our main outcomes. All dependent variables are normalized with respect to the mean and standard deviation in the control group (i.e., participants receiving neither the night-sleep nor the nap treatments). Described in more detail in Section 3.2, the outcomes are:

- Col. 1: index across all of the family-level summary outcomes (earnings and the three index variables) in the table.
- Col. 2-5 show work-related outcomes from the data-entry task. Earnings is the family-level summary outcome for the Work family. The coefficients come from equation 1, including dummies for assignment to nap treatment and assignment to the work condition (leaving the break group as the omitted category).
- Col. 6-8 show impacts on physical and mental well-being. Col. 6 is an index of these two broad classes of well-being. The physical well-being index (col. 7) is a composite of performance in an endline stationary biking task, blood pressure, and self-reports of illness, pain, and the extent to which health limits daily activity. The mental well-being index (col. 8) includes self-reported depression, happiness, life possibility, life satisfaction, and stress. The coefficients come from equation 2.
- Col. 9-11 show impacts on cognitive function. Col. 9 is an index of two different ways of measuring cognitive function: lab measurements of attentiveness, memory, and inhibitory control (col. 10); and attention to piece rates in the data-entry task (col. 11). The coefficients come from equation 2.
- Col. 12-15 show impacts on preferences. Col. 12 is an index of time (savings and present bias, col. 13), social (col. 14), and risk preferences (col. 15). The coefficients come from equation 2.

All indices are a weighted average of their components, taking into account the covariance structure of the components (Anderson, 2008). All work-related regressions are conducted at the participant-day level (Eq 1). All other regressions are at the participant level (Eq 2). All regressions include a dummy capturing the assignment to the nap treatment and control for gender, age, and the baseline outcome variable. In addition, all work related regressions control for day in study and date fixed effects as well as an indicator for whether the participant was assigned to the work or break group that day. Standard errors in parentheses are robust to heteroscedasticity and clustered at the participant-level when applicable. When required, outcomes are flipped so that a positive value aligns with what would be considered a “better” outcome.

Stars next to coefficients reflect unadjusted P-values (* significant at 10%; ** at 5%; *** at 1%). P-values in square brackets are adjusted for multiple hypothesis testing. There are two levels of corrections:

(1) P-values for the family-level summary outcomes (earnings, well-being index, cognition index, and preferences index) correct for the existence of multiple families.

(2) P-values for component outcomes within each family (e.g., productivity, labor supply, and output for the Work family) correct for the multiple outcomes *within* the family.

Table IV: Nap Treatment Effects

	OVERALL	WORK				WELL-BEING		
	Index	**Earnings**	Productivity	Labor Supply	Output	**Index**	Physical	Mental
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nap vs. Pooled Control	0.12***	-0.02	0.04**	-0.09***	-0.01	0.08***	0.05	0.12**
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.05)
		[0.67]	[0.07]	[0.00]	[0.97]	[0.03]	[0.35]	[0.04]
Nap vs. Break Control	0.15***	0.05**	0.04**	0.01	0.05***			
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)			
		[0.06]	[0.07]	[0.89]	[0.02]			
Nap vs. Work Control	0.08**	-0.10***	0.03**	-0.20***	-0.07***			
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)			
		[0.00]	[0.09]	[0.00]	[0.00]			
Participants	451	451	451	451	451	452	452	452

	COGNITION			PREFERENCES			
	Index	Lab Tasks	Work Task	**Index**	Time	Social	Risk
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Nap vs. Pooled Control	0.10**	0.08**	0.20*	0.07**	0.13**	0.03	0.07
	(0.05)	(0.04)	(0.11)	(0.04)	(0.06)	(0.05)	(0.08)
	[0.09]	[0.08]	[0.16]	[0.17]	[0.12]	[0.87]	[0.69]
Participants	452	452	429	452	452	415	415

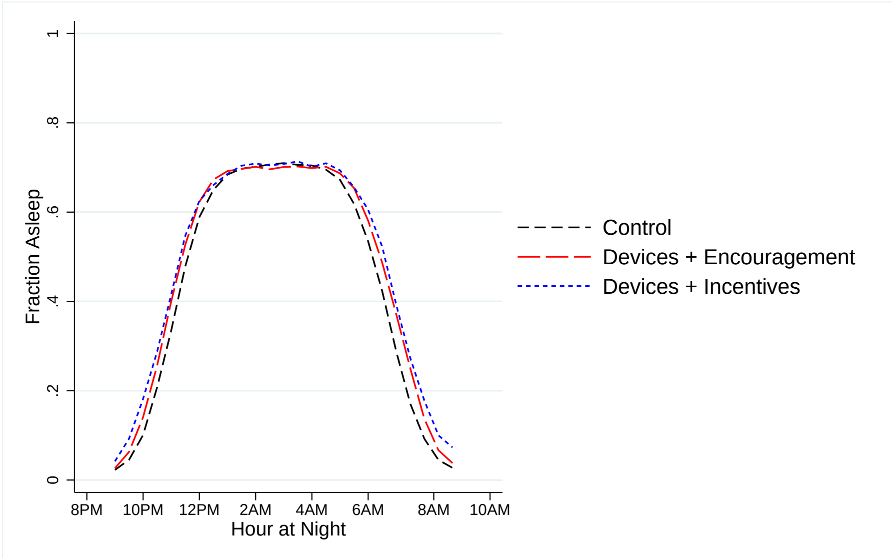
Notes: This table shows the treatment effects of the nap intervention on our main outcomes, closely following the structure of Table III. All dependent variables are normalized with respect to the mean and standard deviation in the control group (i.e., participants receiving neither the night-sleep nor the nap treatments).

- **Each row represents a different regression**. The omitted group varies by regression. Comparisons to the Break Control (Work Control) also include a dummy for the Work Control (Break Control), though it is omitted from the table.
- The outcomes and regressions correspond exactly to those in Table III, described in the table notes of Table III and in more detail in Section 3.2. However, we exclude the last day in study from the analysis because participants do not nap on this day.
- Stars next to coefficients reflect unadjusted P-values (* significant at 10%; ** at 5%; *** at 1%). P-values in brackets are adjusted for multiple hypothesis testing as in Table III. See the table notes in Table III for more details.

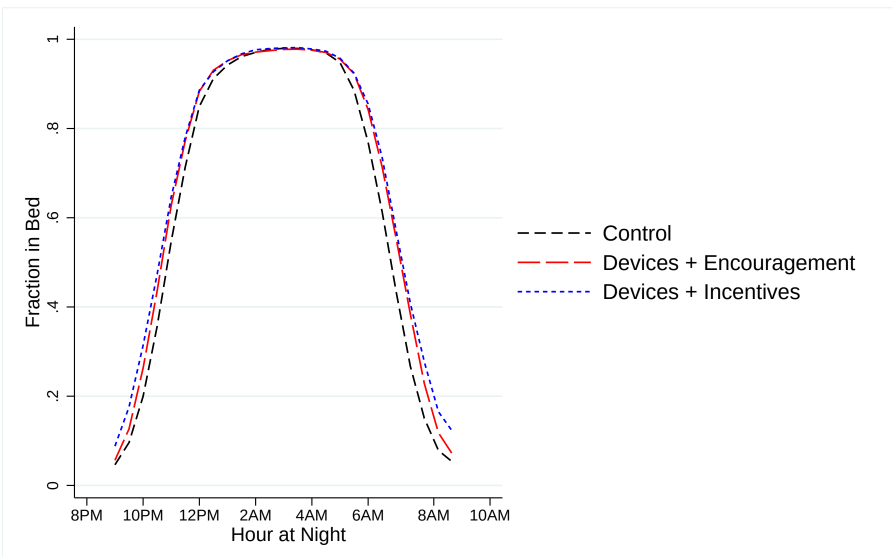
A Online Only Supplementary Tables and Figures

FIGURE A.I: Fraction in Bed and Asleep by Hour of Night

(a) Fraction Asleep by Hour of Night



(b) Fraction in Bed by Hour of Night



Notes: This figure shows the average fraction of participants asleep and in bed over the course of the night during the 19 nights of the treatment period.

- Panel (a) shows the fraction of participants in each night-sleep treatment group that was asleep at any time during the night, as measured by the actigraph.
- Panel (b) shows the fraction of participants in each night-sleep treatment group that was in bed at any given time during the night, as measured by the actigraph.

FIGURE A.II: Measuring and Increasing Sleep

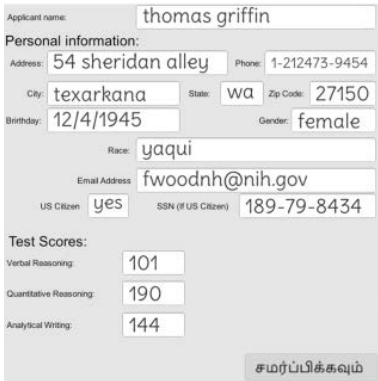
(a) Actigraph watch (b) Devices to Improve Night Sleep Environments (c) Nap Station

Notes: This figure illustrates devices and treatments used to measure and increase study participants' sleep.

- Panel (a) shows an actigraph, the wearable device used to measure study participants' awake/sleep patterns through body motion at all times of the study.
- Panel (b) 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 (not shown in the picture).
- Panel (c) shows the nap station where participants in the nap group were allowed and encouraged to sleep in the early afternoon for up to 30 minutes. The nap stations were located on a separate floor at the study office. The participants in the no nap group were not allowed to use this nap station.

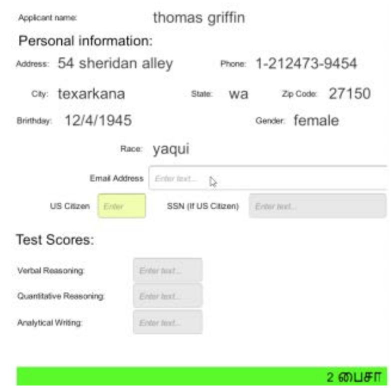
FIGURE A.III: Data-Entry Interface with Salient and Non-Salient Piece Rates

(a) Left side

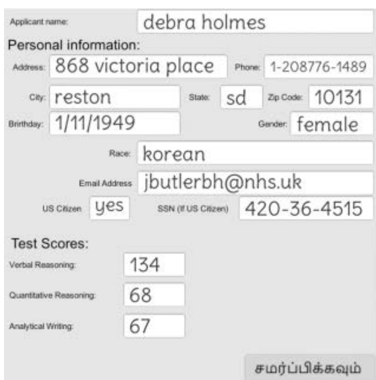


(b) Right side (salient, low)

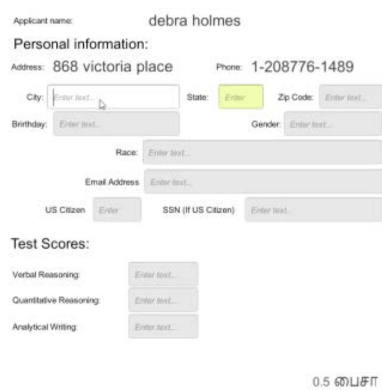
(c) Right side (salient, high)



