The Economic Consequences of Increasing Sleep Among the Urban Poor[*]

Pedro Bessone[†]   Gautam Rao[‡]   Frank Schilbach

Heather Schofield   Mattie Toma

March 15, 2021

Abstract

The urban poor in developing countries face challenging living environments, which may interfere with good sleep. Using actigraphy to measure sleep objectively, we find that low-income adults in Chennai, India 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 by inducing more time in bed. Contrary to expert predictions and a large body of sleep research, increased nighttime sleep had no detectable effects on cognition, productivity, decision-making, or well-being, and led to small decreases in labor supply. In contrast, 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 cognition, but a decrease in work time.

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, water, education, healthcare, and clean air. This paper considers a previously-unexamined challenge faced by the urban poor in developing countries: sleep deprivation. People in these settings 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 Chennai, India. Our two samples of low-income adults 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. People spend about 8 hours per night in bed, but their sleep is highly disrupted, with 31 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 or insomnia in high-income countries (Hedner et al., 2004; Trauer et al., 2015).

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 mood and health (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 119 experts from sleep science and economics who 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 adults in Chennai. We employed participants for a one-month data-entry job with flexible hours, allowing us to precisely measure productivity and labor-supply, as well as physical and psychological well-being, cognition, and time, risk, and social preferences. Two *night-sleep treatments* gave participants (i) items to improve their home-sleep environments, (ii) information and verbal encouragement to increase their night sleep (the 'encouragement' treatment), and (iii), for a subset of participants, additional financial incentives to increase night sleep (the 'incentives' treatment). These treatments were cross-randomized with a *nap treatment* that offered participants the opportunity for a daily half-hour afternoon nap at their workplace.[2]

---

[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]Altogether, our design results in a control group and five treatment groups (encouragement only, incentives only, naps only, encouragement and naps, incentives and naps). In addition, all those who did not receive naps were randomized each day to either work through the nap period or to take a mandatory break, allowing us to compare naps against work days and break days separately.

The night-sleep treatments on average increased nighttime sleep by 27 minutes per night (standard error (SE)=3 mins), with larger effects for the incentives (33 mins) than for the encouragement treatment (20 mins). The increased sleep duration was entirely driven by additional time in bed—on average 38 minutes per night (SE=4 mins)—rather than higher sleep efficiency. These results demonstrate that people do have substantial ability to adjust their nighttime sleep through changes in time in bed, but may not be able to increase their sleep efficiency. The low sleep efficiency increases the opportunity cost of sleep: raising sleep duration by 1 minute requires 1.4 more minutes in bed.

Similarly, the nap treatment increased daytime sleep by about 14 minutes per day on average (SE=0.3 mins), while slightly crowding out nighttime sleep. While both types of treatments increased 24-hour sleep, the 27-minute increase from the night-sleep treatments was significantly higher than the 8-minute increase due to the nap treatment ($p<0.01$).

We first examine the effects of each combination of treatments on an overall summary index which aggregates all outcomes as in Anderson (2008). Each of the night-sleep treatments alone had no significant effect on this overall index: 0.00 standard deviations (SD, SE=0.07) and -0.05 SD (SE=0.07), respectively, for the 'encouragement only' and 'incentives only' groups. In contrast, 'naps only' had a positive, marginally significant effect of 0.11 SD (SE=0.07, $p=0.10$). Those who received a night-sleep treatment in addition to naps had very similar effects to those with naps only.[3] This pattern of results suggests that naps have an overall positive effect on outcomes, while increases in night sleep do not. To increase statistical power and to streamline the discussion of effects on the individual outcomes, we turn to an analysis which pools the two night-sleep treatments and does not allow for an interaction effect of night-sleep and nap treatments. The effects of each treatment described below should thus be interpreted as conditional on the distribution of the other treatment (Muralidharan et al., 2019).

In the pooled analysis, we find no significant effect of increased night sleep on the overall index (-0.01 SD, SE=0.04), or on summary indices corresponding to four families of outcomes: work, well-being, cognition and economic preferences. In fact, the pooled night-sleep treatment had no significant positive effect on *any* outcome other than sleep itself. It did not significantly increase productivity at the data-entry job, 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 reject the median expert prediction of a 7% increase in output ($p<0.001$). Similarly, we find no significant effects on detailed measures of physical and psychological well-being, or standard measures of cognition and 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

---

[3]The overall effect of 'naps only' differs significantly from that of 'incentives only' ($p=0.02$) and suggestively from that of 'encouragement only' ($p=0.11$). Naps combined with the encouragement treatment had an effect of 0.13 SD (SE=0.07) while naps combined with incentives had an effect of 0.09 SD (SE=0.06), neither of which is significantly different from 'naps only'.

in our setting—as proxied by low efficiency and frequent awakenings—explains the lack of benefits of increased nighttime sleep. Returns to increased sleep could be higher 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.

In contrast to night sleep, naps significantly improved a range of outcomes. The nap treatment significantly increased the overall summary index by 0.12 SD (SE=0.04, p=0.00), as well as the index variables corresponding to well-being (0.08 SD, p=0.02), cognition (0.10 SD, p=0.08), and preferences (0.07 SD, p=0.10). The effect of naps on work outcomes depends upon the comparison. Compared to taking a break, naps increased earnings—the summary variable for the work outcomes—by 0.05 SD (p=0.05). However, driven by a decrease in labor supply, naps *reduced* earnings compared to working during the nap time by 0.10 SD (p=0.00), highlighting the importance of taking into account the opportunity costs of sleep.

Considering the individual outcome variables one by one, and adjusting for multiple comparisons, naps have significant positive effects on productivity (0.04 SD, p=0.06), psychological well-being (0.12 SD, p=0.04), lab measures of cognitive function (0.08 SD, p=0.06), and attention at work (0.21 SD, p=0.07). Naps may also increase patience (p=0.11), with point estimates suggesting increases in saving and reductions in present bias in effort choices.

The nap and night-sleep treatments have statistically different effects on the overall index (p=0.02). We are less-powered to detect differences for the family-level indices. Point estimates are larger for the nap treatment for well-being, cognition, and preferences, but the differences are not statistically significant. At the level of individual outcomes, we find statistically significant differences in effects only for psychological well-being (p=0.04). Estimating the *per-minute* effects of the two types of sleep in an instrumental variable analysis, however, we can reject equal per-minute effects of naps and nighttime sleep on the summary index (p<0.01), on the family indices for cognition (p=0.05) and well-being (p=0.08), and on some individual outcomes including labor supply (p=0.04) and psychological well-being (p=0.05). In every case but labor supply, the effects are more positive for naps.

One possible reason for the different effects of the nap and night-sleep treatments is 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 simply 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 nighttime 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).[4] We speculate that the stark difference in sleep efficiency in our setting compared to rich-country populations contributes to this difference. It could also be that increased night sleep simply does not have high benefits at the margin in the field, as in another recent field experiment (Avery et al., 2019). Additional field studies of sleep, including interventions that improve sleep efficiency, may help reconcile these findings.

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 life around the world and are particularly prevalent in tropical countries (Dinges, 1992). 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, or consideration of the opportunity cost of naps (Lovato and Lack, 2010; Ficca et al., 2010). Our work takes a step towards filling this gap, and shows that naps may be an effective way to combat sleep deprivation. The decline of naps as employment in developing countries shifts towards Western schedules could therefore be costly.

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—in this context—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), low-income people in Chennai may be 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), by recording 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 bulky technology is impractical for field studies and may itself interfere with natural sleep patterns at people's homes, thus making measuring sleep outside of sleep labs challenging. Self-reported measures are unreliable and correlate only moderately with objective measures since people tend to report time in bed rather than hours asleep, leading to over-reporting of sleep duration (Lauderdale et al., 2008; Schokman et al., 2018).

---

[4]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.

Actigraphs, which resemble wristwatches and infer sleep/wake states from body movement, 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).[5] 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 contrast, there is scant evidence on sleep patterns in developing countries. 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—which typically overestimate sleep—suggest a substantial share of people in developing countries sleep less than the recommended 7 to 9 hours. 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).

Two recent studies in low-income countries measured sleep using actigraphs and identify significant 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 in a rural population without electricity.

---

[5]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 with high sleep efficiency (Jackson et al., 2018; Cespedes et al., 2016).

2.3 Sleep(less) in Chennai

Our first measure of sleep in Chennai comes from the RCT sample of 452 adults recruited for a full-time data-entry job for one month. To capture reliable and objective measures of sleep beyond self-reports, all participants wore actigraphs continuously throughout the study.[6] Below, we describe sleep during the baseline period (before treatment) in this sample. We then report very similar patterns of sleep in a broader sample in Chennai, which wore actigraphs for three nights.

A Typical Night in Chennai. We first 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. This night closely matches the average time in bed, sleep duration, and sleep efficiency in the RCT sample. The participant spends about 8 hours in bed during this night but only achieves 5.6 hours of highly fragmented and interrupted sleep, involving over 30 awakenings. For comparison, we show a less interrupted night with 90% efficiency in Figure Ib. While this night is unusual in Chennai—only 1% of nights in our sample feature such high sleep—it resembles nights of healthy adults in high-income countries who typically enjoy sleep efficiency of 85 to 95 percent (Carrier et al., 2001; Cole et al., 1992; Walker, 2017).

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.[7] Despite this significant time in bed, study participants only enjoy 5.6 hours *asleep* per night (Table I and Figure IIc), significantly below both time in bed and the recommended 7 to 9 hours. 95% of participants slept less than 7 hours per night, and 71% slept less than 6 hours per night on average. 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. Average sleep efficiency in our sample is 70% (Figure IIe), far below recommended levels by sleep scientists who found that a minimum of 85% is needed to indicate "high-quality" sleep (Ohayon et al., 2017). Like sleep duration, sleep efficiency is much lower than typically found in high-income countries, and instead resembles US-based patients suffering from sleep disorders such as sleep apnea (Roure et al., 2008) or insomnia (Trauer et al., 2015). Sleep efficiency is low throughout the night, remaining around 70% between 1 to 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 col 1), again comparable to insomniacs in the US (Lichstein et al., 2006).[8]

Barriers to Sleep. Why is sleep so inefficient? Survey responses highlight the importance of mental and physical distress (e.g. worries, stress, pain, or hunger) as well as environmental factors

---

[6] Section 3 provides details on the sample and study design. 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 groups, with approximately 6% of participants removing the device on any given day.

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

[8] An awakening is defined as a disruption lasting at least 30 seconds. 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).

(Figure A.IVa). Over 50% of study participants indicate that their sleep is disrupted by heat, noise, and/or light, all of which the night-sleep treatments described below were intended to address.

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 similar: Dinges (1992) finds that, across a broad population of US adults, 61% reported 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 in the US (Lauderdale et al., 2008; Avery et al., 2019).[9] 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.[10] A subset (N=439) completed three nights of actigraph measurements. Despite not using any of the RCT screening criteria for this survey, the nighttime sleep duration and efficiency in this broader sample are similar to that of RCT participants, with an average of 5.5 hours of sleep per night and 71% sleep efficiency (Table I col 2). As in the RCT sample, napping is common, with 25% of participants napping on any given day and an average duration of those nap of roughly 50 minutes.

3 Experimental Design

We designed our experiment with three broad goals in mind: First, we aimed to estimate the impacts of increased sleep over a few weeks in relatively natural sleep environments (as opposed to depriving people of sleep in lab settings). Second, we wanted to precisely measure labor supply, productivity, and earnings, and thus employed study participants full-time in a realistic but closely controlled data-entry job. Finally, to provide a broad view of the impacts of increased sleep, we collected a range of additional outcomes, including cognition, preferences, and well-being.

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.

---

[9]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).

[10]For more detail of the survey and population, see Appendix F.

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 about 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).[11]

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.[12] 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.[13]

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 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,

---

[11]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.

[12]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.

[13]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.

or putting away one's cellphone at night could also plausibly increase sleep efficiency. Finally, the loaned devices could increase both sleep efficiency and time in bed, if the devices made it easier to fall asleep or reduced awakenings, 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 nighttime 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 day, on which we allowed them to work through the 'nap period', or a break day, on which we enforced a half-hour break from data entry during the same period. Break day participants were allowed to engage in any leisure activity they chose, including sitting in a comfortable office break room. By comparing nap and break participants, we isolate the effect of a nap relative to a break of the same length. By instead comparing the nap and work participants, we can estimate the net effect of naps on work output, including the lost work time.

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 psychological well-being, 2) cognition, and 3) economic 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.

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.[14] 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.

---

[14]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.

Labor Supply. Our pre-registered measure of labor supply is the *active* typing time as automatically measured by the data-entry software. As in many workplaces, participants were not forced to arrive or leave the office at precise times. On most days, participants could arrive or depart from the office as they chose between 9:30 am and 8 pm. When in the office, participants could take breaks from work. 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 fair amount 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.[15]

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.

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.[16]

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 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

---

[15]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.

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

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 (in the second half of the study) the rate was only visible for the first 15 seconds of a half-hour slot, fading out slowly. 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.[17]

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 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.[18] 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.

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).

Realism and External Validity. Conducting the study in the context of a month-long data-entry job in a controlled environment follows Kaur et al. (2015) and allows for the provision of afternoon naps and precise measurement of the outcomes described above. However, it also comes with some costs. First, the work environment has some unusual and artificial features, such as regular surveys and laboratory measures of cognition and preferences. Second, any labor-supply responses we

---

[17]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.

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

find might be muted in environments where employers more strictly control schedules. In practice, participants tend to spend about 8 hours at the office each day, with an average arrival time of 10.32 am (std. dev.= 43 minutes) and average departure time of 6.20 pm (std. dev.= 57 minutes). This is quite similar to other jobs in our context, with long commutes and unreliable transportation, such that arriving strictly at a given time is difficult.