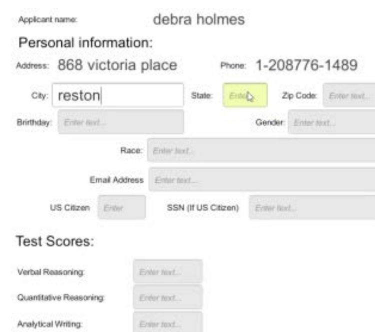
(d) Left side



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

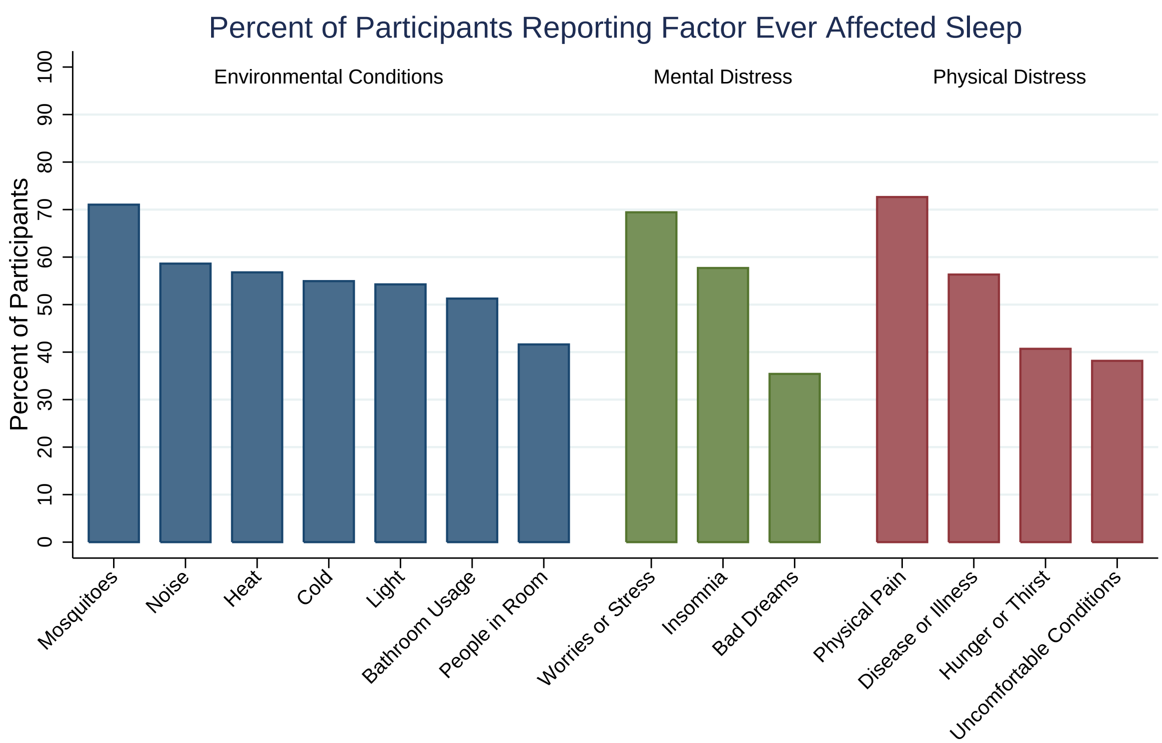


(f) Right side (non-salient, remainder)



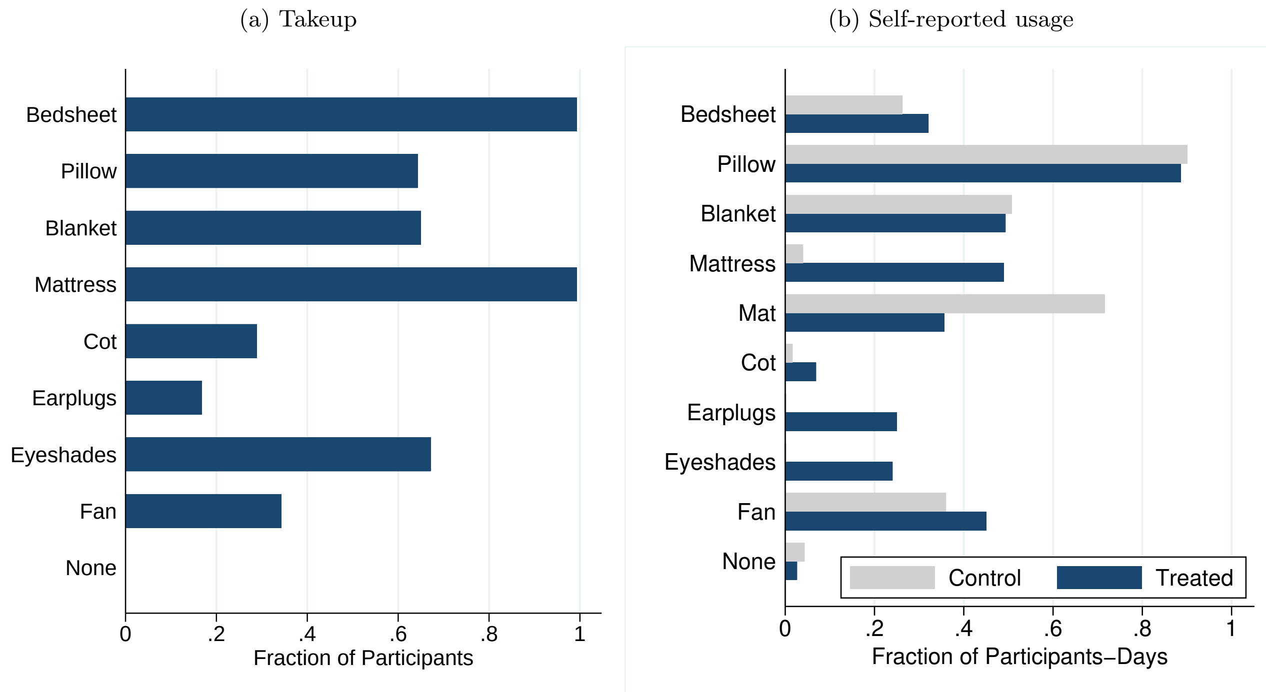
Notes: This figure shows screen shots of the data-entry task interface used by participants. Panels (a) and (d) show the left side of the screen, which contains the (fictional) data to be transcribed by study participants. The remaining panels show versions of the right side of the screen, where participants were supposed to enter the data. Panels (b) and (c) show the right side of the screen under salient low and high incentives, respectively. 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.IV: Factors Interfering with Study Participants' Sleep



Notes: This figure shows the fraction of participants who reported various factors — including environmental conditions, mental distress, and physical distress — impacting their sleep at any point during the study. Participants were asked about sleep disturbances as part of the daily survey, which was administered each morning in the study office. These particular questions were asked at six different points during the study. Participants were asked, "How much does *factor* affect how difficult it is for you to fall asleep?" Responses included "0 - Not at all," "1 - Some," and "2 - A lot."

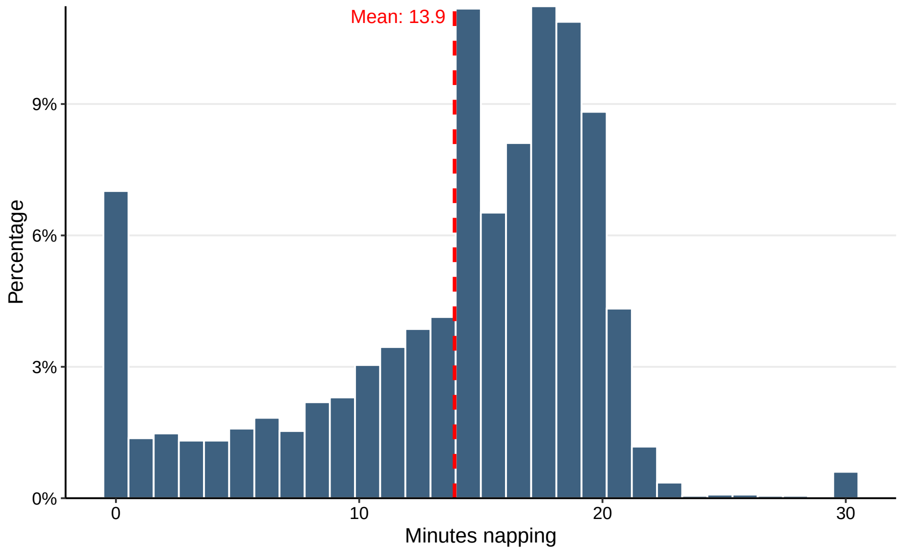
FIGURE A.V: Sleep Aids Usage and Take-Up



Notes: This figure shows the take-up and usage of the sleep devices offered to the participants in the two night-sleep treatments.

- Panel (a) shows the fraction of participants in the night-sleep treatment groups who took home at least one of each offered sleep device.
- Panel (b) shows the share of participants who reports using on any given day each of the devices. These numbers include devices that were not offered by the study, e.g. devices that participants owned prior to joining the study. In blue we show the numbers for the participants in the night-sleep treatment groups. In grey we show the usage of the participants in the corresponding control group.

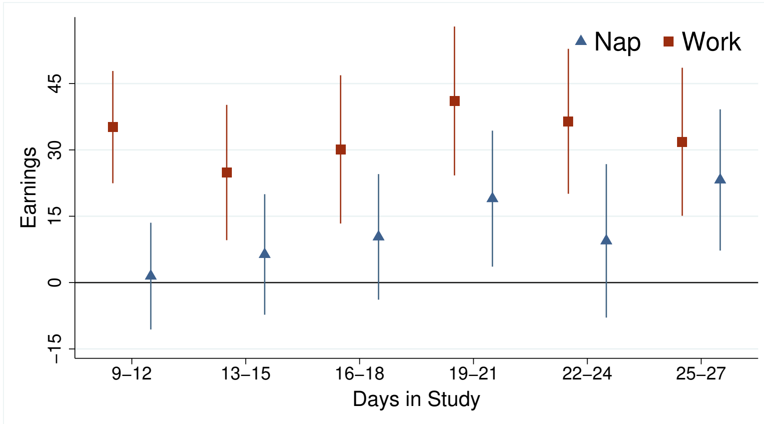
FIGURE A.VI: Distribution of Nap Duration



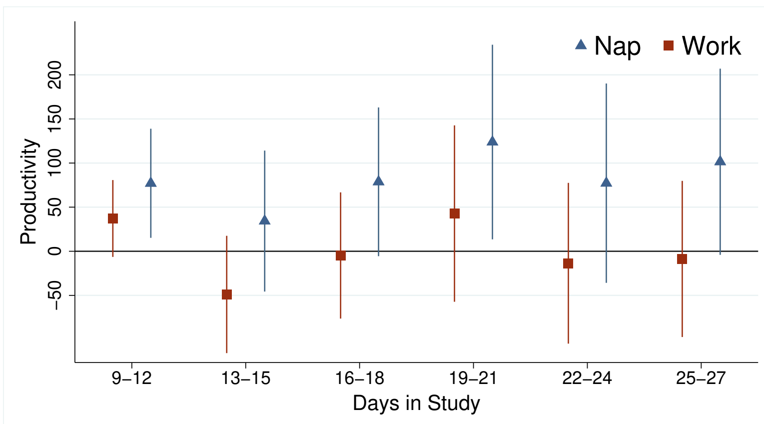
Notes: This figure shows the distribution of nap duration among the nap group during the treatment period. Each observation is the nap duration as measured by actigraph for a participant in a day in the study. We exclude day 28 since naps were not allowed on this day. The red dashed line indicates the average nap duration of 13.9 minutes.