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 the night-sleep interventions included in 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.[19]   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 the survey short, we did not elicit predictions about the effects of the nap treatment. All experts were provided with relevant benchmarks (e.g. the elasticity of effort with respect to the piece-rate) and 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.

---

[19]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.

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 including information about the work task, other experimental measures and surveys, the actigraphs and the randomized treatments. The specific research hypotheses were not shared with participants to avoid demand effects. Instead, the goal of the research was described as work 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.[20]

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:

$$y_{itd} = \beta \mathbf{T}_i + \gamma_1 \bar{y}_i^B + \gamma_2 \mathbf{X}_{it} + \delta_t + \lambda_d + \varepsilon_{itd}, \tag{1}$$

where y_{itd} is the relevant outcome for participant i on their t^{th} day in the study on calendar date d. $\mathbf{T}_i$ is a vector of indicator variables capturing the treatment(s) that participant i was assigned to. β is the vector of coefficients, capturing the impact of each treatment on the outcome of interest.

---

[20]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. Results are robust to a variety of methods to address this imbalance, including baseline controls for this variable.

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 on which participants were absent, since attendance was balanced across groups. $\mathbf{X}_{it}$ includes controls for participants' age (quartiles) and gender and, where specified, 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 estimate the effect of naps separately compared to working and taking a break.[21] Finally, we include day-in-study and calendar-date fixed effects, captured by δ_t and λ_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:

$$y_i = \beta \mathbf{T}_i + \gamma_1 y_i^B + \gamma_2 \mathbf{X}_i + \varepsilon_i \tag{2}$$

where again the outcome variable only uses the observations from the treatment period and we control for the baseline observation of the outcome, y_i^B. The vector $\mathbf{X}_i$ includes the same gender and age controls. Since these outcome measurements span multiple days (e.g. present bias) or are on a fixed day-in-study (e.g. the endline survey), this specification does not include day-in-study or calendar-date fixed effects or control for whether non-nap participants worked or took a break.

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, by residualizing each constituent outcome with respect to day in study and calendar date, standardizing it by the control group's mean and standard deviation, and then taking 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 independent information. Signs of outcome variables are flipped when necessary so positive treatment effects imply an improvement in the outcome.[22] 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 multiple hypothesis testing issues caused by observing many outcomes. Our simplest approach is to examine 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 summary variable

---

[21]For some outcomes, $\mathbf{X}_{it}$ includes additional outcome-specific controls. For work-related outcomes, we 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, which was also randomized). Table notes specify these additional controls.

[22]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.

for each family, while correcting for the existence of multiple families. Finally, at the level of the individual outcomes, we correct for the existence of multiple outcomes *within* each family. In each case, we report adjusted p-values which control the family-wise error rate—the probability of at least one false rejection—using a step-down permutation procedure based on Westfall and Young (1993). The adjusted p-values are displayed for the main results in Tables IV, A.VII, A.VIII and A.IX.[23]

Pre-Analysis Plan (PAP). This study was pre-registered on both the AEA Registry and ClinicalTrials.gov, including a pre-analysis plan (PAP). We deviate from the PAP in some instances. 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 still present the pre-specified estimates in Tables II and III, but complement them with a simplified but higher-powered specification that pools the two night-sleep treatments and omits the interactions between nap and night-sleep treatments in Table IV. 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 deviations are detailed in Appendix Section D.

4.2 Impacts on Sleep

Overview. Figure V and Table II show the impacts of our treatments on the different measures of sleep. We find that the two night-sleep treatments quickly and substantially increase night sleep by 27 minutes per night on average. Offering short afternoon naps increases daytime sleep by 14 minutes per day. Thus, it is possible to substantially increase sleep in the sleep-deprived population we study through encouragement, incentives, and nap opportunities.

Night-sleep Treatments Only. Both night-sleep treatments sharply increased sleep, as measured by the actigraphs. On average, individuals who received the *Devices+Encouragement* and *Devices+Incentives* treatments *only*—that is, without also receiving naps—increased their time asleep at night by 19 and 29 minutes per night compared to the control group, respectively (Figure Va, Table II col 1). Including those who might have additionally received naps, the night-sleep treatments increased night sleep by 27 minutes per night on average (Table A.VI). This is a larger gain than 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 efficiency. Both night-sleep treatment groups increased their time in bed significantly throughout the treatment period—32 minutes for the 'encouragement only' group and 50 minutes for the 'incentives only' group (Figure Vb, Table II col 2).[24] We find no significant changes in sleep efficiency compared

---

[23]More details on our multiple comparisons corrections procedure are in Appendix E. For the main estimates in Table IV, we also provide a range of different corrections, including False Discovery Rate corrections in Table A.XVIII.

[24]On average, night-sleep treatment participants went to bed 17 minutes earlier at night and got out of bed 25

to the control group (Figure Vc, Table II col 3), even in the middle of the night when all participants are likely to be in bed (Appendix Figure A.Ia).

Increasing time spent in bed is a simple and practical way for our study participants to increase their sleep duration. However, improving sleep efficiency appears to be difficult for participants, even with the aid of the devices and tips surrounding sleep hygiene. Participants faced substantial (implicit) incentives for improvement. For example, a participant in the incentives group who improved their sleep efficiency from 70% to 80% would on average earn an additional Rs. 48—about 20% of average typing earnings—in sleep incentives each night (holding fixed eight hours per night in bed). Yet, we saw no changes in sleep efficiency. Improving sleep efficiency may require more substantial interventions , including ways to overcome barriers to sleep—such as mosquitoes, crowding, or psychological distress—that remained unaddressed by our treatments (Figure A.IVa).

Nap Treatment Only. The nap intervention was effective at increasing participants' daytime sleep (Figure Vd, Table II col 4). Nearly all participants in the nap treatment (92%) reported falling asleep during their allotted nap time, consistent with actigraph data showing 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.IId).

The "quality" of nap sleep in the office appears to be 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.[25] 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, Crowd-out, and Heterogeneity. We find only modest interactions between the sleep treatments in terms of their effects on the various sleep measures. The effect on 24-hour sleep of receiving both the encouragement treatment and the nap treatment together is very similar to the sum of the effects of receiving each treatment alone (25 minutes vs. 25 minutes, p=.98). The same is largely true for providing incentives and naps together (42 minutes vs. 35 minutes, p=.32).

We do find evidence that napping modestly crowds out night-time sleep: those treated with naps spent seven minutes less in bed at night and slept, on average, eight minutes less per night (Table II cols 1-2). In contrast, the night-sleep treatments do not interfere with naps. Participants randomized to the night-sleep treatments did not nap any less when offered a nap (Table II col 4). Both treatments increased total time asleep in 24 hours, although naps alone had a substantially smaller and insignificant impact (Table II col 5).[26]

---

minutes later in the morning (Table A.X).

[25]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.

[26]Table A.VI reports estimates from the model which pools the two night-sleep treatments and does not separate the cell which receives both a night-sleep treatment and the nap treatment. These estimates are more precise, and here the effect of naps on 24-hour sleep is significant. Table A.VI also reports Lee Bounds to address imperfect compliance with wearing actigraphs. The conclusions are largely unchanged, consistent with the fact that that non-compliance was relatively low (6%) and balanced across treatment groups.

Finally, the impact of the night-sleep treatments on sleep quantity and efficiency did not differ significantly by baseline sleep patterns, nor by characteristics such as participants' gender, age, or baseline earnings (Table A.V). Nor did these 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 Overall Effect of Each Treatment on Outcomes

Table III presents the treatment effects for each of the five combinations of night-sleep and nap treatments. Given the large number of outcomes and treatments, we focus on the effects on the overall summary index (col 1) which parsimoniously and efficiently averages our outcomes.[27] Each of the night-sleep treatments alone had no effect or a slightly negative (but insignificant) effect on participants: 0.00 SD (SE=0.07) and −0.05 SD (SE=0.07), respectively, for the 'encouragement only' and 'incentives only' groups. In contrast, participants in the 'nap only' treatment experienced positive and marginally significant effects of 0.11 SD (SE=0.07, p=0.10). The effect of 'naps only' differs significantly from that of 'incentives only' (p=0.02) and suggestively from that of 'encouragement only' (p=0.11). Those who received a night-sleep treatment in addition to naps had very similar overall effects to those who received naps only: 0.13 and 0.09 SD for the 'encouragement and nap' and 'incentives and nap' groups respectively, compared to 0.11 SD for 'nap only'.

These results provide evidence that naps have an overall positive effect on outcomes, while increases in night sleep do not. However, since each treatment cell has only around 75 participants, this analysis has limited statistical power. The combination of five treatment cells with numerous outcomes also makes discussion of detailed results unwieldy. We therefore turn to a simplified but higher-powered version of this analysis, which pools the two night-sleep treatments—which typically have similar and statistically indistinguishable effects —and does not include a separate indicator for the group that receives both the night-sleep and nap treatments. The resulting estimates in Table IV should thus be interpreted as weighted averages of treatment effects within the relevant cells. For instance, in this fully pooled specification, the coefficient on the night-sleep treatment 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 probability). In Appendix Table A.VIII, we instead pool the two night-sleep treatments but include a separate indicator for individuals who received a combination of either night-sleep treatment along with the nap treatment.

4.4 Impact of Night-Sleep Treatments

Overview. Experts from sleep science and economics predicted that increased night sleep would result in higher work output and labor supply, improved health and attention, increased financial savings, and reduced present bias (Figure IV). In contrast to these predictions and an influential literature in sleep science, we find no effect of the pooled night-sleep treatments on any of these

---

[27]Table A.VII provides multiple-hypothesis adjusted p-values for the individual outcomes in this specification. The overall index variable aggregates all outcomes, and thus does not require multiple-hypothesis testing corrections.

outcomes. More generally, we find no positive effects of the night-sleep treatments on the four family-level summary variables, or on *any* of the individual outcomes in our pooled specification (Figure VI and Table IV). Instead, increases in night sleep come at the cost of significantly reduced labor supply and therefore a marginally significant reduction in work output.

Work Outcomes. The night-sleep treatments did not cause significant improvements in productivity, labor supply, output, or earnings (Table IV cols 2-5). While the night-sleep treatment groups were 1.3% more productive than the control group (col 3), this difference of 0.02 SD (SE=0.02) 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 SE=0.02, col 4) with no change in days worked (Figure A.VII). This effect on labor supply remains significant with $p < 0.01$ 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.X). While perhaps 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 ($p<0.001$).

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, SE=0.02).[28] 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 results also contrast with those from natural experiments studying the economic consequences of sleep. Gibson and Shrader (2018) exploit variation in sunset times in the US 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 US 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 ($p<0.01$).

Well-being. Increased night sleep did not significantly improve physical or psychological well-being (Table IV cols 6-8). 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 shows 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, SE=0.06, Table IV col 8). The individual components of this measure also show no significant improvements (Table A.XI).

Similarly, we find no significant impacts of increased night sleep on an index of physical well-

---

[28]We also do not find any evidence that the effects of the night-sleep treatments become more positive as the length of treatment increases (Figure A.VIa).

being (col 7) which combines objective and self-reported measures of health status. We do find positive (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.XI). 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 (e.g. Strine and Chapman, 2005; Cappuccio et al., 2008; Giuntella and Mazzonna, 2019). It is also worth noting that the disaggregated analysis paints a slightly different picture when it comes to the impacts on physical (but not psychological) well-being. Table A.VIII reports that the 'night-sleep only' treatment has a positive effect on physical well-being (0.12 SD, $p=0.07$)—as does the 'nap only' treatment (0.16 SD, $p=0.03$)—when one allows the 'night-sleep and nap' cell to have a distinct effect. The interpretation is that the night-sleep treatments do increase physical well-being in a population which does not receive any naps, but have no average impact on physical well-being in a population where half the people already receive naps. Night-sleep and naps thus appear to have a negative interaction effect on physical well-being.

Cognition. We find no impacts of increased night sleep on cognition (Table IV cols 9-11). There is no significant effect on an index of laboratory measures of attention, memory, and inhibitory control (Table IV col 10), or on any of the individual outcomes, which closely mimic measures used in a laboratory studies (Table A.XII). A large body of sleep studies shows that inducing sleep deprivation—typically by keeping participants up all night—substantially worsens performance on these tasks (Lim and Dinges, 2010; Killgore, 2010). The more modest but sustained and policy-relevant increases in sleep in our setting 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 IV col 11). Consistent with limited attention, participants in the control group reacted 16% *less* to high incentives when piece-rates were *non*-salient (Table A.XIII). Participants in the night-sleep treatments behave quite similarly, reacting 15% less to incentives when they were non-salient. Increased night sleep did thus not close the gap between responses to salient and non-salient incentives.

Preferences. Consistent with the lack of positive impacts of increased night sleep on the above outcomes, we find no evidence of the night-sleep treatments affecting time, risk, or social preferences, or on an index combining these outcomes (Table IV cols 12-15). We detect no significant effect on an index of two measures of time preferences: savings and present bias (Table IV col 13). The night-sleep treatments did not meaningfully affect savings behavior, leaving deposits and accumulated interest unaffected (Table A.XIV, Panel A). Similarly, we find no impacts on present bias estimated from effort choices. We do detect present bias in the control group ($\beta=0.92$, Table A.XIV Panel B). However, increased sleep does not significantly shift this parameter, in contrast to the view that sleep replenishes self-control (Vohs and Baumeister, 2016).

We also find no evidence of altered risk aversion, loss aversion, or social preferences in standard experimental tasks (Table IV col 14-15), 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 can rule out changes greater than 0.16 SD at the 95%

level for each of these outcomes (Table A.XV).