FIGURE A.VII: Treatment Effects of Naps throughout the Study

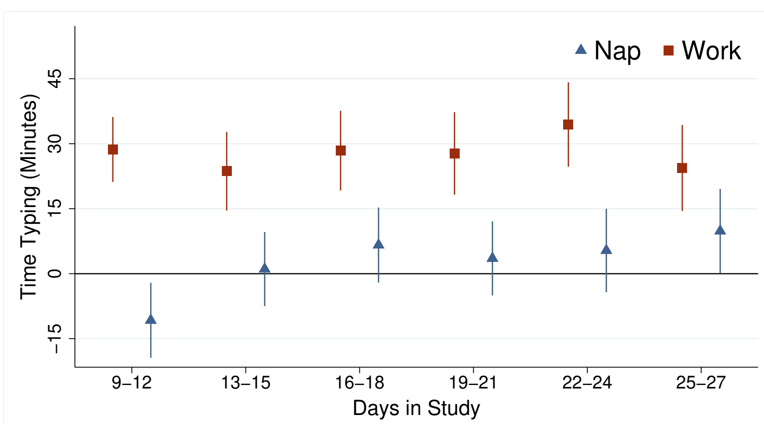
(a) Earnings



(b) Productivity



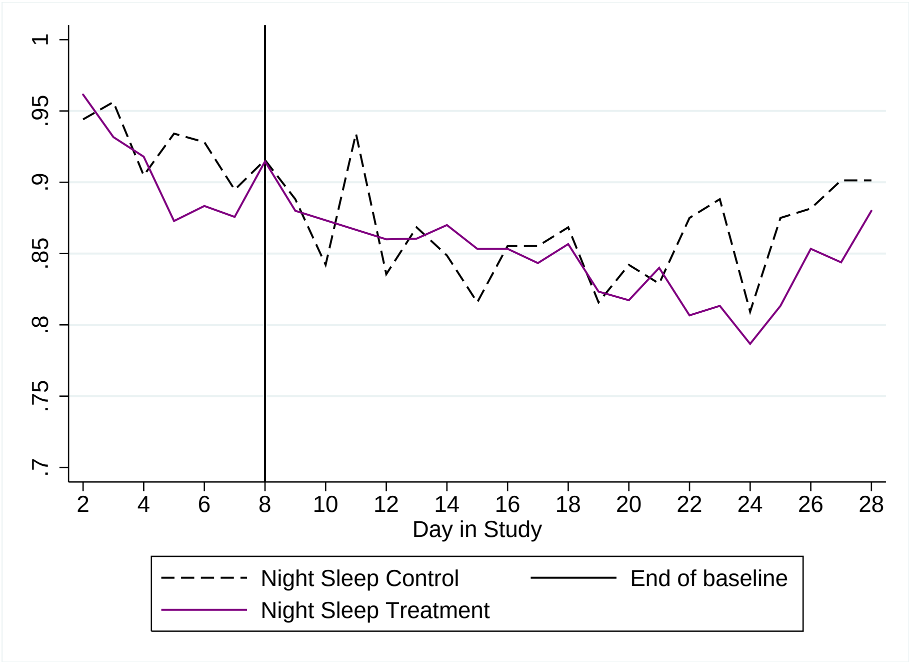
(c) Hours Typing



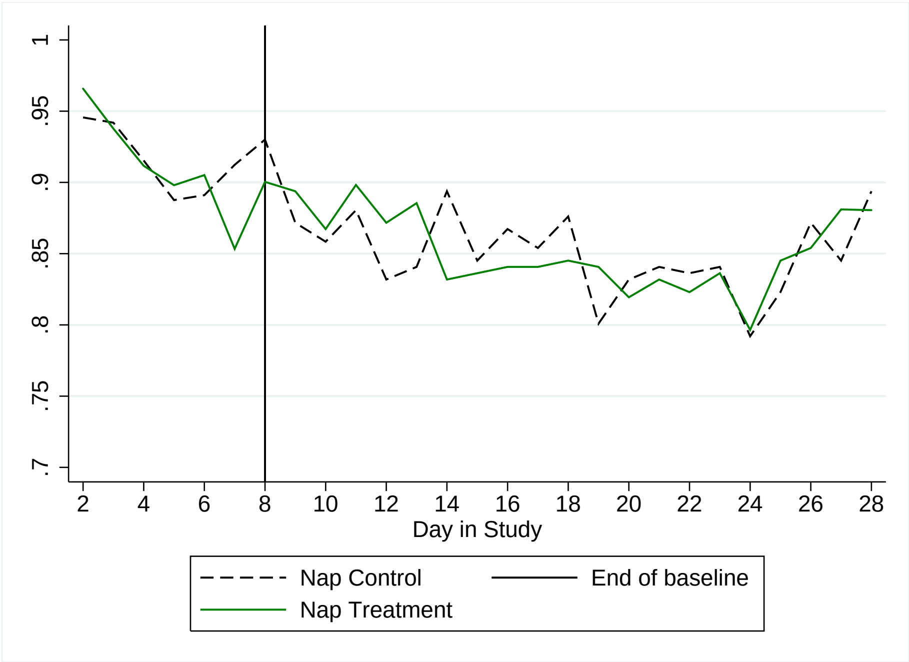
Notes: This figure shows the dynamics of the nap treatment effect throughout the treatment period. The blue series (triangles) shows the nap treatment effect in different in comparison to the Break group (i.e., not assigned to the nap treatment and not allowed to work during the nap break). The red series (squares) shows the difference in productivity between the Work group (i.e., participants not assigned to nap but allowed to work during the nap time) and the Break group. Each graph shows regression coefficients of the outcome variables productivity (panel a), hours typing (panel b), and earnings (panel c) on the indicators of nap and work conditions. The regression follows specification (1), except that we interact the Nap and Work indicators and controls with how long participants have been in the study (3-days dummies). The bars represent 95% confidence intervals.

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

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



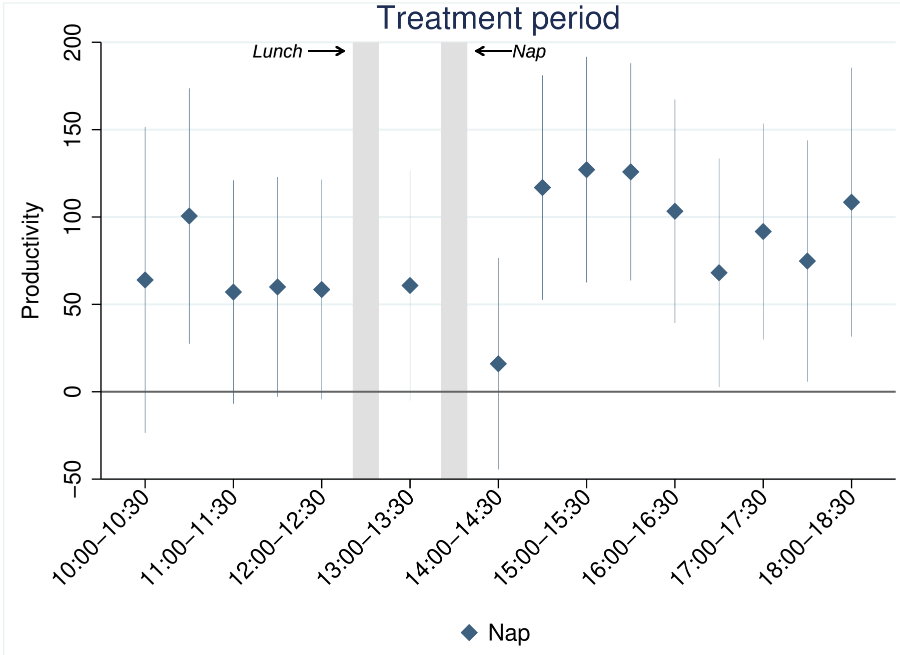
(b) Fraction of Participants Present by Nap Intervention Groups



Notes: This figure shows the fraction of participants present by day of the study and treatment group.

- In panel (a), the solid purple line represents participants in the two night-sleep treatment groups while the dashed black line represents the corresponding control group.
- In panel (b) the solid green line represents the nap group dashed black line represents the corresponding (no nap) control group.

FIGURE A.IX: Nap Treatment Effects on Productivity throughout the Day



Notes: This figure shows the nap treatment effect over the course of the day during the treatment period. Each point shows the coefficients of a regression of productivity on nap interacted with time of the day (30-minute bins). The omitted group is the no nap group, including participants who worked or took a break during the allotted nap time. The control variables in this regression are the same as indicated in equation (1) for the work outcomes, except that we interact them with dummies for each the 30-minute window. The bars represent 90% confidence intervals. The grey rectangles capture the time allotted to lunch and to the nap.

B Appendix Tables

Table A.I: Balance Across Each Experimental Treatment Cell (Part 1/2)

	Averages						***p-val* of differences**				
	Control, No Nap	Encouragement, No Nap	Incentives, No Nap	Control, Nap	Encouragement, Nap	Incentives, Nap	1 = 2	1 = 3	1 = 4	1 = 5	1 = 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A. Demographics											
Female	0.68	0.64	0.64	0.64	0.64	0.74	0.65	0.60	0.65	0.65	0.42
	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)					
Age	35.91	35.04	33.82	35.77	35.52	33.62	0.46	0.08	0.91	0.74	0.05
	(0.86)	(0.80)	(0.74)	(0.89)	(0.85)	(0.83)					
Number of Children	1.40	1.33	1.16	1.44	1.36	1.42	0.69	0.16	0.83	0.80	0.91
	(0.13)	(0.12)	(0.11)	(0.12)	(0.12)	(0.13)					
Years of Education	10.31	10.17	10.53	10.39	9.83	9.88	0.77	0.65	0.87	0.30	0.36
	(0.32)	(0.32)	(0.33)	(0.34)	(0.36)	(0.33)					
Familiar with Computer	0.27	0.33	0.44	0.32	0.24	0.30	0.69	0.23	0.75	0.81	0.85
	(0.10)	(0.11)	(0.10)	(0.10)	(0.10)	(0.11)					
Unemployed	0.96	0.93	0.95	0.95	0.95	0.93	0.46	0.69	0.70	0.70	0.47
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)					
Panel B. Baseline Sleep											
Self-Reported Night Sleep (Hrs)	7.25	7.28	7.20	7.20	7.15	7.08	0.81	0.76	0.74	0.51	0.27
	(0.13)	(0.10)	(0.11)	(0.08)	(0.10)	(0.10)					
Actigraph Night Sleep (Hrs)	5.50	5.59	5.63	5.64	5.54	5.57	0.53	0.35	0.33	0.76	0.60
	(0.10)	(0.10)	(0.09)	(0.10)	(0.11)	(0.09)					
Actigraph Time in Bed (Hrs)	7.99	8.13	8.16	8.23	8.04	8.12	0.37	0.25	0.12	0.73	0.38
	(0.11)	(0.10)	(0.10)	(0.10)	(0.12)	(0.10)					
Sleep Efficiency	0.69	0.70	0.70	0.69	0.69	0.70	0.72	0.72	0.97	0.96	0.92
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)					
Number of Sleep Devices Owned	2.62	2.64	2.34	2.42	2.79	2.33	0.93	0.30	0.45	0.52	0.27
	(0.20)	(0.20)	(0.15)	(0.16)	(0.22)	(0.17)					
Number of Participants	77	75	74	75	75	76					

Notes: This table considers whether there are any underlying differences between the fully disaggregated randomized experimental arms at baseline.

- Columns 1-6 show baseline means and standard errors (in parentheses) for each treatment arm.
- Columns 7-11 show p-values of t-tests between column 1 and each of the other columns, respectively.