Heterogeneity. We do not find significant evidence of heterogeneity in the effects of the night-sleep treatments. Table A.XVI considers effects on the overall index variable, and tests for heterogeneity by a number of baseline covariates. Baseline night-sleep duration, sleep efficiency and propensity to nap prior to the study do not interact significantly with the night-sleep treatments. Nor do demographics such as gender and age.[29]

4.5 Impact of Naps

Overview. Naps improved outcomes across a range of domains (Figure VI). Table IV col 1 reports that naps had an economically meaningful and statistically significant effect on the overall summary index (0.12 SD, SE=0.04, $p<0.01$). As shown in Table A.IX, the effect is significant whether naps are compared to taking enforced breaks (0.15 SD, $p<0.01$) or to working through the afternoon (0.08 SD, $p=0.03$). Given the lack of evidence on the impacts of naps on economically meaningful outcomes in real-world settings (Lovato and Lack, 2010; Ficca et al., 2010), 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.

Work Outcomes. Naps increased productivity. Participants randomized to naps were 2.3% (0.04 SD, SE=0.02, $p=0.06$) more productive on average over the day (Table IV col 3).[30] This effect is sizable, given that productivity is quite inelastic: quadrupling the piece rate increased productivity by only 14%. The productivity effects of naps are similar when compared to the break or the work counterfactuals, suggesting that the impacts are due to sleep rather than merely resting (Table A.IX). Figure A.VIII shows that the effects are larger in the afternoon (2.7%) but also evident in the morning (1.9%), suggesting either cumulative effects of regular napping or that participants work harder in the morning in anticipation of the nap.[31]

By design, nap participants could not work during the a 30-minute nap period. They could adjust their labor supply outside of this period, say by staying at the office longer. We find no evidence of such adjustments (Table IV col 4): The nap group spends almost exactly as much time typing as those in the control group on their 'break days', and 26 minutes (0.20 SD) fewer than the control group on its 'work days' (Table A.IX).

The impacts of naps on output and earnings depend on the comparison group (Table A.IX). Compared to taking a break, naps increased total output by 0.05 SD ($p=0.02$). Comparing to working, naps instead *reduce* output by 0.07 SD ($p<0.01$). Earnings closely track output: naps increased overall earnings by Rs. 11 per day (0.05 SD, $p = 0.07$) compared to taking a break, a sizable increase of 4.1%.

---

[29]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 duration or sleep efficiency typically observed in high-income countries.

[30]Table A.VIII reveals that the estimated positive effect of naps on productivity in Table IV is driven by those individuals who also receive the night-sleep treatments. Naps alone, instead, have no effect on productivity.

[31]The unadjusted p-values of effects in the afternoon and morning are 0.01 and 0.05, respectively. 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).

In contrast, taking time to nap lowered earnings by Rs. 23 (8.3% or 0.10 SD, $p<0.01$) compared to simply working through the break.[32]

Well-being. Naps significantly improved well-being (Table IV cols 6-8). The overall effect of 0.08 SD ($p=0.02$) is driven by a 0.12 SD ($p=0.04$) improvement in psychological well-being. The point estimates for most individual components of psychological well-being are positive, with significant effects on happiness, life satisfaction, and sense of life possibility (Table A.XI). The lack of significant impacts on physical well-being is perhaps unsurprising, given the limited impacts of naps on total sleep. Moreover, physical health benefits of sleep may require more time to emerge and some of the outcomes, such as the cycling task, were conducted at the end of the study on a day without naps.[33]

Cognition. Naps boost the cognition index—which combines lab measures of cognition with a measure of attention at work—by 0.10 SD ($p=0.08$; Table IV col 9). The lab measures of cognition increase by 0.08 SD ($p=0.06$; col 10) on average, driven by an effect on attention (Table A.XII). We find no significant impacts on inhibitory control or memory, 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). We also find that naps increased participants' attention to work incentives by 0.20 SD ($p=0.07$) (Table IV col 11). The nap group was nearly fully attentive to non-salient incentives, reacting to them about as much as they reacted to salient incentives (Table A.XIII). This result illustrates the improved attentional resources provided by naps in a real-world work environment.

Preferences. Naps have a positive and marginally significant effect (0.07 SD, $p=0.10$) on the index corresponding to the preferences family (Table IV col 12). Within this family, we find suggestive evidence that naps increase an index of patience by 0.13 SD ($p=0.11$, Table IV col 13). This index combines two real-stakes measures of time preferences: savings at the study office and present bias in an effort-discounting task. Naps caused a 14% increase in deposits and a 13% increase in daily net savings (deposits minus withdrawals, Table A.XIV Panel A). These effects are sizable: randomly providing a 1 percentage point higher *daily* interest rate increased deposits by 31%. The nap group earned 19% more interest over the course of the study, although this effect is imprecisely estimated.[34]

Naps also reduce present bias in a real-effort task (Table A.XIV, Panel B). We estimate an average present bias parameter of $\beta=0.92$ in the control group.[35] The nap treatment significantly (unadjusted $p<0.05$) 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 or risk preferences (Table IV, cols 14-15). The nap group is not more willing to accept risk or probabilistic losses (Table A.XV). The same table reports

---

[32]The negative effect of naps on earnings (compared to work days) appears to diminish over time. On "long" days, when participants are not restricted to artificially short work hours, the earnings of the nap group converge with the earnings of participants on work days by the end of the study (see Figure A.VI).

[33]Recall, however, that 'naps only' had a positive effect on physical well-being (0.16 SD, $p=0.01$) when we allow the 'night-sleep and nap' cell to have a separate effect, as reported in Table A.VIII.

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

[35]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.XIX).

that nap participants send 0.16 SD more (SE=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. It is worth noting, however, that the measures of social and risk preferences were elicited in the morning, before daily naps, while time preferences were instead measured in the afternoon (Table A.III).

Heterogeneity. As with the night-sleep treatments, we do not detect statistically significant heterogeneity in the overall effects of naps by baseline sleep or demographic characteristics (Table A.XVI). Notably, those who self-reported napping in the week prior to the study did not see higher effects from the nap treatment. This partly alleviates concerns that the treatment effects capture mainly the effect of reducing naps among the control participants rather than increasing naps in the treatment group.

4.6 Nap vs. Night-sleep Treatments

Table IV reports p-values for differences in the effects of the night-sleep and nap treatments on each outcome. We find clear evidence that naps and night-sleep treatments have significantly different effects on the *overall* index variable which combines all outcomes ($p=0.02$). We are relatively underpowered to instead compare the effects on individual outcomes. The only individual outcome on which naps have a significantly different effect than the night-sleep treatments is psychological well-being ($p=0.04$). The differences in effects on the indices of well-being, cognition and preferences are not significant, although the point estimates are larger for naps in each case. For earnings, the point estimates for naps and night-sleep treatments are similar and not significantly different.

The nap treatment has larger overall effects than the night-sleep treatment, despite causing smaller increases in total (24-hour) sleep. Figure A.IX plots the overall treatment effects for each of the five disaggregated treatments against the increase in 24-hour sleep generated by the treatment. The figure shows no evidence of a dose response to 24-hour sleep or to night sleep: treatments with larger effects on overall sleep do *not* have larger effects on the summary of downstream outcomes. In contrast, treatment combinations involving naps have distinctly larger effects than the ones without naps. Naps at the office and nighttime sleep at home are not one-for-one substitutes in our setting, in contrast to some evidence from sleep-lab experiments (Mollicone et al., 2007).

To formally test for differences in the effects of nighttime sleep and nap sleep per minute of sleep, we turn to an instrumental variable (IV) analysis. Specifically, we estimate an IV regression with duration of nighttime sleep and nap sleep as endogenous variables and the five different combinations of treatments as instruments (Table A.XVII).[36] As expected given the results above, each additional minute of nap sleep has a substantial impact on the overall index (0.008 SD per minute, or 0.47 SD per hour), significantly different from a minute of nighttime sleep ($p<0.01$). However, the per-minute benefits of naps and nighttime sleep also differ significantly for the well-being ($p=0.08$) and cognition ($p=0.05$) indices, and suggestively for the preferences index ($p=0.11$).[37] The per-minute effects are

---

[36]This analysis makes strong assumptions about the linearity of effects (per minute of sleep) and involves exclusion restrictions. While some of these assumptions might be questionable, we believe Table A.XVII provides valuable information on the differences in effects of the two types of sleep.

[37]The permutation-based Westfall-Young procedure we use to control the family-wise error rate in our other tables

also statistically different—more positive for nap sleep—for psychological well-being (p=0.05) and the lab measures of cognition (p=0.08). However, naps also lead to larger *reductions* in labor supply per minute of sleep than nighttime sleep (p=0.04). Understandably, a minute of daytime sleep crowds out more work than a minute of nighttime sleep.

5 Discussion and Conclusion

Using state-of-the-art objective measurements, we document a novel fact about sleep in India: low-income adults in Chennai are severely sleep-deprived by usual standards. The strikingly low sleep duration and efficiency in our sample could be a widespread but underappreciated feature of the lives of the urban poor in developing countries. Despite the lack of short-run benefits from sleeping more in our setting, severe sleep deprivation may have serious long-run effects on health and well-being, and calls for policymakers' and researchers' attention. More systematic research on sleep in developing countries is needed to establish the external validity of our findings and to study longer-term effects.

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 our outcomes, contrasting with predictions made by sleep scientists and economists and evidence from many sleep lab experiments (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 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 instead study the impacts of a more modest and arguably policy-relevant *increase* in sleep on highly incentivized outcomes. As one of the first studies of 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

---

cannot be applied to the IV setting, since the permuted first-stage leads to a weak instruments problem. Therefore, we report adjusted p-values calculated using Hochberg's step-up procedure for controlling the FWER for this table. For the main intent-to-treat results in Table IV, the Hochberg and Westfall-Young adusted p-values are essentially identical, as reported in Table A.XVIII.

perspective on sleep, which considers sleep as a choice variable, and measures both the benefits and the opportunity costs of sleeping more.

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. 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 afternoon work may be less productive (Dinges, 1992). Yet we know very little about the economic costs and benefits of naps in real-world settings, as concluded by Ficca et al. (2010) in a recent review. As work in developing countries shifts from self-employment to working in firms with set hours, naps may be crowded out. Our findings suggest that this could be costly, since naps provide workers with benefits that night sleep does not. Large firms may be beginning to recognize this, as evidenced by the adoption of nap policies by employers such as Google and Nike (Walker (2017), chapter 15). Given the foregone work time and logistical costs of offering naps, an employer's decision to provide naps may depend on how much they value their workers' psychological well-being and attention in the particular work setting.

An obvious question raised by our results is why naps were effective when a larger increase in night sleep had no effect. Naps and night-sleep are clearly not substitutes in our setting (Figure A.IX and Table A.XVII). 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).

Perhaps the broadest question our research raises is to what extent the sizable impacts of sleep in lab studies generalize to field settings. Measuring sleep in the field—made possible by wearables—allows such experiments linking sleep to real-world economic outcomes in a wider range of environments (Handel and Kolstad, 2017). 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.

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.

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.

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.

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.

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.

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

Handel, Benjamin and Jonathan Kolstad, "Wearable technologies and health behaviors: new data and new methods to understand population health," *American Economic Review*, 2017, *107* (5), 481–85.

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.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-Control at Work," *Journal of Political Economy*, 2015, *123* (6), 1227–1277.

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

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

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

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

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

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

Lee, David S, "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *The Review of Economic Studies*, 2009, *76* (3), 1071–1102.

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.

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.

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.

Westfall, Peter H and S Stanley Young, *Resampling-based multiple testing: Examples and methods for p-value adjustment*, Vol. 279, John Wiley & Sons, 1993.

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.

6 Figures and Tables

FIGURE I: Typical Sleep in Chennai

(a) A Typical Night

(b) A Good Night
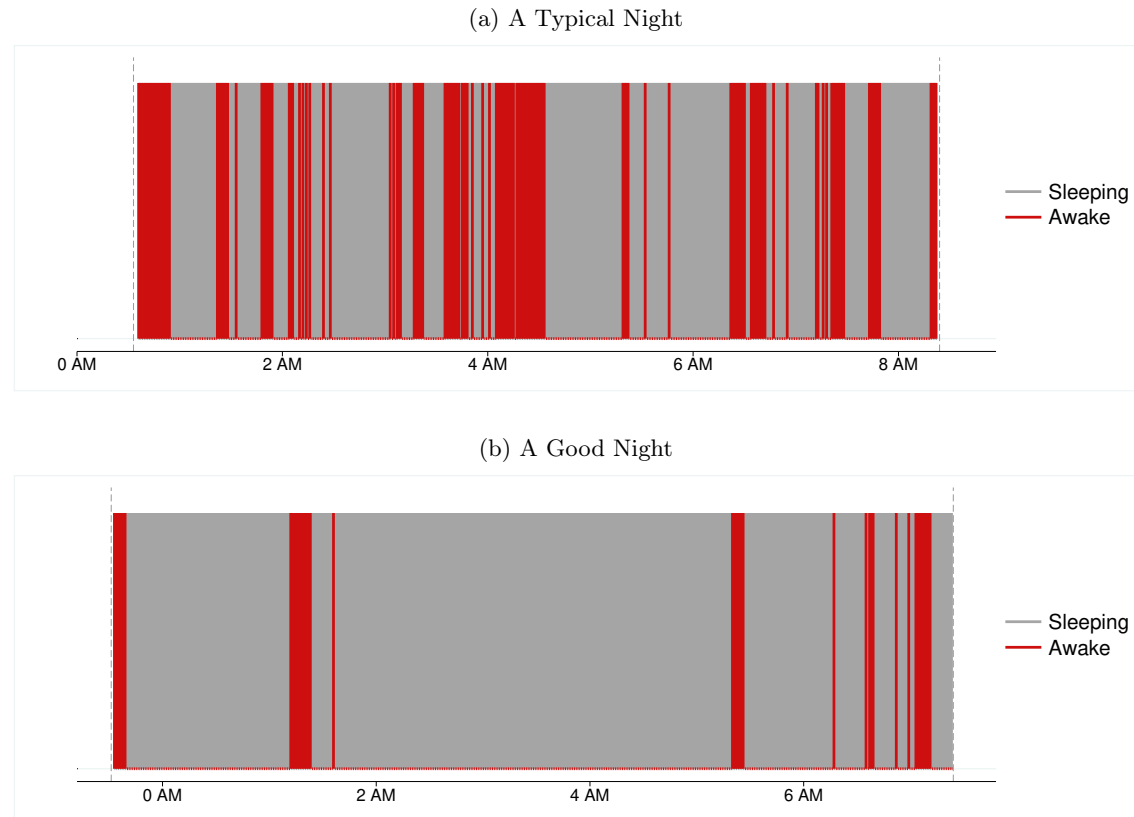
Notes: This figure represents actigraph-measured sleep-wake patterns of two particular nights of two selected study participants. Gray areas indicate one-minute periods in which the participant was asleep and red areas indicate periods in which the participant was awake according to the actigraph. The light gray dashed lines indicate when the participant got into or out of bed.