Table A.I: Balance Across Each Experimental Treatment Cell (Part 2/2)

	Averages						p-val of differences				
	Control, No Nap (1)	Encouragement, No Nap (2)	Incentives, No Nap (3)	Control, Nap (4)	Encouragement, Nap (5)	Incentives, Nap (6)	1 = 2 (7)	1 = 3 (8)	1 = 4 (9)	1 = 5 (10)	1 = 6 (11)
Panel C. Health, Well-Being, Cognition											
Baseline Health	-0.04	0.01	-0.04	0.06	-0.04	0.04	0.46	0.98	0.16	0.98	0.25
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)					
Baseline Wellbeing	-0.03	0.06	0.02	0.00	0.02	0.01	0.33	0.60	0.75	0.63	0.64
	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)	(0.06)					
PVT Payment	18.60	18.59	19.24	18.58	18.89	18.13	0.99	0.26	0.97	0.60	0.41
	(0.36)	(0.41)	(0.36)	(0.44)	(0.43)	(0.40)					
HF Payment	20.49	20.89	20.48	20.19	20.50	19.86	0.53	0.99	0.65	0.99	0.33
	(0.48)	(0.46)	(0.43)	(0.50)	(0.47)	(0.42)					
Corsi Payment	20.92	21.29	20.48	21.89	21.02	21.25	0.64	0.59	0.22	0.90	0.68
	(0.57)	(0.59)	(0.57)	(0.57)	(0.56)	(0.54)					
Panel D. Baseline Work and Savings											
Typing Time (Hrs)	4.52	4.49	4.43	4.46	4.61	4.37	0.87	0.56	0.72	0.55	0.32
	(0.08)	(0.07)	(0.07)	(0.07)	(0.20)	(0.07)					
Time in Office (Hrs)	8.01	7.90	7.87	7.90	7.94	7.85	0.33	0.23	0.35	0.58	0.18
	(0.08)	(0.07)	(0.09)	(0.08)	(0.09)	(0.07)					
Productivity	2475.56	2625.09	2577.37	2268.74	2277.16	2361.89	0.57	0.70	0.43	0.45	0.66
	(200.62)	(216.10)	(198.58)	(157.03)	(180.61)	(147.09)					
Earnings	243.13	247.78	239.99	225.90	223.60	225.04	0.78	0.85	0.31	0.25	0.28
	(13.66)	(14.19)	(12.34)	(10.08)	(10.81)	(9.69)					
Attendance	0.94	0.94	0.93	0.95	0.92	0.92	0.93	0.61	0.56	0.15	0.09
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)					
Savings (Rs.)	84.82	102.99	108.77	108.89	78.53	116.62	0.35	0.22	0.22	0.75	0.10
	(11.51)	(15.41)	(12.97)	(14.02)	(12.31)	(16.09)					
Prior Savings (Rs. 1000)	24.73	13.07	45.64	29.67	19.09	20.04	0.43	0.16	0.74	0.70	0.75
	(8.11)	(4.86)	(21.09)	(7.74)	(6.77)	(5.06)					
Joint Orthogonality Test							0.96	0.57	0.60	0.73	0.37
Number of Participants	77	75	74	75	75	76					

Notes: This table considers whether there are any underlying differences between the fully disaggregated randomized experimental arms at baseline.

- Columns 1-6 show baseline means and standard errors (in parentheses) for each treatment arm.
- Columns 7-11 show p-values of t-tests between column 1 and each of the other columns. respectively.
- 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 the treatment arms being compared across all variables in the table.

Table A.II: Balance Checks Corresponding to Main Regression Specifications (Part 1/2)

	Night-Sleep Treatments			**Nap Treatments**		
	Control	Treatment	1 = 2	Control	Treatment	4 = 5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Demographics						
Female	0.66	0.66	0.91	0.65	0.64	0.62
	(0.04)	(0.03)		(0.03)	(0.04)	
Age	35.84	34.50	0.06	34.94	35.28	0.97
	(0.62)	(0.40)		(0.46)	(0.58)	
Number of Children	1.42	1.32	0.34	1.30	1.35	0.29
	(0.09)	(0.06)		(0.07)	(0.08)	
Years of Education	10.35	10.10	0.39	10.34	10.00	0.26
	(0.23)	(0.17)		(0.19)	(0.24)	
Familiar with Computer	0.30	0.33	0.67	0.35	0.28	0.41
	(0.07)	(0.05)		(0.06)	(0.07)	
Unemployed	0.95	0.94	0.54	0.95	0.94	0.84
	(0.02)	(0.01)		(0.01)	(0.02)	
Panel B. Baseline Sleep						
Self-Reported Night Sleep (Hrs)	7.22	7.18	0.63	7.24	7.21	0.24
	(0.08)	(0.05)		(0.07)	(0.07)	
Actigraph Night Sleep (Hrs)	5.57	5.58	0.85	5.57	5.57	0.89
	(0.07)	(0.05)		(0.06)	(0.07)	
Actigraph Time in Bed (Hrs)	8.11	8.11	0.94	8.09	8.08	0.66
	(0.07)	(0.05)		(0.06)	(0.08)	
Sleep Efficiency	0.69	0.70	0.75	0.70	0.70	0.77
	(0.01)	(0.00)		(0.01)	(0.01)	
Number of Sleep Devices Owned	2.52	2.53	0.97	2.54	2.71	0.87
	(0.13)	(0.09)		(0.11)	(0.15)	
Number of Participants	152	300		226	226	

Notes: This table considers whether there are any underlying differences between the randomized experimental arms at baseline.

- Columns 1-2 show baseline means and standard errors (in parentheses) by night-sleep treatments status. Column 3 show the p-value of a t-test between columns 1 and 2.
- Columns 4 and 5 show baseline means and standard errors (in parentheses) by nap treatment status. Column 6 shows the p-value for the t-test between the no nap group (control) and nap group.

Table A.II: Balance Checks Corresponding to Main Regression Specifications (Part 2/2)

	Night-Sleep Treatments			**Nap Treatments**		
	Control	Treatment	1 = 2	Control	Treatment	4 = 5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel C. Health, Well-Being, Cognition						
Health Index	0.01	-0.00	0.74	-0.02	-0.01	0.29
	(0.04)	(0.03)		(0.03)	(0.04)	
Well-Being Index	-0.01	0.03	0.46	0.02	0.04	0.90
	(0.04)	(0.03)		(0.04)	(0.05)	
Low Incentive PVT Pay (Rs.)	12.71	12.58	0.48	12.74	12.54	0.21
	(0.14)	(0.11)		(0.11)	(0.16)	
Low Incentive HF Pay (Rs.)	13.56	13.67	0.43	13.65	13.72	0.83
	(0.12)	(0.08)		(0.09)	(0.11)	
Low Incentive Corsi Pay (Rs.)	13.95	13.83	0.54	13.80	13.87	0.47
	(0.15)	(0.12)		(0.13)	(0.16)	
Panel D. Baseline Work and Savings						
Typing Time (Hrs)	4.49	4.47	0.86	4.48	4.55	0.99
	(0.05)	(0.06)		(0.04)	(0.11)	
Time in Office (Hrs)	7.95	7.89	0.36	7.92	7.92	0.70
	(0.06)	(0.04)		(0.05)	(0.06)	
Productivity	2373.51	2459.66	0.59	2558.52	2451.13	0.09
	(127.65)	(93.42)		(118.05)	(141.07)	
Earnings	234.63	234.05	0.96	243.64	235.69	0.05
	(8.52)	(5.93)		(7.72)	(8.94)	
Attendance (baseline)	0.95	0.94	0.06	0.94	0.94	0.28
	(0.01)	(0.00)		(0.00)	(0.01)	
Savings (Rs.)	96.70	101.76	0.67	98.69	90.76	0.81
	(9.07)	(7.17)		(7.71)	(9.88)	
Prior Savings (Rs. 1000)	27.17	24.38	0.76	27.71	16.08	0.57
	(5.59)	(5.76)		(7.63)	(4.16)	
Joint Orthogonality Test			0.57			0.79
Number of Participants	152	300		226	226	

Notes: This table considers whether there are any underlying differences between the experimental arms at baseline.

- Columns 1-2 show baseline means and standard errors (in parentheses) by night-sleep treatments status. Column 3 show the p-value of a t-test between columns 1 and 2.
- Columns 4 and 5 show baseline means and standard errors (in parentheses) by nap treatment status. Column 6 shows the p-value for the t-test between the no nap group (control) 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 the treatment arms being compared across all variables in the table.

Table A.III: Timing of Tasks in the Study

	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
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 presents information on the timing of the experimental tasks. Further information about the tasks can be found in Section 3 and Appendix Section C. In the Present Bias Task, we show the dates the task was performed by the end of our study. In the first months of the study, the participants completed 4 rounds of the present bias task (instead of 2). More details are provided in Section C.6.3.

Table A.IV: Survey of Experts: Summary Statistics

	All					General Economists					Behavioral Economists					Sleep Experts				
	Mean (1)	p25 (2)	Median (3)	p75 (4)	N (5)	Mean (6)	p25 (7)	Median (8)	p75 (9)	N (10)	Mean (11)	p25 (12)	Median (13)	p75 (14)	N (15)	Mean (16)	p25 (17)	Median (18)	p75 (19)	N (20)
Correct Entries	0.16	0.05	0.10	0.20	119	0.07	0.04	0.06	0.08	28	0.04	0.02	0.03	0.07	19	0.23	0.10	0.14	0.28	72
Hours Working	0.13	0.01	0.09	0.18	115	0.05	0.00	0.05	0.09	27	0.04	0.00	0.04	0.08	19	0.19	0.02	0.18	0.24	69
Savings	0.08	0.01	0.02	0.04	45	0.12	0.00	0.02	0.04	26	0.03	0.01	0.03	0.05	19	.	.	.	.	.
Present Bias	0.02	0.00	0.02	0.05	18	.	.	.	.	.	0.02	0.00	0.02	0.05	18	.	.	.	.	.
Attention	0.33	0.10	0.29	0.47	65	.	.	.	.	.	.	.	.	.	.	0.33	0.10	0.29	0.47	65
Physical Health	0.12	0.00	0.09	0.22	62	.	.	.	.	.	.	.	.	.	.	0.12	0.00	0.09	0.22	62

Notes: This table describes survey responses predicting the effects of the Night-Sleep Treatments from experts in economics and sleep science.

- Each row presents a different outcome about which the experts were asked to make a prediction.
- Experts were only asked to respond to topics within their likely expertise. Items that were not asked of that group are left blank.
- Respondents were provided with information about the increase in sleep among the treated participants and the control group mean for each outcome. Section C.1 provides additional details on the survey.
- The values in the table are standardized to reflect the intention-to-treat (ITT) parameter predictions for each outcome divided by the control group's standard deviation. The sleep science experts predicted the ITT directly. The economists predicted the impact of a one hour increase in sleep duration estimated by an IV approach. Given the differences in response format, we multiply the economists' predictions by the first stage to recover the ITT prediction.
- Correct Entries (row 1) refers to the number of correct characters in the data-entry task each day.
- Hours Working (row 2) refer to the number of hours working in the typing task (excluding voluntary and scheduled pauses) each day.
- Savings (row 3) refers to the amount of money (in Rupees) deposited minus the amount withdrawn by the participants in the office's savings box during the experiment each day.
- Present Bias (row 4) refers to the present-bias parameter β. Unlike the other variables, the predictions and point-estimate refers to the level of present bias rather than a normalized outcome, for ease of interpretation.
- Attention (row 5) refers to an index pooling inverse response times (IRT) and minor lapses (ML) in the Psychomotor Vigilance Task (PVT).
- Physical Health (row 6) refers to a variable that pools predictions of both systolic and diastolic blood pressure.

Table A.V: Heterogeneous Treatment Effects for Sleep Outcomes

	Night Sleep Duration in Hours					**Night Sleep Efficiency in %**					**Nap Sleep Duration in Minutes**				
		X=Sleep Duration	X=Sleep Efficiency	X=Baseline Naps	X=Female		X=Sleep Duration	X=Sleep Efficiency	X=Baseline Naps	X=Female		X=Sleep Duration	X=Sleep Efficiency	X=Baseline Naps	X=Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Night-Sleep Treatments	0.48***	0.46***	0.46***	0.38***	0.37***	-0.14	-0.18	-0.17	-1.29	-0.59					
	(0.05)	(0.08)	(0.08)	(0.12)	(0.11)	(0.42)	(0.65)	(0.65)	(0.80)	(0.79)					
X		-0.04	0.15*	-0.11	-0.11		0.15	0.13	-0.62	0.71					
		(0.11)	(0.09)	(0.11)	(0.10)		(0.73)	(0.81)	(0.78)	(0.82)					
Night-Sleep Treatments*X		0.03	0.04	0.12	0.17		0.09	0.06	1.56	0.68					
		(0.10)	(0.10)	(0.14)	(0.12)		(0.82)	(0.86)	(0.97)	(0.95)					
Nap Treatment											13.84***	13.52***	13.41***	13.28***	13.51***
											(0.30)	(0.42)	(0.42)	(0.62)	(0.52)
Nap Treatment*X												0.66	0.84	0.76	0.49
												(0.55)	(0.57)	(0.69)	(0.61)
DV Control Group Mean	5.60	5.60	5.60	5.60	5.60	69.83	69.83	69.83	69.83	69.83					
Participants	451	451	451	451	451	451	451	451	451	451	452	452	452	452	452

Notes: This table considers heterogeneous treatment effects of the night-sleep interventions on night sleep duration and efficiency, and the nap intervention on nap duration.