- In panel (a) we show 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 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 (RCT Sample)
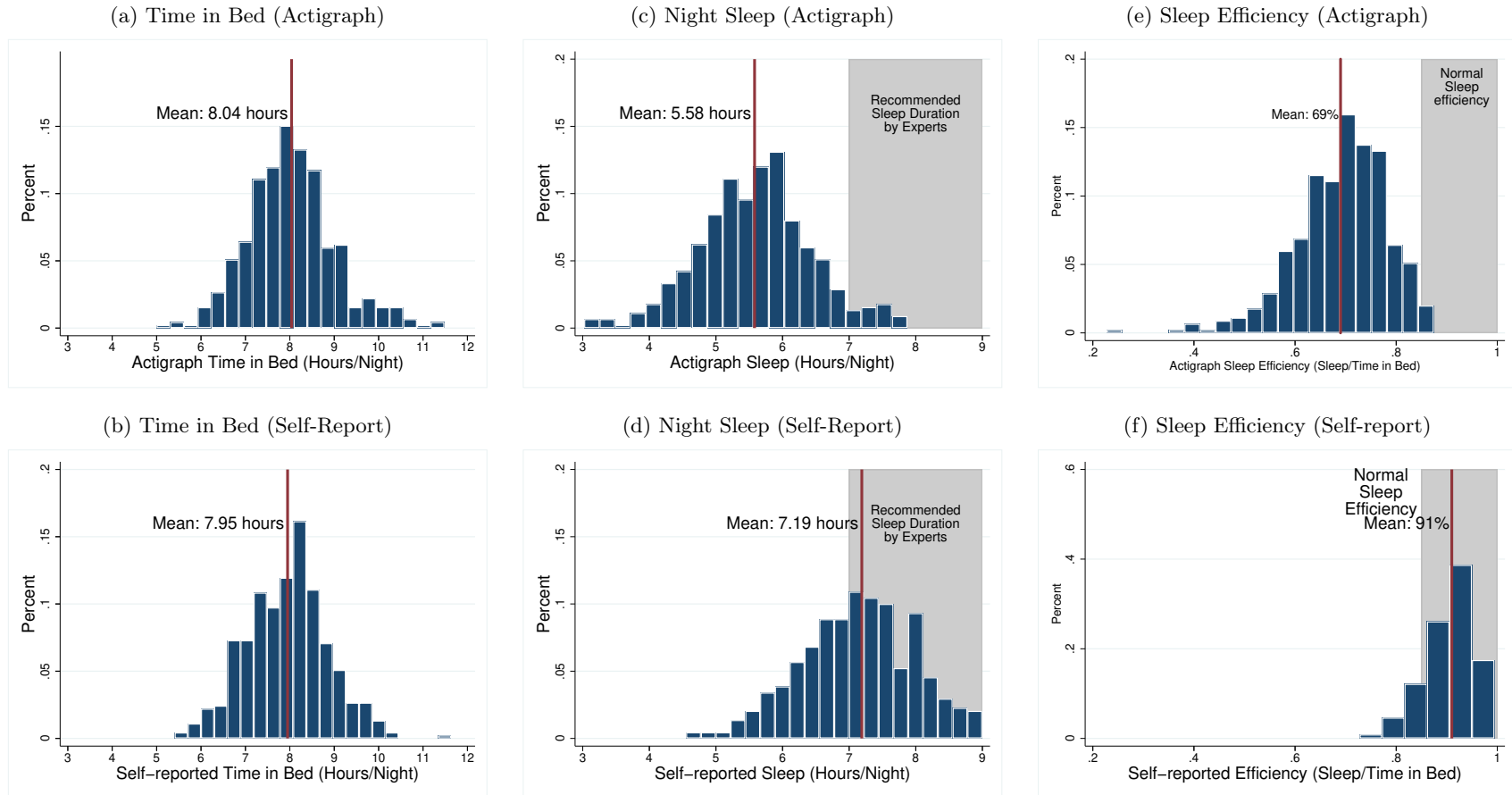
Notes: This figure shows the distribution of the sleep-related variables averaged at the participant-level over the baseline period (7 nights) in the RCT sample (N = 452). The upper three panels show distributions of actigraph-measured sleep patterns while the lower three panels show the corresponding distributions based on self-reports.

- 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: RCT Timeline and Experimental Design
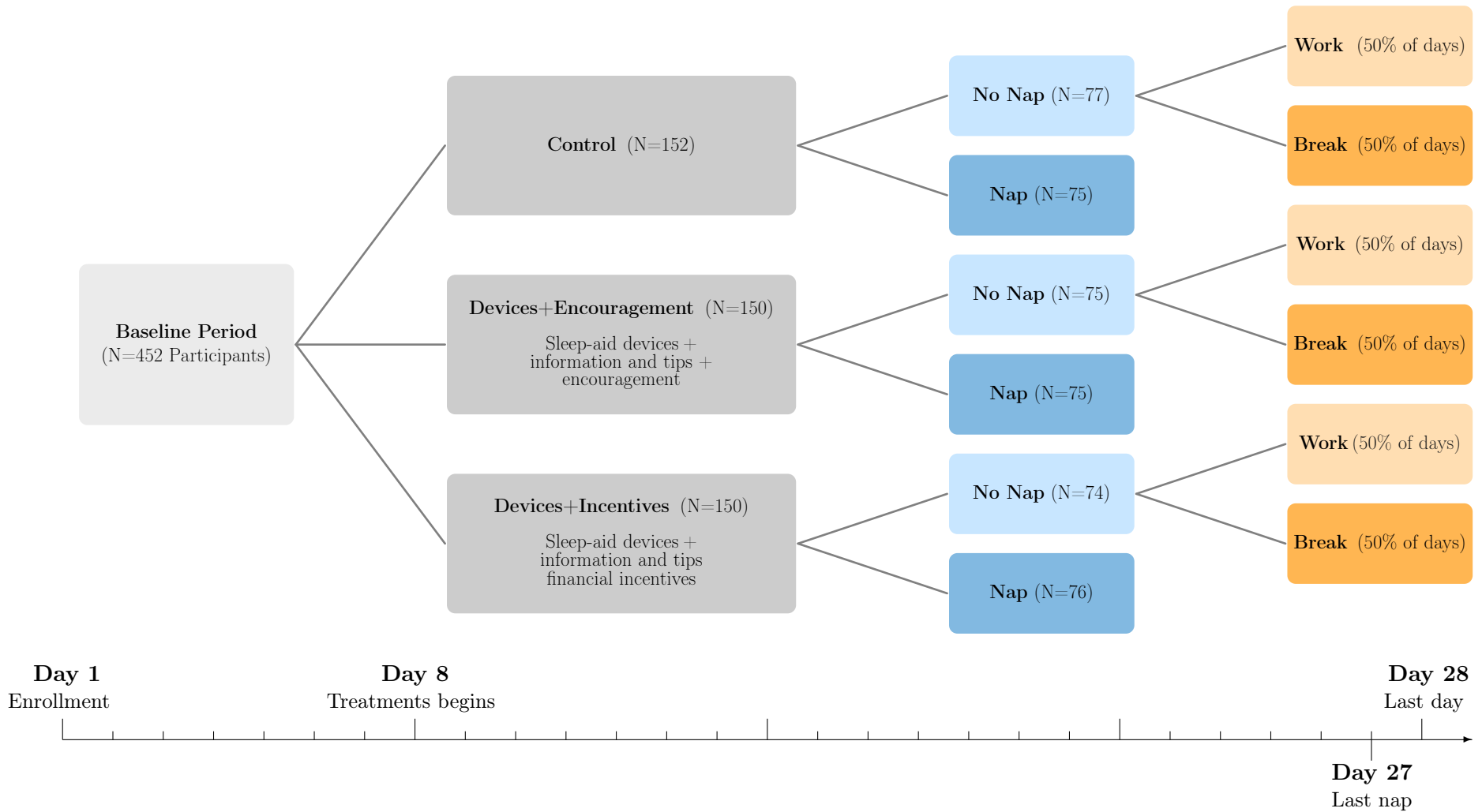
Notes: This figure presents an overview of the timeline and experimental design of the study.

- After 8 baseline days, the 452 participants were divided into three night-sleep treatment groups: Control, Sleep Devices+Encouragement, and Sleep Devices+Incentives.

- Participants in each of these groups were further randomized into 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.

- Participants in the No Nap group were further randomized on a daily level (i.e. within individual) either to being allowed to work during the nap period or to take a mandatory pause during that time, with equal probability.

- 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: Expert Predictions vs. RCT Results
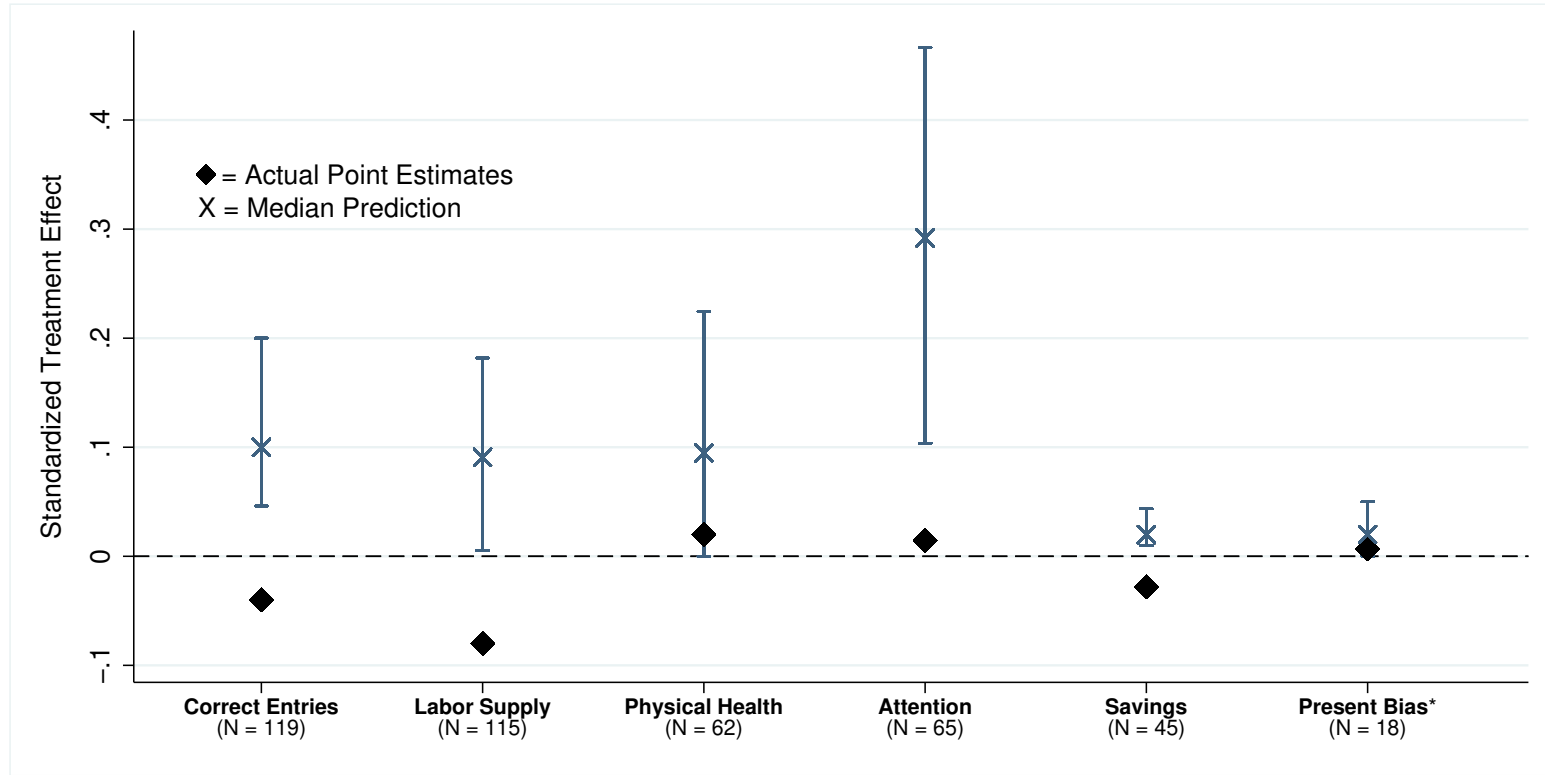
Notes: This figure summarizes the predictions made by experts in economics and sleep science about the expected treatment effect of our (pooled) night-sleep interventions. We normalize each prediction, dividing them by by the control group's standard deviation. For comparison, we also present the actual estimated treatment effect.