- The outcome variables are:
  - Columns 1-5: Night sleep duration (in hours) as measured by the actigraph.
  - Columns 6-10: Night sleep efficiency (sleep duration/time in bed) as measured by the actigraph.
  - Columns 11-15: Nap duration (in minutes) as measured by the actigraph.
- Columns 1, 6, 11 provide the treatment effect for the whole sample, without interactions, as a point of reference.
- The remaining columns interact the treatments with a dummy that indicates whether the participant is above median for a variable, X, measured during the baseline period. The X variables are:
  - Columns 2, 7, 12: Sleep duration (in hours) as measured by the actigraph during baseline.
  - Columns 3, 8, 13: Sleep efficiency (sleep duration/time in bed) as measured by the actigraph during baseline.
  - Columns 4, 9, 14: Whether the participant reported napping at least once a day prior to the beginning of the study.
  - Columns 5, 10, 15: Whether the participant is female.
- Regressions control for the baseline outcome, gender, and age. The nap treatment effects should be interpreted as the difference between treatment and the pooled nap control group (extra work and break jointly).
- Regressions are at the participant-day level and standard errors (in parentheses) are clustered at the participant level.

Table A.VI: Impacts on Time Allocation

	Labor Supply				Sleep	
	Time at Office	Arrival Time	Leave Time	Work Breaks	Get Up Time	Bed Time
	(1)	(2)	(3)	(4)	(5)	(6)
Night-Sleep Treatments vs. Control	-0.15***	0.10***	-0.05	0.03**	0.42***	-0.29***
	(0.04)	(0.03)	(0.03)	(0.01)	(0.04)	(0.05)
Nap vs. Break Control	0.06	-0.01	0.05	0.02*	-0.04	0.12**
	(0.04)	(0.03)	(0.03)	(0.01)	(0.05)	(0.05)
Nap vs. Work Control	0.06	0.00	0.06**	-0.04***	-0.06	0.16***
	(0.04)	(0.03)	(0.03)	(0.02)	(0.04)	(0.05)
Control Mean	6.71	10.53	18.33	0.26	7.13	23.04
Control SD	2.93	0.72	0.95	0.52	1.15	1.17
Participants	451	451	451	451	450	450

Notes: This table shows the impact of our treatments on different measures of participants' daily time allocation.

- Each row shows an OLS coefficient for a different regression. Row 1 shows the night-sleep treatment effect and rows 2 and 3 show the nap treatment effect in comparison to participants assigned to not work during the nap period (row 2) and in comparison to those assigned to work during the nap period (row 3).
- The dependent variables in each column are: (1) time spent in the office in hours (even if not working); (2) time of arrival to the office; (3) time of departure from the office; (4) time spent in voluntary breaks from work in hours; (5) time the participants gets out of bed in the morning; (6) time participant goes to bed at night.
- Time of day measures are represented using a 24-hour clock starting at midnight.
- Regressions are run at the participant-day level and standard errors are clustered at the participant level. We control for participants' age, gender, and baseline outcome variable, whether they were randomized to the break or extra work group, the day type, and the fraction of high piece rates received in the typing task.

Table A.VII: Treatment Effects: Fully Disaggregated Treatment Arms

	OVERALL	WORK				WELL-BEING		
	Index	**Earnings**	Productivity	Labor Supply	Output	**Index**	Physical	Mental
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Devices + Encouragement	0.07	-0.07**	-0.02	-0.06	-0.07**	0.15***	0.16**	0.13**
	(0.07)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.06)	(0.06)
Devices + Incentives	-0.01	-0.08**	-0.02	-0.08**	-0.08**	0.07*	0.08	0.08
	(0.07)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.06)	(0.07)
Devices + Encouragement × Nap	-0.05	0.07	0.07*	-0.02	0.08*	-0.18***	-0.21**	-0.17*
	(0.10)	(0.05)	(0.04)	(0.05)	(0.04)	(0.06)	(0.09)	(0.09)
Devices + Incentives × Nap	0.01	0.07*	-0.02	0.07	0.07	-0.10	-0.12	-0.09
	(0.10)	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)	(0.09)	(0.10)
Nap Treatment	0.15**	0.01	-0.01	0.03	0.01	0.17***	0.16**	0.19***
	(0.07)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.07)	(0.07)
Participants	451	451	451	451	451	452	452	452

	COGNITION			PREFERENCES			
	Index	Lab Tasks	Work Task	**Index**	Time	Social	Risk
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Devices + Encouragement	-0.00	-0.00	0.04	-0.09	-0.05	-0.10	-0.15
	(0.07)	(0.07)	(0.17)	(0.06)	(0.12)	(0.09)	(0.13)
Devices + Incentives	-0.03	0.04	-0.12	-0.04	0.11	-0.16*	-0.06
	(0.08)	(0.07)	(0.19)	(0.07)	(0.12)	(0.08)	(0.13)
Devices + Encouragement × Nap	-0.04	0.06	-0.16	0.19**	0.16	0.21	0.19
	(0.11)	(0.10)	(0.25)	(0.09)	(0.16)	(0.14)	(0.18)
Devices + Incentives × Nap	0.08	-0.02	0.30	0.07	-0.12	0.15	0.07
	(0.11)	(0.09)	(0.25)	(0.09)	(0.16)	(0.13)	(0.19)
Nap Treatment	0.09	0.07	0.15	-0.01	0.12	-0.09	-0.01
	(0.07)	(0.07)	(0.17)	(0.07)	(0.12)	(0.09)	(0.14)
Participants	452	452	429	452	452	415	415

Notes: This table considers the treatment effects shown in Tables III and IV for all five fully disaggregated treatments arms. All dependent variables are normalised with respect to the control group's mean and standard deviation. The outcomes are identical to to the outcomes in Tables III and IV:

- Col. 1: index across all of the family-level outcomes (earnings and the three index variables) in the table.
- Col. 2-5: outcomes related to the data-entry work. The outcomes are: (2) productivity; (3) active typing time; (4) output; and (5) earnings from the data-entry task. Earnings is the family-level outcome for the Work family.
- Col. 6-8: outcomes related to health and well-being. Column 6 is an index of our two broad measures of well-being. The physical well-being index (col. 7) is a composite of performance in an endline stationary biking task, blood pressure, and self-reports of illness, pain, and the extent to which health has limited daily activity. The mental well-being index (col. 8) is a composite of self-reported depression, happiness, life possibility, life satisfaction, and stress. The coefficients come from equation 2.
- Col. 9-11: outcomes related to cognitive function. Col. 9 is an index of two different ways of measure cognitive function: lab measurements of attentiveness, memory, and inhibitory control (column 10); and attention to piece rates in the data-entry task (column 11). The coefficients come from equation 2.
- Col. 12-15: outcomes related to preferences. Col. 12 is an index of three different categories of preferences: time – including savings and present bias – (col. 13), social (col. 14), and risk preferences (col. 15). The coefficients come from equation 2.

All indices are a weighted average of its components, in which the weights take into account the covariance structure of the components (Anderson, 2008). All work-related regressions are conducted at the participant-day level. All other regressions are at the participant level. Standard errors in parentheses are robust to heteroscedasticity and clustered at the participant-level when applicable. When required, outcomes are flipped so that a positive value aligns with what would be considered a "better" outcome. All regressions control for gender and age of the participant. Regressions for work outcomes (columns 2-5) also control for day in study, date, fraction of high piece rates for the day, whether the day was long, and whether the participant was assigned to the work or break group that day.

Table A.VIII: Treatment Effects on Mental Well-being

	Standardized Components				
	Depression (1)	Happiness (2)	Life Possibility (3)	Life Satisfaction (4)	Stress (5)
Night-Sleep Treatments	-0.11	0.03	-0.00	0.01	-0.06
	(0.10)	(0.04)	(0.06)	(0.06)	(0.06)
Nap Treatment	0.05	0.13***	0.20***	0.11*	0.01
	(0.09)	(0.04)	(0.06)	(0.06)	(0.06)
Participants	445	452	445	445	445

Notes: This table considers the treatment effects of the interventions on our mental well-being outcomes. All dependent variables are normalized with respect to the control group's mean and standard deviation.

- The outcomes are: (1) self-reported depression from the PHQ-9 survey; (2) self-reported happiness on a scale from 1 to 5; (3) self-reported response to the "Cantril Scale" ladder of life possibility; (4) self-reported life satisfaction on a scale from 1 to 10; (5) self-reported stress on a scale from 1 to 6. Additional details of the outcome measures are located in Section C.3.
- All regressions are run at the level of the participant-day and control for gender, age, the baseline outcome variable, and day in study and date fixed effects.
- Standard errors in parentheses are robust to heteroscedasticity and clustered at the participant-level.
- When required, outcomes are flipped so that a positive value aligns with what would be considered a "better" outcome.

Table A.IX: Treatment Effects on Physical Well-being

	Standardized Components				
	Biking (1)	Illness (2)	Pain (3)	Daily Act. (4)	BP (5)
Night-Sleep Treatments	0.11	0.07	0.09	0.07	0.00
	(0.10)	(0.06)	(0.09)	(0.09)	(0.04)
Nap Treatment	-0.12	0.06	-0.06	0.13	0.04
	(0.09)	(0.04)	(0.08)	(0.08)	(0.04)
Participants	370	445	445	452	443

Notes: This table considers the treatment effects of the interventions on our physical well-being outcomes. All dependent variables are normalized with respect to the control group's mean and standard deviation.

- The outcomes are: (1) the average of standardized distance and standardized maximum speed recorded during the endline biking task; (2) number of self-reported illnesses over the last seven days; (3) self-reported pain on a scale from 1 to 10; (4) self-reported extent to which health has limited daily activities; (5) the average of standardized and winsorized systolic and standardized and winsorized diastolic blood pressure. Additional details of the outcome measures are located in Section C.4.
- The regressions control for gender, age, and the baseline outcome variable.
- Standard errors in parentheses are robust to heteroscedasticity and clustered at the participant-level for outcomes with multiple measures.
- When required, outcomes are flipped so that a positive value aligns with what would be considered a "better" outcome.

Table A.X: Treatment Effects on Lab Tasks Measuring Cognitive Function

	Inhibitory Control			**Memory**	**PVT**			
	(1) Payment	(2) Frac. Correct	(3) Avg. Reaction	(4) Payment	(5) Payment	(6) Inverse RT	(7) Minor Lapses	(8) False Starts
Night-Sleep Treatments	0.04	0.08	0.00	0.01	0.01	-0.02	0.02	-0.03
	(0.05)	(0.07)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)
Nap Treatment	0.05	-0.06	0.11**	-0.02	0.16***	0.12***	0.12***	0.02
	(0.04)	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
Participants	449.00	449.00	449.00	449.00	452.00	452.00	452.00	452.00

Notes: This table considers the treatment effect of the night-sleep and nap interventions on the three laboratory measures of cognition: inhibitory control, memory, and attention (PVT).