- The bars show the inter-quartile range (25th to 75th percentiles) of the predictions for a given outcome variable. We also show the median prediction (X) and the actual point estimates (diamond) of the treatment effects. N refers to the number of responses for each outcome variable. This number varies by outcome because different types of experts (e.g. sleep researchers, behavioral economists) were asked about some different outcomes.

- *Correct Entries* refers to the number of daily correct characters in the data-entry task, i.e. a measure of overall output per day.

- *Labor Supply* refers to the daily number of hours working in the typing task, i.e. time at the office excluding voluntary and scheduled pauses.

- *Physical Health* refers to a variable that averages (normalized) systolic and diastolic blood pressure. We flip the sign of this variable so a positive value means an improvement in health (i.e., a reduction in blood pressure).

- *Attention* refers to an index pooling inverse response times and minor lapses in the Psychomotor Vigilance Task (PVT), a standard lab measure of attention in the sleep literature. Signs are flipped such that higher values refer to increased attention.

- *Savings* refers to the daily amount deposited minus the amount withdrawn in the savings box during the experiment.

- *Present Bias* refers to the present-bias parameter β. Unlike the other variables, the predictions and point-estimate refer to the reduction in present bias (increase in β) rather than a normalized outcome, for ease of interpretation.

FIGURE V: Impacts on Nighttime and Nap Sleep

(a) **Night Sleep**

(b) **Time in Bed**

(c) **Sleep Efficiency**

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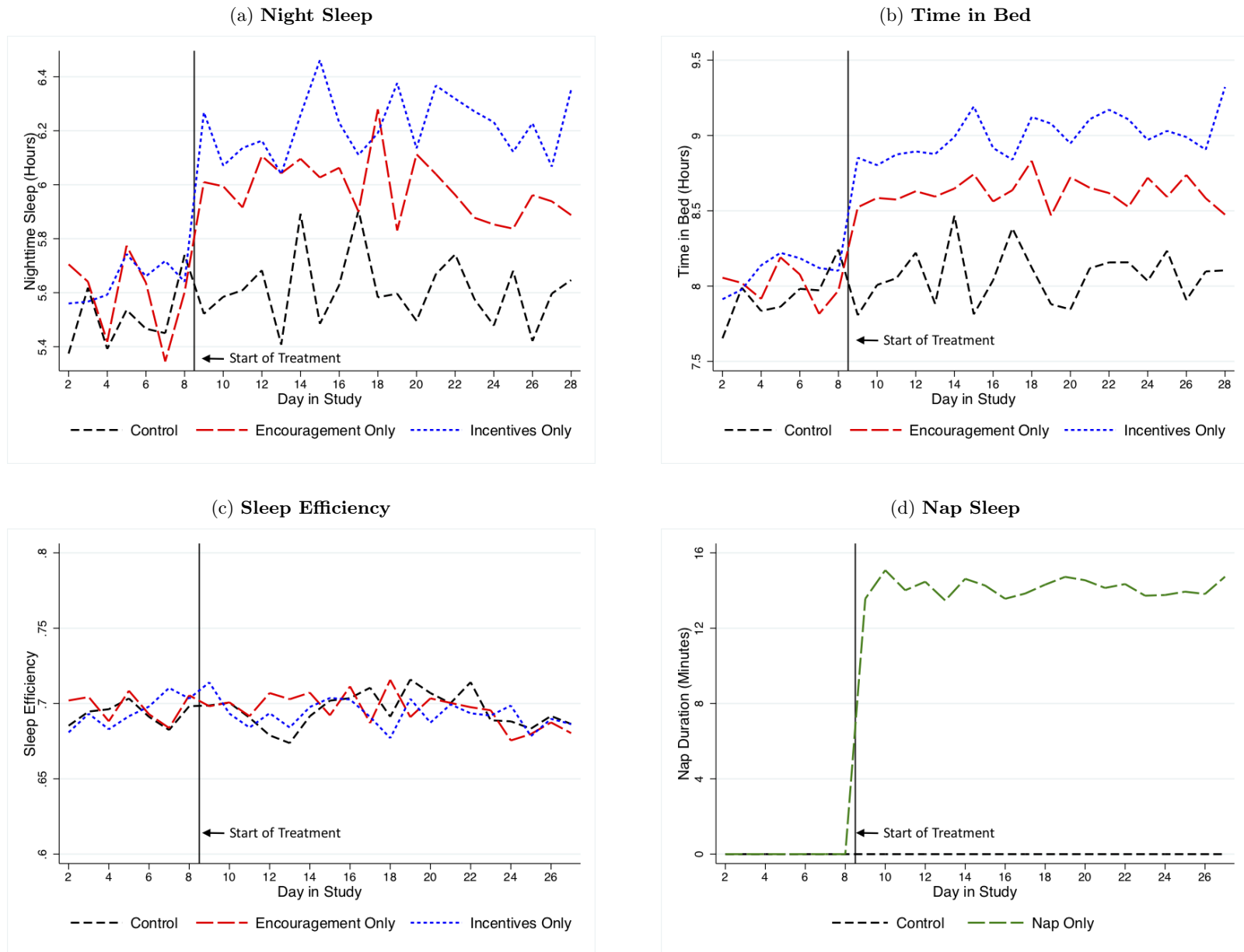
Notes: This figure shows the average of different sleep-related variables for different treatment arms by day in study of the RCT. All outcomes are actigraph measures. We only include workday nights and days in the sample. In panels (a) and (b), we plot hours of night sleep and hours in bed at night,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

FIGURE VI: Summary of Treatment Effects
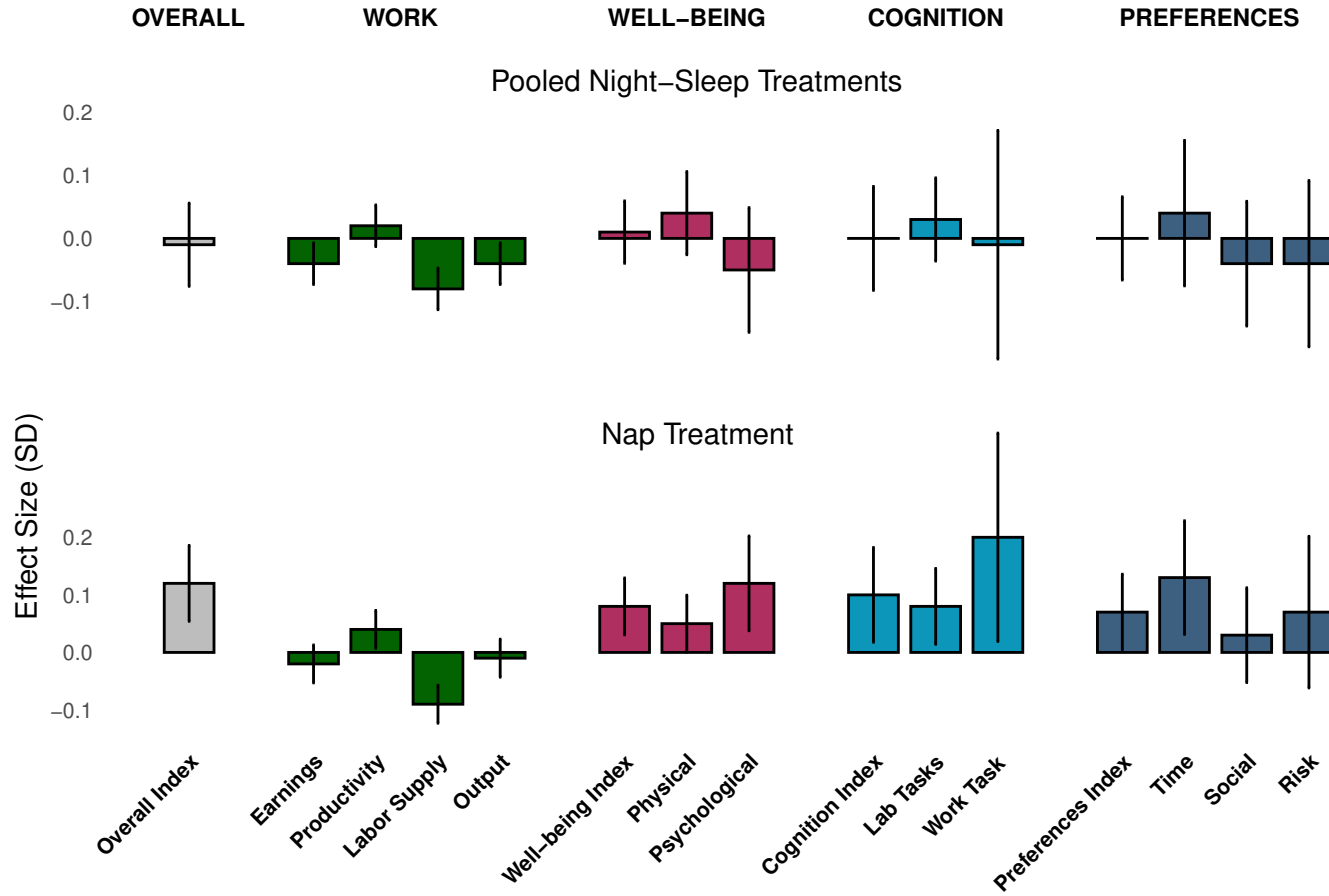
Notes: This figure summarizes the treatment effects in our study. We plot the point estimates and 90% confidence intervals for the pooled night-sleep interventions in Panel (a) and the nap intervention in Panel (b). All outcomes are standardized, i.e. we subtract the mean and divide by the standard deviation 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 Table 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; (ii) productivity; (iii) active typing time; and (iv) output from the data-entry task.
- **Well-being:** (i) "Well-being Index" - a composite index of the physical and mental well-being indices; (ii) "Physical" - a 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; and (iii) "Psychological" - a mental well-being index, a composite of self-reported depression, happiness, life possibility, life satisfaction, and stress.
- **Cognition:** (i) "Cognition Index" - composite index of a lab-based and a work-based measure of cognitive function; (ii) "Lab Tasks" - index of lab measures of attention, memory, and inhibitory control; and (iii) "Work Task" - measure of attention to piece rates in the data-entry task
- **Preferences:** (i) "Preferences Index" - composite index of time, social, and risk preference indices; (ii) "Time" - index capturing time preferences, including savings and present bias; (iii) "Social" - index representing social preferences; and (iv) "Risk" - index representing risk preferences.

Table I: Sleep Statistics in Two Samples in Chennai

	RCT Sample (pre-treatment)	Broader Sample
	(1)	(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)	27.4 (10.14)
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 average sleep characteristics in two samples in Chennai. Standard deviations are in parentheses.

- Column 1 presents summary statistics from the RCT sample, only using data from the 7 nights in the baseline period (i.e. before any treatments were implemented). 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 outcomes are objectively measured by actigraphs unless indicated otherwise. All means and standard deviations are at the participant-level (i.e. we collapse the data at the participant level by averaging across nights) unless indicated otherwise.

- The variables shown in the table are: (i) hours in bed (per night, regardless of whether awake or asleep); (ii) hours asleep at night (per 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 (in hours) 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)
Devices+Encouragement Only	0.32***	0.53***	-0.54	0.00	0.33***
	(0.08)	(0.09)	(0.66)	(0.00)	(0.08)
Devices+Incentives Only	0.49***	0.84***	-1.14	0.00	0.49***
	(0.08)	(0.09)	(0.70)	(0.00)	(0.08)
Nap Only	-0.13*	-0.11	-0.72	0.25***	0.09
	(0.08)	(0.09)	(0.72)	(0.01)	(0.08)
Devices+Encouragement and Nap	0.21***	0.40***	-1.06	0.24***	0.42***
	(0.07)	(0.08)	(0.66)	(0.01)	(0.08)
Devices+Incentives and Nap	0.48***	0.58***	0.90	0.24***	0.70***
	(0.08)	(0.09)	(0.68)	(0.01)	(0.08)
Control Mean	5.61	8.07	69.86	0.00	5.61
Control SD	1.20	1.37	11.28	0.00	1.20
Participant-nights	8454	8454	8454	7191	8035
Participants	451	451	451	450	451

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

- Each column shows the OLS estimates of equation (1) including dummies for each treatment arm and controlling for the average baseline measure of the dependent variable (ANCOVA), age, sex, and day-in-study and date fixed effects.

- Each row shows coefficients of treatments cells compared to the control group that receives no sleep-related treatments.

- Night sleep, time in bed, nap sleep and 24-hour sleep (cols 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 in the office and night sleep.

- Standard errors are clustered at the participant level. Stars next to coefficients reflect unadjusted p-values (* significant at 10%; ** at 5%; *** at 1%).

Table III: Fully-Disaggregated Treatment Effects of Night-Sleep and Nap Treatments

	OVERALL	WORK				WELL-BEING		
	Index (1)	Earnings (2)	Productivity (3)	Labor Supply (4)	Output (5)	Index (6)	Physical (7)	Psychological (8)
Devices+Encouragement Only	0.00 (0.07)	-0.07** (0.03)	-0.02 (0.03)	-0.06* (0.04)	-0.07** (0.03)	0.13*** (0.05)	0.16** (0.06)	0.02 (0.09)
Devices+Incentives Only	-0.05 (0.07)	-0.07* (0.04)	-0.02 (0.03)	-0.08** (0.04)	-0.08** (0.03)	0.05 (0.05)	0.08 (0.06)	-0.02 (0.09)
Nap Only	0.11* (0.07)	-0.07 (0.04)	-0.01 (0.03)	-0.07* (0.04)	-0.05 (0.04)	0.18*** (0.06)	0.16** (0.07)	0.19** (0.10)
Devices+Encouragement and Nap	0.13* (0.07)	-0.07* (0.04)	0.04 (0.03)	-0.15*** (0.04)	-0.04 (0.03)	0.13*** (0.05)	0.12* (0.06)	0.11 (0.09)
Devices+Incentives and Nap	0.09 (0.06)	-0.07* (0.04)	0.04 (0.03)	-0.17*** (0.04)	-0.07** (0.03)	0.11** (0.05)	0.12* (0.06)	0.07 (0.09)
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)
Devices+Encouragement Only	-0.00 (0.07)	-0.00 (0.07)	0.04 (0.17)	-0.09 (0.06)	-0.05 (0.12)	-0.10 (0.09)	-0.15 (0.13)
Devices+Incentives Only	-0.03 (0.08)	0.04 (0.07)	-0.12 (0.19)	-0.04 (0.07)	0.11 (0.12)	-0.16* (0.08)	-0.06 (0.13)
Nap Only	0.09 (0.07)	0.07 (0.07)	0.15 (0.17)	-0.01 (0.07)	0.12 (0.12)	-0.09 (0.09)	-0.01 (0.14)
Devices+Encouragement and Nap	0.05 (0.08)	0.13* (0.07)	0.03 (0.18)	0.09 (0.07)	0.23* (0.12)	0.03 (0.10)	0.03 (0.13)
Devices+Incentives and Nap	0.14** (0.07)	0.09 (0.07)	0.32** (0.16)	0.01 (0.07)	0.11 (0.12)	-0.10 (0.09)	0.01 (0.14)
Participants	452	452	429	452	452	415	415

Notes: This table shows the treatment effects of the five fully-disaggregated treatment arms on the overall index as well as the four families of outcomes: work, well-being, cognition, and preferences. Each row shows coefficients of treatments cells compared to the control group that receives no sleep-related treatments. All work-related regressions are conducted at the participant-day level using equation (1). All other regressions are at the participant level using equation (2).

- The overall index (col 1) aggregates across the four family-level outcomes. The work outcomes include data-entry earnings (the summary variable for the work family, col 2); productivity (output per hour typing, col 3); active typing time (col 4); and output (col 5). Well-being outcomes include an overall index (col 6) of the two broad measures of well-being: a physical well-being index (col 7) and a mental well-being index (col 8), as described in section 3.2. Cognition measures include an overall index (col 9) of two measures: lab measurements of attentiveness, memory, and inhibitory control (col 10); and attention to piece rates in the data-entry task (col 11). Preference measures include an index (col 12) of three different categories: time preferences (savings and present bias, col 13), social preferences (col 14), and risk preferences (col 15).

- All indices are a weighted average of their components, in which the weights take into account the covariance structure of the components (Anderson, 2008). All dependent variables are normalized with respect to the control group's mean and standard deviation. When required, outcomes are flipped so that a positive value aligns with what would be considered a "better" outcome.

- Standard errors in parentheses are robust to heteroscedasticity and clustered at the participant-level when applicable. Stars next to coefficients reflect unadjusted p-values (* significant at 10%; ** at 5%; *** at 1%). Table A.VII shows p-values that take into account multiple-hypothesis corrections.

## Table IV: Pooled Treatment Effects of Night-Sleep and Nap Treatments

	OVERALL	WORK				WELL-BEING		
	Index (1)	**Earnings** (2)	Productivity (3)	Labor Supply (4)	Output (5)	**Index** (6)	Physical (7)	Psychological (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)
	{0.79}	{0.08}	{0.30}	{0.00}	{0.05}	{0.66}	{0.34}	{0.36}
		[0.30]	[0.32]	[0.00]	[0.09]	[0.96]	[0.36]	[0.36]
Nap Treatment	**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)	(0.03)	(0.05)
	{0.00}	{0.23}	{0.03}	{0.00}	{0.77}	{0.01}	{0.12}	{0.02}
		[0.24]	[0.06]	[0.00]	[0.78]	[0.02]	[0.14]	[0.04]
Participants	451	451	451	451	451	452	452	452
Unadjusted p-value NS vs. Nap	{0.02}	{0.62}	{0.49}	{0.65}	{0.23}	{0.14}	{0.74}	{0.03}
FWER-corrected p-value NS vs. Nap		[0.63]	[0.65]	[0.65]	[0.46]	[0.30]	[0.76]	[0.04]

	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)
	{0.93}	{0.51}	{0.95}	{0.95}	{0.58}	{0.45}	{0.65}
	[0.96]	[0.76]	[0.95]	[0.96]	[0.65]	[0.65]	[0.65]
Nap Treatment	**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)
	{0.03}	{0.03}	{0.07}	{0.05}	{0.04}	{0.52}	{0.33}
	[0.08]	[0.06]	[0.07]	[0.10]	[0.11]	[0.51]	[0.51]
Participants	452	452	429	452	452	415	415
Unadjusted p-value NS vs. Nap	{0.12}	{0.31}	{0.21}	{0.15}	{0.30}	{0.30}	{0.32}
FWER-corrected p-value NS vs. Nap	[0.30]	[0.30]	[0.30]	[0.30]	[0.33]	[0.33]	[0.33]

Notes: This table shows treatment effects of the pooled night-sleep and the nap interventions on the overall index as well as the four families of outcomes: work, well-being, cognition, and preferences. Each row shows coefficients of treatments cells compared to the control group that receives no sleep-related treatments. All work-related regressions are conducted at the participant-day level using equation (1). All other regressions are at the participant level using equation (2). All outcome variables are the same as in Table III.

- In contrast to Table III, in this table: (i) we pool the Devices and Encouragement and the Devices and Incentives treatments to estimate the pooled effect of the two night-sleep treatments, and (ii) we do *not* include a separate indicator for the group that receives both the night-sleep and nap treatments. Hence, these estimates should be interpreted as weighted averages of treatment effects within the relevant cells. For instance, the coefficient on the night-sleep treatment 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 probability).

- Below each coefficient, we report (i) the corresponding standard errors in parentheses (·), robust to heteroscedasticity and clustered at the participant-level when applicable, (ii) the unadjusted p-value in curly brackets {·}, and (iii) the Westfall-Young FWER-adjusted p-value in square brackets [·], as described in more detail in Appendix E.

A Online Only Supplementary Figures

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

(a) **Fraction Asleep by Hour of Night**
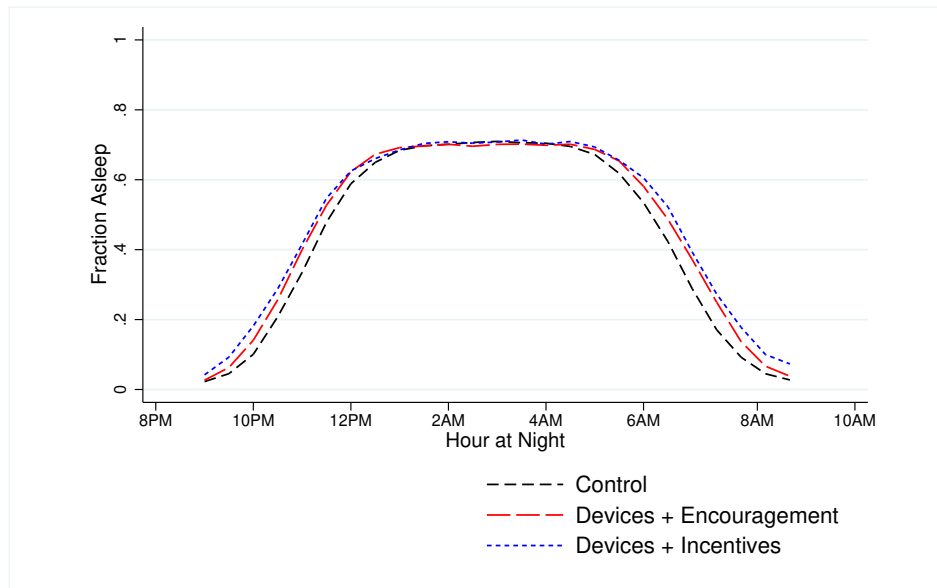
(b) **Fraction in Bed by Hour of Night**
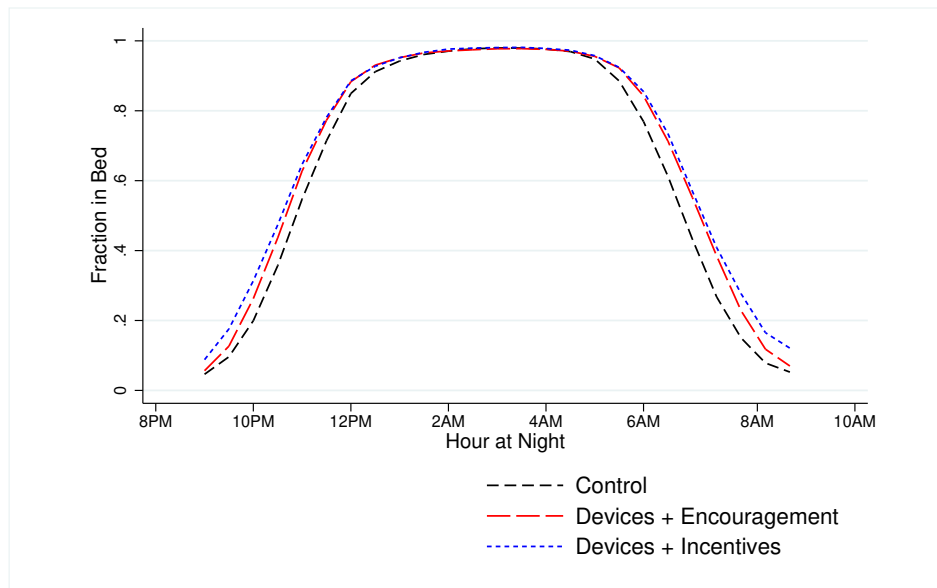
Notes: This figure shows the average fraction of participants asleep and in bed over the course of the night, as measured by the actigraphs. The figures use data from the 19 nights of the treatment period.

- Panel (a) shows the fraction of participants in the two Night-sleep Treatment Groups and in the corresponding Control Group that was asleep at any time during the night, as measured by the actigraph.
- Panel (b) shows the fraction of participants in the two Night-sleep Treatment Groups and in the corresponding Control 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**

(d) **Distribution of Nap Duration**
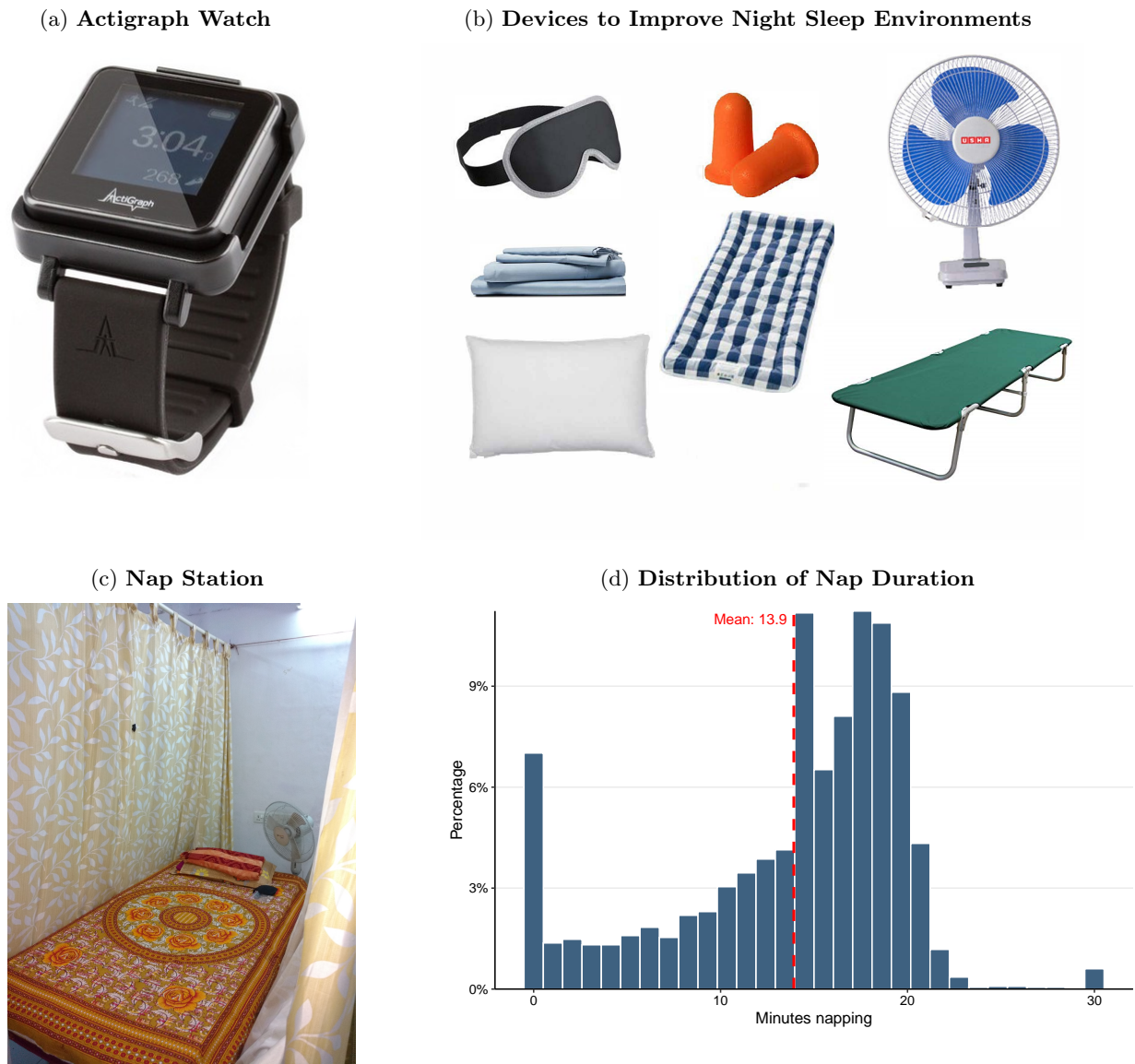
Notes: This figure illustrates devices and treatments used to measure and increase study participants' sleep.