- All dependent variables are normalized with respect to the pure control group's mean and standard deviation (i.e., participants in neither the night-sleep nor the nap treatment group). In addition, signs are flipped when needed to ensure that for all variables higher indicates better performance.
- 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 (faster) reaction times. Columns 2 and 3 decompose performance by the fraction of correct entries, out of 40, and average reaction time (with the sign flipped), respectively.
- The outcome variable in column 4 is the payment participants earn for completing the Corsi blocks task, which measures working memory. Payment is a function of the maximum number of blocks the participant can recall.
- Column 5 is the overall payment for the PVT task. The payment, which was determined before the study began, is a function of three performance metrics in columns 6-8. Column 6 shows treatment effects for the inverse reaction time (reaction time captures how fast participants react to each stimulus). Column 7 depicts minor lapses (significant delays between when the signal appears and the participant acts). The outcome variable in column 8 is the number of false starts (when the participant acts before the signal is displayed). Signs are flipped for columns 7 and 8 such that positive values indicate *fewer* minor lapses and false starts (more desirable outcomes).
- All columns show the OLS estimates of equation (1), 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. Columns 5-8 also control for whether participants were randomized to the work group in the afternoon.
- Standard errors are clustered at the participant level.

Table A.XI: Treatment Effects on Attention to Work Incentives

	Overall		Morning		Afternoon	
	Output	Minutes	Output	Minutes	Output	Minutes
	(1)	(2)	(3)	(4)	(5)	(6)
Night-Sleep Treatments	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 Treatment	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 N-S vs Control	0.89	0.83	0.58	0.57	0.88	0.68
P-value Nap vs Control	0.01	0.11	0.28	0.71	0.02	0.06
Participants	451	451	450	450	451	451

Notes: This table considers the treatment effects of the night sleep and nap interventions on attention in the typing task. The task is described in greater detail in Section 3.2 under the sub-heading "Attention to Work Incentives".

- Each column shows the attention parameter (Gabaix, 2019) described in detail in Section C.5. This parameter varies from 0 to 1: a value of 0 means that participants don't react to high piece-rate incentives at all when the incentives are non-salient; a value of 1 means that the participants reacts equally to high piece-rate incentives under salient and non-salient conditions.
- Columns 1 and 2 include the entire day. Columns 3 and 4 only use observations from the morning (i.e., pre-naps). Columns 5 and 6 include observations from the afternoon (i.e., after the nap).
- The two dependent variables are Output and Minutes Actively Typing ("Minutes"), each of which are captured at the 30-minute incentive-session level.
- We consider attention for three groups: (1) the control group, consisting of individuals assigned to both the night-sleep control and the nap control; (2) the night-sleep treatments, pooling across both night-sleep interventions, and (3) the nap treatment group.
- In the lower portion of the table, Rows 4 and 5 depict the p-value of a test of differences between the coefficients estimated between the Night-Sleep Treatments and Control and the Nap Treatment and Control, respectively.
- Regressions include participant, date, and day-in-study fixed effects.
- Standard errors are clustered at the participant level.

Table A.XII: Treatment Effects on Time Preferences

Panel A: Savings

	Savings		Interest Accrued	
	Deposits	Net Savings	Real Pos. Rates	Hypothetical 1%
	(1)	(2)	(3)	(4)
Night-Sleep Treatments	-2.72	-9.10	0.36	0.03
	(9.25)	(11.77)	(1.71)	(0.98)
Nap Treatment	15.92*	9.60	2.05	1.47*
	(8.27)	(11.10)	(1.58)	(0.89)
Interest Rate	35.07***	39.88***	17.17***	4.01***
	(8.60)	(11.26)	(3.48)	(0.93)
Control Mean	113.29	71.97	10.63	8.70
Control SD	166.68	325.68	19.66	15.43
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 Treatments	0.01	0.05	0.01	0.05	0.01	0.05
	(0.03)	(0.05)	(0.04)	(0.06)	(0.04)	(0.06)
Nap Treatment	0.06**	0.08*	0.06	0.09	0.03	0.05
	(0.03)	(0.05)	(0.04)	(0.06)	(0.04)	(0.05)
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
Participants	352	214	352	214	398	252

Notes: This table considers the treatment effects of the Night-sleep and Nap Treatments on time preference.

- **Panel A: Savings.** This task is described briefly in Section 3.2 subsection "Savings", and in greater detail in Section C.6.2, along with a description of the related "defaults task."
  - 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 utilizing the full sample, but assuming all participants faced a 1% interest rate.
  - Each column shows the OLS estimates of equation (1), controlling for the baseline average of the dependent variable, age, gender, the fraction of high piece rates in the typing task , interest rate, maximum payment from cognitive tasks, a dummy for whether it is a risk and social activity day, the randomized piece rate for the present bias task, surveyor fixed effects, and the amount defaulted for savings. The regressions also include date and day-in-study fixed effects.
  - Standard errors are clustered at the participant level.
- **Panel B: Present-bias** This task is described briefly in Section 3.2 subsection "Effort Discounting", and in more detail in Section C.6.3.
  - The dependent variable in columns 1 and 2 is our preferred structurally-estimated present bias parameter, β. We exclude individuals for whom the structural estimator did not converge.
  - The dependent variable in columns 3-6 is the OLS present bias parameter, the percentage decrease in effort chosen on "work-days". In columns 3 and 4, we exclude the participants for whom the structural estimator did not converge. In columns 5 and 6 the sample includes all participants who completed the present bias task successfully at least once in the treatment period.
  - In Columns 2, 4, and 6 we also present results restricting the sample to the participants which engaged in the revised version of the present-bias task (See Appendix C.6.3 for more details).
  - In all columns, we control for the baseline value of the dependent variable and the gender and age of the participant.

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

	Risk Preferences		**Social Preferences**				
	Risk Aversion (1)	Loss Aversion (2)	Dictator Send (3)	Ultimatum Send (4)	Trust Send (5)	Ultimatum Receive (6)	Trust Send Back (7)
Night-Sleep Treatments	-0.11	0.01	-0.05	-0.01	-0.15	-0.01	-0.02
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)
Nap Treatment	-0.02	0.09	0.16*	-0.01	0.05	0.00	0.06
	(0.09)	(0.09)	(0.10)	(0.09)	(0.10)	(0.09)	(0.10)
Participants	383	403	415	415	415	415	415

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.
- Risk preferences components use the point at which the participant switched from the risky to safe choice in the risk aversion game (column 1) and the point at which the participant switched from the risky to safe choice in the loss aversion game (column 2) as the dependant variable.
- Social preferences components include the amount of money (in Rs) the sender sent in the dictator game (column 3), the amount of money (in Rs) the sender sent in the ultimatum game (column 4), the amount of money (in Rs) the sender sent in the trust game (column 5), the average amount the receiver would choose to accept versus reject in the ultimatum game, where a higher propensity to accept is considered the "good" outcome (column 6), and the average amount of money the recipient would send back to the sender in the trust game (column 7).
- Each column shows the OLS estimates of equation (2). Standard errors in parentheses are robust to heteroscedasticity.
- These Risk Preferences specifications differ compared to those underlying the indices in Tables III and IV. In particular, non-monotonic observations are excluded in the component risk preferences regressions but not in the indices. This accounts for the differences in the number of participants across columns, and this change was made to avoid dropping a large number of observations due to a single task and to simplify the indices.

Table A.XIV: Heterogeneous Treatment Effects for Main Outcomes of Interest

	Overall Index						
		X=Sleep Duration	X=Sleep Efficiency	X=Baseline Outcome	X=Baseline Naps	X=Female	X=Age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Night-Sleep Treatments	-0.01	0.01	-0.00	0.06	0.08	0.02	-0.08
	(0.04)	(0.06)	(0.05)	(0.05)	(0.07)	(0.07)	(0.06)
X		0.14*	0.06	0.57***	0.07	-0.02	-0.07
		(0.08)	(0.08)	(0.09)	(0.07)	(0.08)	(0.08)
Night-Sleep Treatments*X		-0.05	-0.02	-0.08	-0.12	-0.04	0.13
		(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)
Nap Treatment	0.12***	0.16***	0.12**	0.05	0.11	0.19***	0.12**
	(0.04)	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.05)
Nap Treatment*X		-0.08	-0.00	0.11	0.02	-0.11	-0.00
		(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)
Participants	452	452	452	452	452	452	452

Notes: This table considers the treatment effect of the night-sleep and nap interventions for different groups in the sample. The outcome variable in all columns is the overall index that aggregates over the four family-level outcome variables (corresponding to column 1 in Tables III and IV).

- Column 1 displays the treatment effect for the whole sample as a point of reference.
- In columns 2-7, we interact the treatments with a dummy that indicates whether the participant is above the median for a variable X measured during the baseline period. The X variables are:
  - Column 2: sleep duration as measured by the actigraph during baseline.
  - Column 3: Sleep efficiency (sleep duration/time in bed) as measured by the actigraph during baseline.
  - Column 4: The overall index itself during baseline.
  - Column 5: Whether the participant reported napping at least once a day before the beginning of the study.
  - Column 6: Whether the participant is female.
  - Column 7: The participants' age
- Regressions control for gender, age, and the baseline outcome variable. The nap treatment effects should be interpreted as the difference between treatment and both nap control groups (work and break jointly).
- Standard errors in parentheses are robust to heteroscedasticity.

C Detailed Description of Outcomes

C.1 Survey of Experts

C.1.1 Design

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): (i) a survey for general economists, (ii) one for behavioral economists, and (iii) one for sleep experts. 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 economist surveys also elicited predictions on savings, while the behavioral economist 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. The sleep science experts predicted the intention-to-treat (ITT) effect of the intervention, but the economists predicted the impact of a 1-hour increase in sleep duration. For the economists, we multiply their predictions by the first stage they were presented in the survey to recover the ITT prediction.

The survey has three main parts. In the first part, we introduce important information necessary to be able to take the survey. 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 SD of night sleep in the control and treatment groups; (iv) explanation of the data-entry task; (v) a benchmark, 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 the second part of the survey, we elicited the experts' predictions. The participants were informed that the treatment effect of the pooled night-sleep treatments on sleep was 32 minutes (the point estimate we had estimated with the available data at the time). Respondents were then informed about the level of the outcome variable during the treatment period for the RCT's control group participants and a table at the bottom of the screen mapped participant's answers to percentage and standard deviation changes for ease of interpretation. While all participants were asked to input their numeric prediction as the difference in levels of the outcome variable between treatment and control groups, the framing of the effect sizes varied according to common practices by field.

Finally, respondents were thanked for their time and invited to add their email address if they wished to receive information about the final results of the study or had comments about the survey.

C.1.2 Data Collection

Following a pilot among PhD students at Harvard and MIT, we sent 68 personalized emails to researchers known 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. In addition, a link to the survey was distributed to sleep scientists via multiple professional listserves.[38] In total, we gathered 122 surveys divided between sleep medicine experts ($N = 76$), behavioral economists ($N = 27$), and non-behavioral economists ($N = 19$).

Importantly, the results of the study 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.

[38]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.

C.2 Sleep Surveys in the RCT

Brief daily surveys about sleep quantity and quality were administered to all RCT participants each morning at the study office. These surveys elicited information about sleep the previous night including time to bed, time asleep, disruptions and their causes, time of awakening, time out of bed, subjective sleep quality, causes of poor sleep, and use of sleep devices. Participants were also asked about the timing and duration of any naps.

C.3 Well-being

We elicited a variety of outcomes related to mental well-being and mental health over the course of the study:

1. *Self-reported happiness:* Participants reported their happiness "today," where a score of 1 means "not at all happy" while a score of 4 means "very happy". Responses were recorded each morning the participant was in the office as part of the daily survey.

2. *Ladder of life possibility (Cantril Scale):* 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?" This question was included in the daily survey once every four days, where the particular day was randomly assigned for each participant.

3. *Life satisfaction (Gallup Survey):* Participants were asked, "All things considered, how satisfied are you with your life as a whole?" (1 Dissatisfied to 10 Satisfied). This question was included in the daily survey once every four days, where the particular day was randomly assigned for each participant.

4. *Self-reported stress (Cohen et al., 1983):* Participants reported their stress "in the last three days," where an answer of 1 means "none" of the time while 4 means "very often." This question was included in the daily survey once every four days, where the particular day was randomly assigned for each participant.

5. *Self-reported depression (PHQ-9):* Participants reported depressive symptoms using the PHQ-9. Responses were recorded during the baseline and endline surveys.