- 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 two night-sleep treatment groups. 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 of the study and retrieved on day 28 by surveyors. A subset of the participants in the control group received household goods unrelated to sleep such as a wall calendar (not shown in the picture) in order to allow us to test for (and if needed, estimate) experimental demand or reciprocity effects.

- 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 gender-separated nap stations were located on a separate floor at the study office. The participants in the No Nap Group were not allowed to use the nap stations.

- Panel (d) 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.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, start)**

(f) **Right side (non-salient, remainder)**
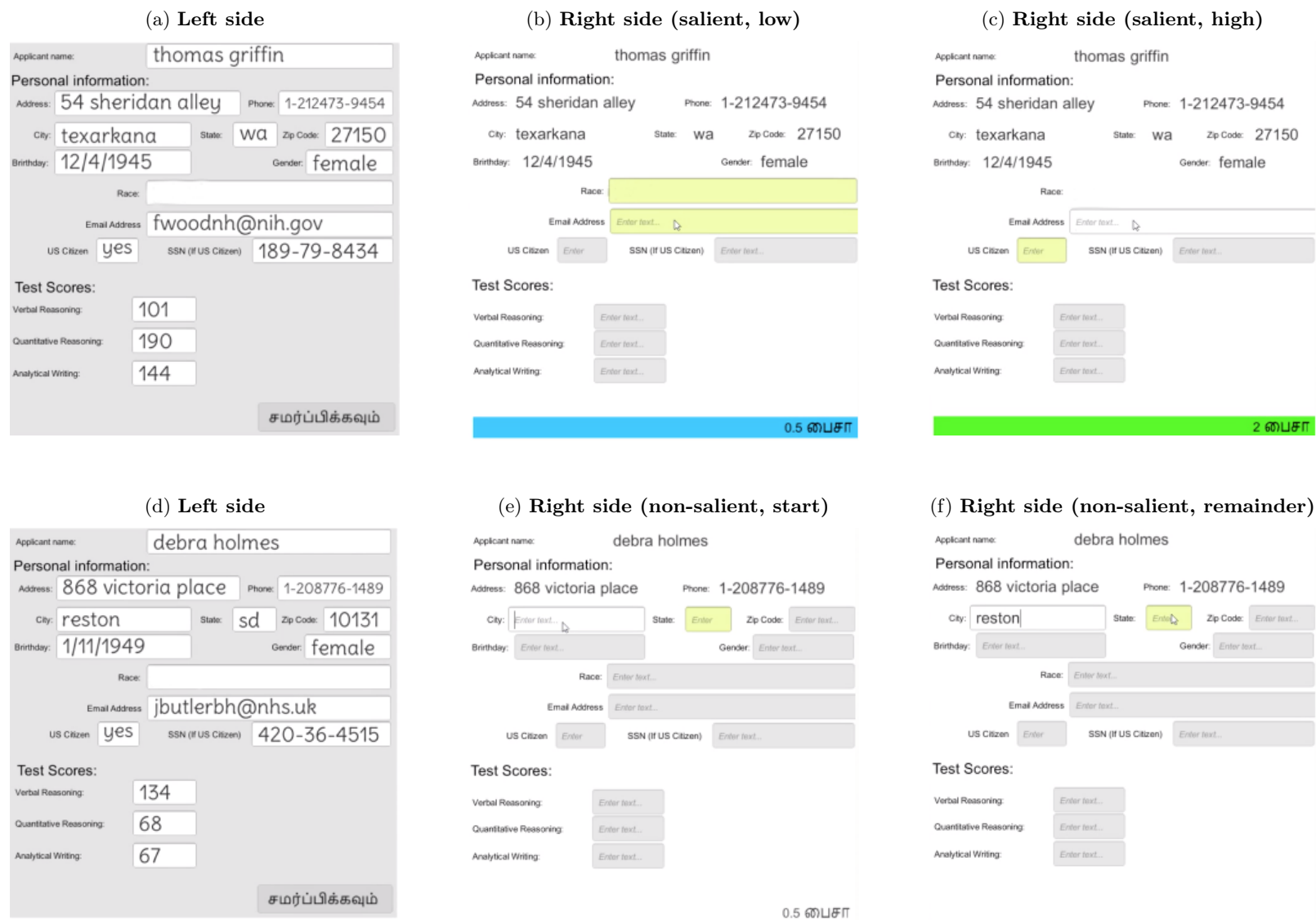
Notes: This figure shows screen shots of the data-entry task interface used by participants.

- 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 Sleep; Take-up and Usage of Devices to Improve Sleep

(a) **Factors Interfering with Study Participants' Sleep**
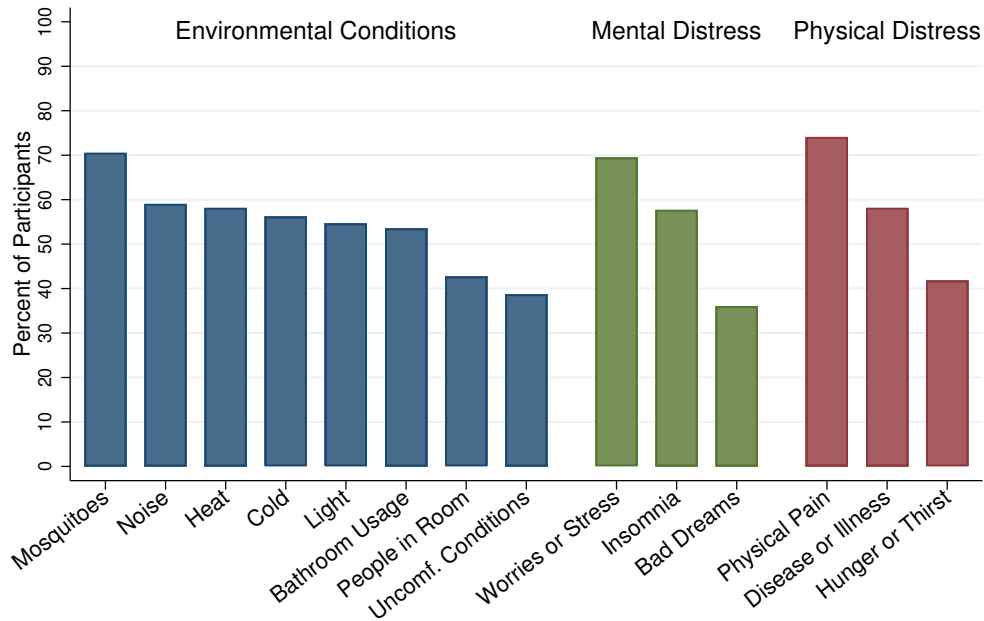
(b) **Take-up of Night-Sleep Devices**
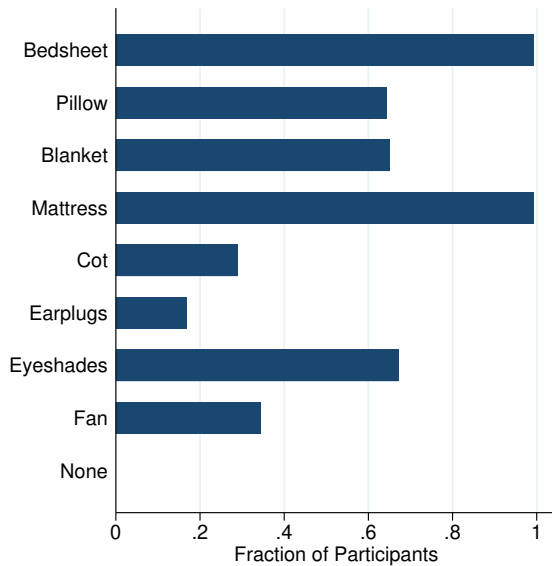
(c) **Self-reported Usage**
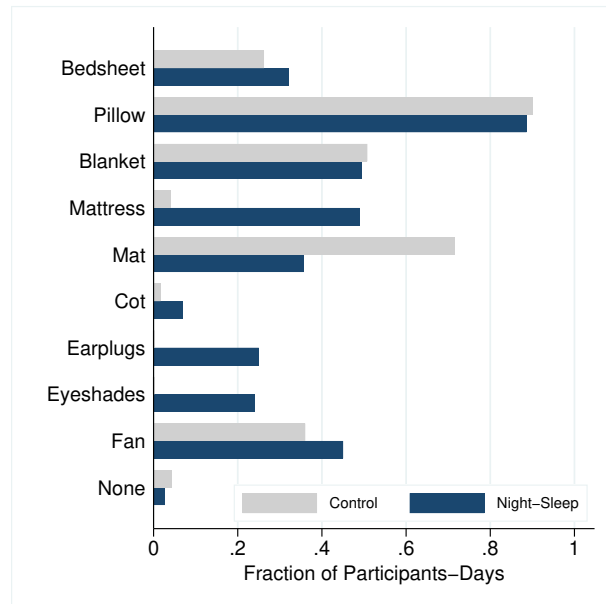
Notes: This figure illustrates the factors interfering with participants' sleep (panel a) as well as take-up and self-reported usage of devices to improve night sleep (panels b and c).

- Panel (a) figure shows the fraction of participants who reported various factors — including environmental conditions, mental distress, and physical distress — impacting their sleep "some" or "a lot". These questions were asked of each participant six times throughout the study, and a factor is recorded as impacting a participant's sleep if they report it at least once 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."

- Panels (b) shows the fraction of participants in the two Night-sleep Treatment Groups who took home at least one of each offered sleep devices.

- Panel (c) shows the share of participants who reports using the device on any given day. These numbers include devices that were not offered by the study, e.g. devices that participants owned prior to joining the study. Usage by participants in the two Night-sleep Treatment Groups are in blue. Usage by Control Group participants is shown in grey.

FIGURE A.V: Fully-Disaggregated Impacts on Nighttime and Nap Sleep

(a) **Night Sleep**

(b) **Time in Bed**

(c) **Sleep Efficiency**

(d) **Nap Sleep**
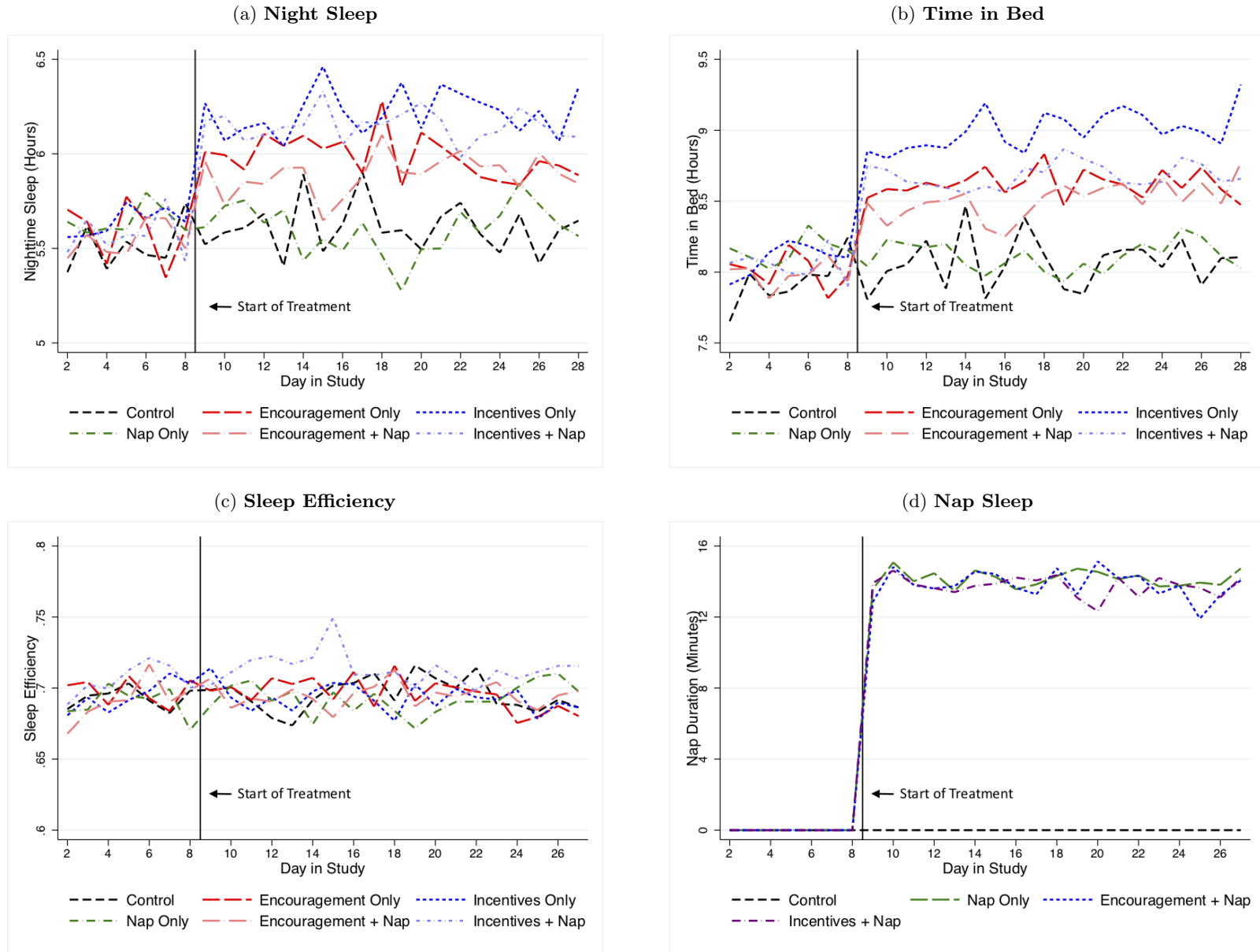
Notes: This figure shows the average of key sleep-related variables for different treatment arms by day in study of the RCT. The different panels show exactly the same variables as in Figure V but the different treatment groups are fully disaggregated. All outcomes are actigraph measures. In panels (a) and (b), we plot hours of night sleep and hours in bed at night, respectively. In panel (c), we plot 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 A.VI: Night-Sleep and Nap Treatment Effects throughout the Study