C.4 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*

- *Blood pressure:* Systolic and diastolic blood pressure were 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*

- *Self-reported illness:* Participants were asked about any symptoms of sickness (e.g., fever, cold, headache, etc.) they had 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 were asked to self report pain on a scale of 1 to 10, recorded at baseline and endline. *Pre-registered*

- *Daily Activity:* Participants were 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 were not limited at all in their daily life by their health to 36 for people who were 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*

C.5 Measures of Cognitive Function

C.5.1 Lab Tasks

Participants completed three laboratory-style tasks to measure attention, inhibitory control, and memory. The tasks, described in more detail in Dean et al. (2019), were conducted at varying frequencies in the afternoon.

1. *Psychomotor Vigilance Task (PVT).* Participants completed the PVT daily as a measure of simple attention. Developed by sleep scientists, the task asks participants to react to a series of randomly timed 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).

2. *Hearts and Flowers.* Participants completed the Hearts and Flowers task, a measure of inhibitory control (or one's ability to override impulses), every two days. The task includes three rounds during which participants are asked to touch keys in response to stimuli appearing on the screen. In the first round, participants are asked to touch a key on the same side of the screen as the stimulus appears. In the second round, participants, they are asked to touch a key on the opposite side of the screen as a stimulus. In the third round – which is scored and incentivized – participants continue the same reactions while the stimuli are intermixed. Performance is compensated based on a pre-specified mix of accuracy and speed in the third round.

3. *Corsi Block Span.* This task was also completed once every two days. The task measures visual memory by asking respondents to view a series of blocks which flash in a random order, and then repeat the series back in the same order using a touchscreen. Performance is compensated based on accuracy (the longest span remembered correctly).

C.5.2 Work Task

In addition to these laboratory style tasks, we embedded a measure of attention to incentives into the data-entry task, in an effort to provide a more economically relevant measure of attention. This task and the approach we take to measure attention in the participant-level is described in the body of the text in Section 3.2. Here, we provide an alternative estimation strategy which allows us to estimate the attention parameter in the spirit of Gabaix (2019).[39]

For each of the treatment groups j (i.e. night sleep, nap, and control), we estimate the (average) 'reaction' of output, productivity, and labor supply to the high piece rate, i.e. the difference in performance when piece rates are high compared to when pieces rates are low. We estimate this difference for days with salient incentives and for days with non-salient incentives, and denote it by ϵ_j^S and ϵ_j^{NS}, respectively.

[39]The reason this is not our main measure is because it is not amenable to being transform into participant-level indices.

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 $\theta_j = 0$.

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, and Sal_{it} whether participant i was randomized to the salient condition on day t.

This equation differs from the benchmark reduced-form regression (1) 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 $\hat{\epsilon}_j^{S} = \hat{\beta}_1^j + \hat{\beta}_3^j$. We then 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.

C.6 Preferences

We gathered data on three types of preferences: risk and loss preferences, social preferences, and time preferences. Time preferences included two measures: a savings opportunity and a real-effort task. In addition, the savings opportunity was overlaid with variation to examine whether one's propensity to overrule defaults was influenced by sleep. Each of these tasks is described in greater detail below.

C.6.1 Risk and Social Preferences

We measure risk and social preferences via standard tasks in the behavioral economics literature. We elicited these preferences twice, once during the pre-treatment period (day 7) and once at the conclusion of the study (day 26).

Risk preferences. 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). Following the literature in this space, the point at which the participant switched from the safe choice to the risky choice is taken as the primary outcome of interest.

Social preferences. Social preferences are measured via dictator, ultimatum, and trust games (Camerer, 2003). Participants were randomly matched and did not know who their specific partner was. Outcome measures were chosen to be consistent with the literature and included: the amount of money the sender sent in the dictator game, ultimatum, and trust games, 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.

C.6.2 Savings Task

Additional details of task design. As described in Section 3.2, participants were offered the opportunity to save at the study office at a favorable interest rate. These 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. 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. As described in more detail in Table A.XII, our main outcome measures are (i) daily deposits; (ii) daily net savings (deposits minus withdrawals); and (iii) daily interest accrued on savings.

Construction of counterfactual interest accrued variable. Our measure of savings accrued due to interest excludes participants randomized to 0% and disproportionately weights individuals who were assigned to 2% interest rate. To avoid this bias, we built an alternative measure of accrued savings by applying an hypothetical homogeneous 1% interest rate. 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. For instance, that would be the case for participants deciding to withdraw quantities at day 10, $x_{10} < 0$. 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 above procedure for the baseline period, setting $s_1 = 0$. We then regress savings during the treatment indicators, controlling for the total interest accrued during baseline.

Default task. 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 strength 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 strength 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. Additional details and results are available upon request.

C.6.3 Present Bias

Overview. Our design follows Augenblick and Rabin (2019) and Augenblick et al. (2015). The participants completed a real-effort task, making decisions about how many pages to type on a fixed date ("work day") under different piece rates. The work was very similar to the data-entry work completed each day, except that the pages were shorter to allow for a finer choice set for the participants. The work for this task was completed at a fixed time after the completion of their regular working day, but before their daily payment.

Choices. Participants had to make a total of 14 decisions. For each choice, the participants were offered a piece rate w^c per page completed and needed to choose how many pages they would like to type at that piece rate. Participants had to choose at least 5 pages, which we imposed to avoid fixed costs associated with moving from 0 pages to 1 page (Augenblick et al., 2015). We also imposed a participant-specific upper limit to the number of pages the participants could choose, $\max_i$, to ensure the task could be completed on time.[40] Immediately after the participant made their last decision, we randomly selected one of the decisions to be the one that counts. For example, if decision c was selected, the piece rate associated with

[40]The limit of pages was calculated based on their typing speed up to that point in the study. We imposed this limit because sleep could impact risk-aversion, which would then affect participants' decision-making.

that choice, w^c, and the participant's choice, e^c, would be the piece rate and the output target of the participant for the task.

Timeline. The decisions were made on two different dates: on a date prior to the work day (prospective date) and on the work day. The prospective date was chosen to be 1 to 5 days before the work date. The payment date was always at least one day after the work day. Moreover, the payment date was a function of the randomly selected choice. We designed it so the payment distance was fixed between the date of a given choice and the payment if that choice was randomly selected to be the one that counted. Participants completed the present bias experiment once during the baseline period and at least once during the treatment period.

Earnings from the task. Earnings from the tasks consisted 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 participants were only paid if they completed all the work they had committed to within two hours, otherwise they received nothing from the present bias task.

Changes during the study. Debriefing of participants who had already completed their participation revealed that they would often make the same choices across the two dates in an effort to stay consistent. Since such behavior would make it difficult for us to identify present-biased preferences, we made two modifications to the task during the study. First, instead of offering the same piece rates in the two dates of the task, the piece rates on each day were slightly modified. We randomized which of the piece rates were offered on each day of the task. Second, to allow more time to elapse between the two choices, we reduced the number of times participants completed the present bias task in the treatment period from three to one.

Exclusion criteria. Of the 452 participants in the study, we cannot estimate a present bias parameter for 54 individuals. 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 chose the maximum or always chose the minimum number of pages during the treatment period. We exclude these participants since we cannot identify time preferences parameters for them. Of the remaining 398 participants, we cannot estimate the structural β for 46 because the algorithm does not converge. In our preferred specification we also exclude them, leaving us with a final sample of 352 participants. The fraction of participants excluded for these criteria is balanced across groups.

Structural estimation of present bias. We estimate individual-level short-term discounting parameters β assuming participants chose the number of pages they would like to type by maximizing the utility function

$$U(e, w, k, t, T) = -\beta^{-D_{k,t}}\delta^{t-k}C(e) + \delta^{T-k}U_m(e \cdot w), \tag{4}$$

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 Augenblick and Rabin (2019) (AR, henceforth), we assume the cost function has a power form in our benchmark specification, i.e.

$$c\left(e\right) = \frac{1}{\gamma}e^{\gamma}. \tag{5}$$

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

$$U_m(e \cdot w) = \phi \cdot w \cdot e + \alpha \cdot e. \tag{6}$$

The first term of this function captures the utility of money. We found that some participants also appear to have an intrinsic motivation for working, which based on participants' debriefings is often linked to either reputation building (although we were explicit that we just want to know their preferences) or gift exchange (DellaVigna and Pope, 2018). We capture this effect with the term $\alpha \cdot e$ above.

In this model, optimal effort is given by

$$e^* \equiv e^*(k,t,T,w) = \left[(\phi \cdot w + \alpha) \, \frac{\delta^{T-t}}{\cdot \beta^{\{t>k\}}} \right]^{\dfrac{1}{\gamma - 1}} \tag{7}$$

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 number of pages we observe being chosen is determined by

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

We estimate the utility parameters in (4) using a 2-sided Tobit model, with cost function (5) and return to effort (6). We also impose that $\delta = 1$. We do this because due to absences, some participants performed the second day of the task on later than originally planned, thus creating non-random variation in the timing between the two days of the task.

We estimate the model twice per participant: (i) using data from the baseline period; (ii) and using data from the treatment period. We thus estimate one baseline and one treatment period estimate of present bias per participant. The structural estimation does not converge for 46 participants in the treatment period in our preferred specification, so we drop those from the sample. The structural estimation also does not converge for 10 participants in the baseline period. We replace those missing values with the average value across participants during baseline.

Correlates of present bias. The structurally estimated present-bias parameter correlates with behaviors that one might expect to be affected by time preferences (Table A.XV). More present-biased participants (i.e. those with lower β) saved less (columns 1-2) and arrived late in short days more often (columns 3-4) than less present-biased participants.[41] Interestingly, our estimates of present bias do not correlated with labor supply (columns 5-6) and sleep duration (columns. 7-8). The latter suggests that self-control may be a less important determinant of low sleep duration than found in rich countries (Avery et al., 2019).

[41] On Short Days participants received a financial incentive to arrive on time, as described in Section 3.2.

Table A.XV: Relationship between Present Bias (β) and Behaviors Involving Time Preferences

	Daily Deposits		Lateness		Typing Time		Night Sleep	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Beta Structural	43.45**	42.33*	-3.65**	-3.74**	0.19	0.21	0.05	0.08
	(21.15)	(21.60)	(1.82)	(1.88)	(0.17)	(0.17)	(0.15)	(0.15)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	111.79	111.79	5.36	5.36	4.27	4.27	5.59	5.59
Control SD	103.10	103.10	12.20	12.20	0.90	0.90	0.83	0.83
Participants	352	352	352	352	352	352	352	352

Notes: This table reports the OLS coefficient between the structurally estimated present bias coefficient (β) and participant behaviors that we expect would be affected by present bias.

- 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 (in Rs) in columns 1 and 2; lateness in minutes on "short days" (i.e. the maximum between zero and arrival time - 11am) in columns 3 and 4; typing time (measured in hours) in columns 5 and 6; and hours of night sleep (measured by actigraph) in columns 7 and 8. All dependent variables are study-long averages (including the baseline period).
- Columns 1, 3, 5, and 7 have no controls. Columns 2, 4, 6, and 8 include controls for participants' age and sex.

C.6.4 Treatment Effect on Present Bias

To estimate the treatment effect of the night sleep and the nap interventions, we estimate equation 2 with two different outcome variables: (i) the individual-level structurally estimate of present bias, β; (ii) the OLS estimate $\hat{\beta}_i^{raw}$ from the regression

$$\log e_{cit} = \beta_i^{raw} \text{Now}_{cit} + \gamma_i^0 + \gamma_i^1 \log w_{cit} + \varepsilon_{cit} \tag{8}$$

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 results can be found in Table A.XII.

C.7 Willingness to Pay

Overview. 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.

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. Low willingness to pay could also be consistent with beliefs that the devices themselves are not productive in generating additional sleep.

D Deviations from Pre-Analysis Plan and Original Study Design

This study was pre-registered on the AEA RCT Registry (ID: AEARCTR-0002494) under the title "Sleepless in Chennai: The Consequences of Sleep Deprivation Among the Urban Poor." Pre-registration took place on December 8, 2017, shortly after the start of our study. By the time of the pre-registration, only 7 participants had completed the study (recall the rolling enrollment scheme), and we had not started analyzing any of the data. 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 because of 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.

D.1 Family of Outcomes and MHT Correction

The PAP defined two core families of outcomes, work and decision-making, and noted that multiple hypothesis corrections would be run within these families. Given that we realized that some of the outcomes in reality do not fit well under the umbrella of "decision-making", and to include some of the additional outcomes that we had pre-registered separately, we decided to instead create three families in addition to the work family: well-being, cognition, and preferences, as reported in Tables III and IV. We ran multiple hypothesis corrections both across the relevant family indices as well as within each family, among the component outcomes that comprise that family (for more details, see Appendix E).

D.2 Data-Entry Task

- **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 well balanced across treatment groups (Figure A.VIII) and excluding missing observations improves statistical power without changing results qualitatively (results in the working version of this paper, Bessone et al. (2020)).

- **Typing earnings variable.** The PAP specifies 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.

- **Output.** Earnings, labor supply, and productivity were part of our original work family. In addition, we also report output given that this outcome was of interest to some readers and referees.

D.3 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 large withdrawals close to the end of the study. We believe these withdrawals were driven by the study design rather than participants underlying savings behavior. Hence, in addition to this measure, we use daily deposits and interest accrued.

- **Interest rates.** Interest rates were changed to improve participant understanding and to allow us to estimate 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.
- **Default**. The study included an outcome capturing adherence to a default amount that was automatically added to the participants' lockbox. We pre-registered that we would analyze the treatments effects on adherence to default. However, the main effect (the adherence to default) was relatively small and we were hopelessly under-powered to detect effects on top of the main effect. The treatment effects would have needed to be almost 100 percentage points in adherence to the default to be statistically significant. Following the prescription in Duflo et al. (2020), we exclude the adherence to default outcome from the paper. We show the results of the default task in the working paper version of this paper (Bessone et al., 2020).

D.4 Preferences and Cognitive Function

- **Present bias.** There are three deviations from the pre-registered analysis:
  1. We assume that $\delta = 1$, rather than estimating it from data. The reason we do this is that often the variation in distance between the decision day and the payment day was driven by absent days, which are not-random.
  2. We do not estimate treatment-group-specific parameters, as we pre-registered we would. Instead, we estimate individual-level present-bias parameters (which we also said we would do in the pre-registration). We do not estimate treatment-group-specific parameters because the specification with individual-level parameters is economically more sound than assuming homogeneous preferences withing treatment groups.
  3. In equation (5) of the PAP, we specify that we will run a semi-parametric specification for present bias. We estimate it but with two modifications. First, instead of using the number of pages chosen as an outcome, we use the number of pages chosen divided by the maximum number of pages participants can choose. This approach ensures that we do not give more weight to participants who could select more pages. Second, we do not include date, day in study, and surveyor FEs. That was a mistake, since the tasks occurs over multiple days, which does not allow us to control for these FEs.
- **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. In the first version of the task, the only difference between the salient and the non-salient conditions was that in the salient condition, the incentives were shown in different colors in the bottom of the screen, while in the non-salient condition, incentives were always show in the same color. In the second version of the task, we added two additional features. First, in the salient condition the screen blinks twice when the incentives change to ensure that participants would notice the change in piece rate quickly. Second, in the non-salient condition, the incentives faded away after 15 seconds, thus allowing for more scope for participants to miss incentive changes in the non-salient condition.

D.5 Risk and Social Preferences

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

D.6 Well-being

- **Outcome components.** Our AEA pre-registration included two measures to rely on when creating the subjective well-being index - happiness and life possibilities. We also registered our study at ClinicalTrials.gov (NCT03322358) and included depression in this registration. We later added questions on life satisfaction and self-reported stress.

E Multiple Hypothesis Corrections

We applied multiple hypothesis corrections both across and within our four families of outcomes: (i) work, (ii) physical and mental well-being, (iii) cognition, and (iv) preferences. The outcomes and adjusted p-values are reported in Tables III and IV.

To apply all corrections we ran simulations to control the Family-Wise Error Rate. The corrections are applied separately for each treatment. We took this approach rather than applying a formulaic correction (e.g. Holm or Bonferroni) in order to 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 1,000 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 1000 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 adjusted p-value.

4. Repeat for each family of outcomes.

F Broader 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. Households were approached starting from those locations and walking in a predetermined pattern. 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.XVI, and the demographic characteristics across the first two stages can be found in Appendix Table A.XVII.

Findings. The first key takeaway is that this broader sample of Chennai is severely sleep deprived, sleeping just 5.5 hours on average per night according to the Actigraph.[42] 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, as measured by 71% sleep efficiency.

In our sample, sleep characteristics do not vary substantially by household income, education, or employment status (Table A.XVIII).[43] Despite these similarities, sleep does vary with some demographic factors. Women sleep more than men and households with more children sleep less. However, these differences are small, between 5 and 15 minutes. Finally, middle-aged individuals sleep approximately 30 minutes less than younger or older adults. The survey also revealed that daytime naps are common in this population. 37% of individuals report napping on any given day. Higher-income individuals are roughly 10 percentage points less likely to nap, but conditional on napping spend more time asleep. Older participants are also more likely to nap on any given day.

[42] 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.

[43] 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. Hence, no strong strong conclusions about the sleep of higher income populations in this context should be drawn.

Table A.XVI: Sleep Survey Stagewise Take Up

	Percent of Last Stage (1)	Percent of Total (2)	Frequency (3)
Census	49.93	49.93	3833
Baseline Survey	44.18	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 conducted among a broader population in Chennai.

- N represents the total number of participants approached for the study, including all refusals to participate in any portion of the survey.
- "Census" indicates participants willing to respond to a basic demographic questionnaire.
- "Baseline Survey" captures people who completed the full baseline survey, including information about their sleep.
- "Interest to hear about actigraph" and "Willingness to wear actigraph" indicate that the participant listened to a description of the actigraph request and accepted, respectively.
- "Actigraph installation" captures whether the participant was loaned an actigraph to wear. Not all willing participants were given an actigraph for multiple reasons such as non-availability of the participant on the day of installation, shortage of actigraph devices at our disposal, and compliance with the upper limit of installing 20 actigraphs per locality.
- “Actigraph component participants (all)" includes all participants who wore the actigraph for at least one night.
- “Actigraph component participants (completed)" includes only those participants who complied with the study’s requirement of wearing the actigraph for three nights.

Table A.XVII: 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	0.17
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	45.80	45.17	45.76
	(15.72)	(15.07)	(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 sleep survey - the census, the baseline survey, and the actigraph component.

- With the exception of Age and Highest grade attended, all statistics represent the fraction of respondents in each category.
- Gender, age, employment status, and education are all self-reported by the participant. Employed individuals are coded as "1" while those who report being unemployed, housewives, and retired without a pension are coded as a "0".
- Income was categorized in three ways: (1) the participants' self-report in the baseline survey; (2) an estimate based on the surveyor's observation of the participant's house; (3) an estimate based on the surveyor's observation of the participant's neighborhood.
- 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 percent reporting each income category do not sum to 100 because participants could respond that they "do not know" and "do not want to disclose."

Table A.XVIII: Sleep Survey - Sleep Correlates

	Self-reported Night Sleep		**Actigraph Night Sleep**			**Actigraph Total Sleep**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	PW
Self-reported Sleep Q2			0.54***			0.41***			
			(0.10)			(0.11)			
Self-reported Sleep Q3			0.59***			0.55***			
			(0.10)			(0.11)			
Self-reported Sleep Q4			0.88***			0.80***			
			(0.10)			(0.11)			
Middle income	0.07	0.05		0.20	0.17		0.15	0.13	0.15
	(0.10)	(0.10)		(0.12)	(0.13)		(0.12)	(0.13)	(0.13)
Higher income	0.06	0.00		0.04	0.03		0.02	0.04	0.09
	(0.11)	(0.11)		(0.13)	(0.14)		(0.13)	(0.14)	(0.14)
Female	-0.04	-0.07		0.19**	0.18*		0.14	0.12	0.20
	(0.05)	(0.06)		(0.09)	(0.10)		(0.09)	(0.11)	(0.13)
Age 34 - 45	-0.54***	-0.50***		0.01	0.01		0.03	0.00	-0.11
	(0.06)	(0.06)		(0.11)	(0.11)		(0.11)	(0.12)	(0.12)
Age 46 - 58	-0.58***	-0.52***		-0.55***	-0.56***		-0.49***	-0.53***	-0.75***
	(0.07)	(0.07)		(0.11)	(0.12)		(0.12)	(0.13)	(0.13)
Age 59 - 92	-0.45***	-0.40***		-0.06	-0.08		-0.02	-0.06	-0.31**
	(0.07)	(0.07)		(0.11)	(0.11)		(0.12)	(0.12)	(0.14)
Children (#)	-0.07***	-0.06**		-0.10**	-0.11**		-0.09*	-0.10**	-0.11**
	(0.03)	(0.03)		(0.05)	(0.05)		(0.05)	(0.05)	(0.05)
Some school		-0.11			0.26*			0.21	0.14
		(0.11)			(0.13)			(0.13)	(0.17)
College		0.05			0.12			0.02	-0.17
		(0.12)			(0.15)			(0.16)	(0.19)
Employment		-0.09			-0.01			0.01	-0.01
		(0.06)			(0.09)			(0.10)	(0.12)
Constant	6.90***	6.99***	4.98***	5.42***	5.28***	5.27***	5.69***	5.60***	5.79***
	(0.12)	(0.17)	(0.07)	(0.16)	(0.20)	(0.08)	(0.16)	(0.22)	(0.26)
Mean of DV	6.49	6.49	5.45	5.45	5.45	5.69	5.69	5.69	5.65
N	3389	3387	1367	1367	1367	1367	1367	1367	1367
Participants	3387	3387	439	439	439	439	439	439	439

Notes: This table considers correlations between participant demographics and sleep habits.

- Column 1 considers nighttime sleep as self-reported by the participant in hours. Columns 2 and 3 examine actigraph measurements of nighttime sleep in hours. Columns 4 and 5 include total hours of sleep per 24 hour period, summing actigraph measures of night sleep and naps. Column 6 is an indicator for whether a participant naps on a given day, as measured by the actigraph. Column 7 measures the duration of the nap in hours, conditional on napping, as measured by the actigraph.
- Covariates in the rows include: (1) sleep quartiles based on the participants' self-reported sleep, (2) income categories derived from the surveyor's assessment of the income level of the participant's neighborhood, (3) the participant's self-reported gender, (4) the participant's self-reported age (binned), (5) the number of children in the household, (6) the participant's completed education binned as "never attended school" (omitted), some school up to but not including college, some college or more, (7) whether the individual reports being employed (where unemployed, housewives, and retired without a pension are the omitted category).
- Dependent variables are recorded on the participant-day level. Standard errors are clustered at the participant level.
- N indicates the total number of observations (participant-days) and "Participants" indicates the number of participants.

Appendix References

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

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

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

Bessone, Pedro, Gautam Rao, Frank Schilbach, Heather Schofield, and Mattie Toma, "The Economic Consequences of Increasing Sleep Among the Urban Poor," Technical Report, National Bureau of Economic Research 2020.

DellaVigna, Stefano and Devin Pope, "What motivates effort? Evidence and expert forecasts," *The Review of Economic Studies*, 2018, *85* (2), 1029–1069.

Dinges, David F., 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.

Duflo, Esther, Abhijit Banerjee, Amy Finkelstein, Lawrence F Katz, Benjamin A Olken, and Anja Sautmann, "In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in economics," Technical Report, National Bureau of Economic Research 2020.

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.

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