(a) **Earnings**

(b) **Hours Typing**

(c) **Productivity**
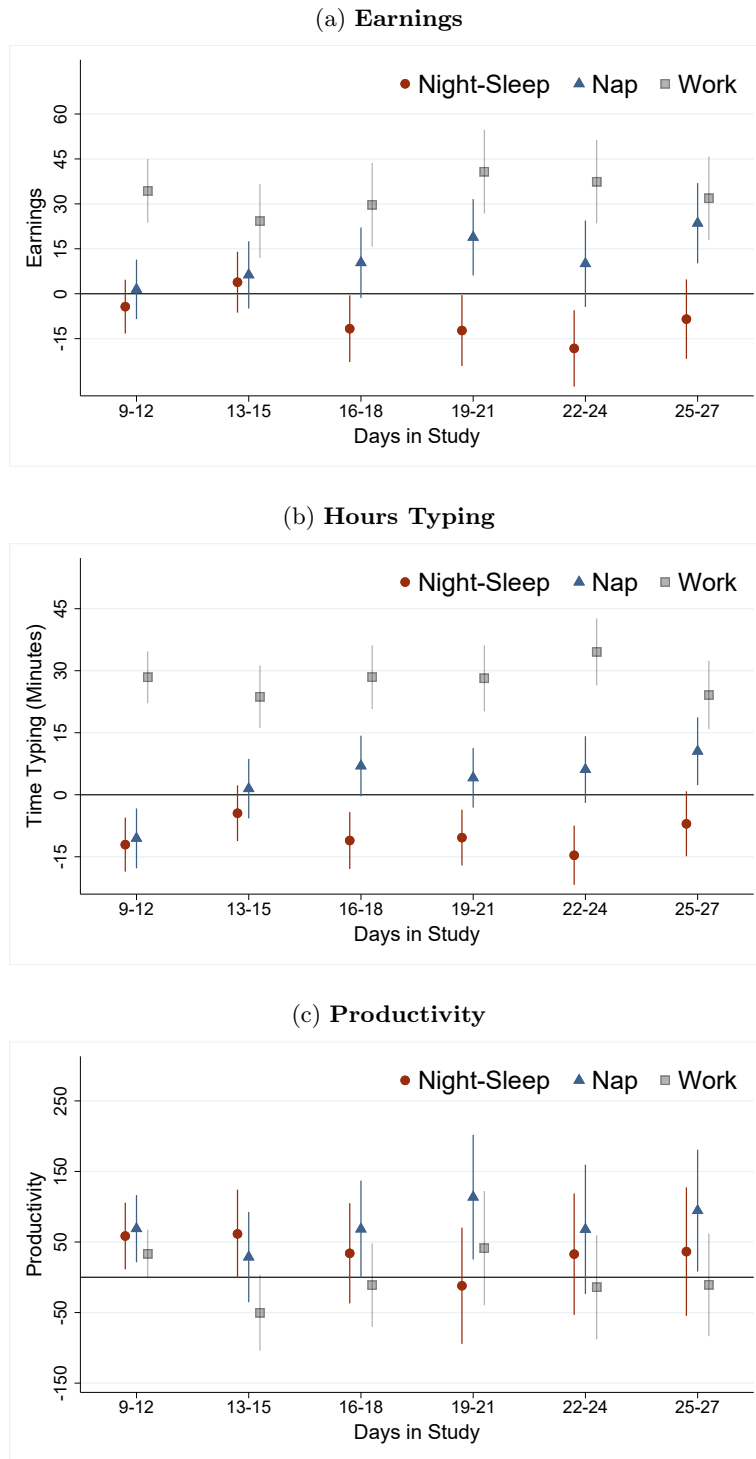
Notes: This figure shows the dynamics of the nap and night-sleep treatment effects throughout the treatment period.

- The blue series (triangles) shows the nap treatment effect in comparison to participants on Break days (i.e., not assigned to the nap treatment and not allowed to work during the nap break that day). The grey series (squares) compares participants on Work days (i.e., participants not assigned to nap but allowed to work during the nap time that day) with participants on Break days. The red series (circle) shows the night-sleep treatments effect.

- Each graph shows regression coefficients of the outcome variables earnings (panel a), hours typing (panel b), and productivity (panel c). The regressions follow the specification in equation (1), except that we interact different treatment indicator variables and controls with how long participants have been in the study (3-days dummies). The bars represent 90% confidence intervals.

FIGURE A.VII: Attendance by Treatment Group over the Course of the Study

(a) **Fraction of Participants Present by Night-Sleep Group**

(b) **Fraction of Participants Present by Nap Group**
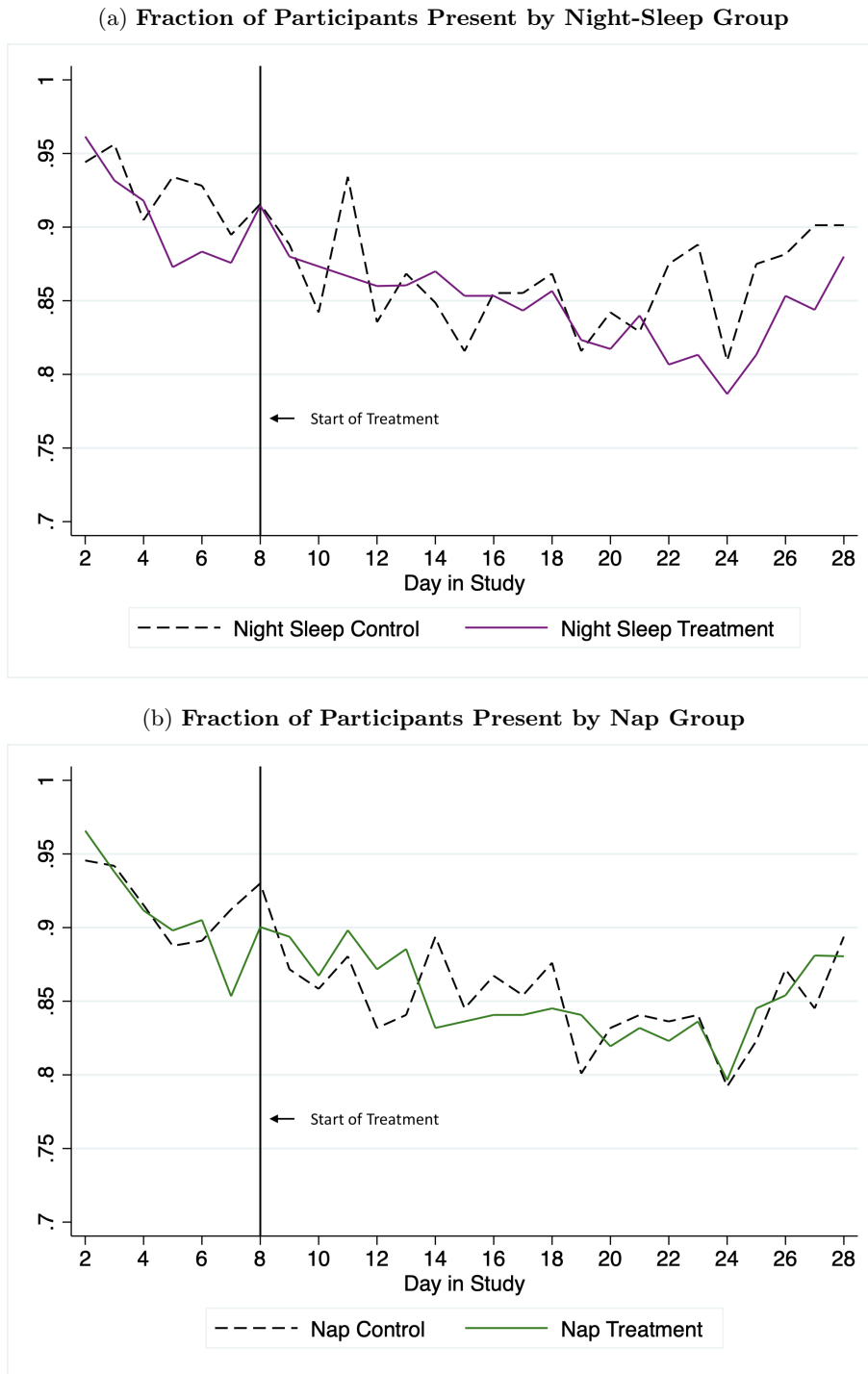
Notes: This figure shows the fraction of participants present by day of the study and treatment group.

- 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 while the dashed black line represents the corresponding (no nap) control group.

FIGURE A.VIII: Nap Treatment Effects on Productivity throughout the Day
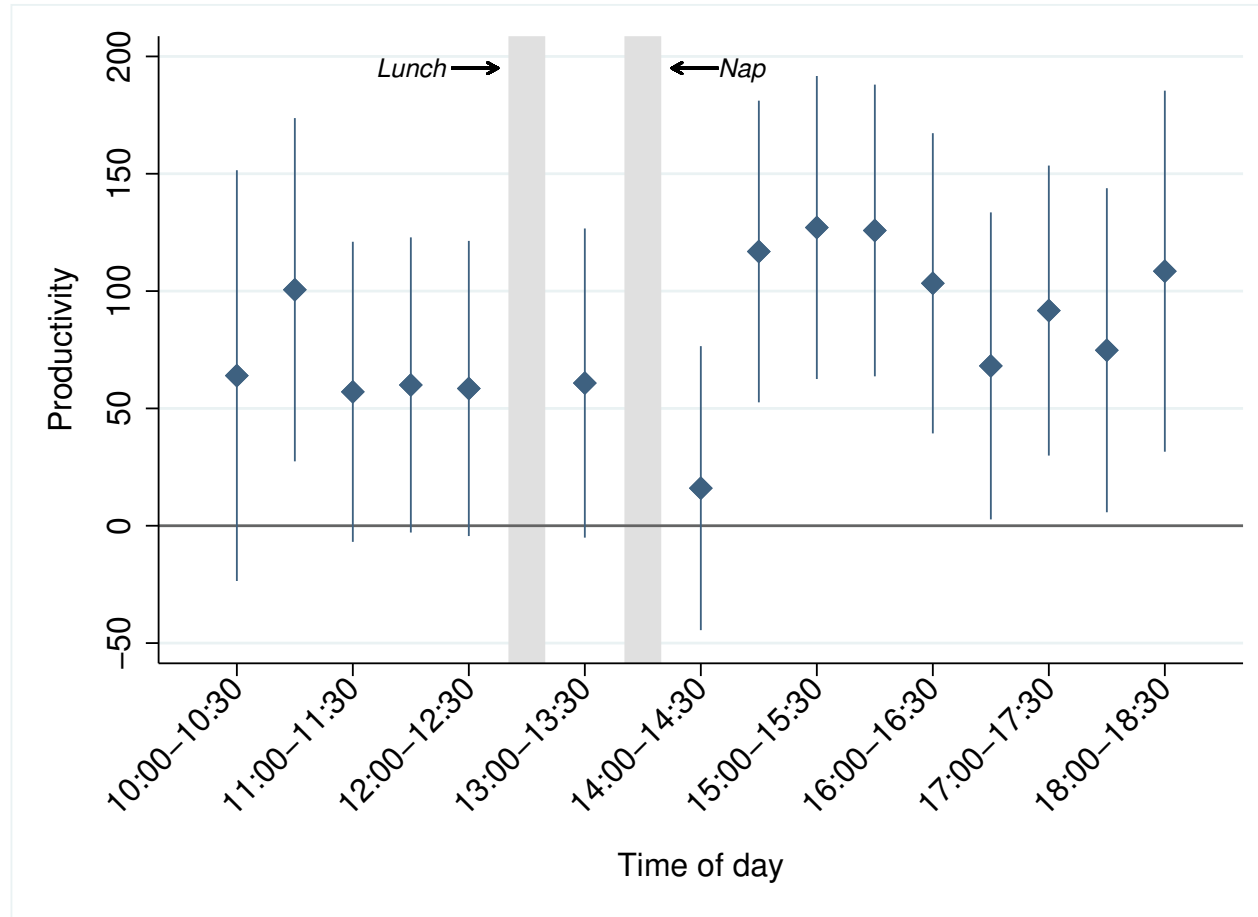
Notes: This figure shows the nap treatment effect over the course of the day during the treatment period. Each point shows the coefficient of a regression of productivity on a nap treatment status dummy interacted with time of the day (30-minute bins). The omitted group is the no nap group, including 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.

FIGURE A.IX: Relationship Between Increases in 24-Hour Sleep and Treatment Effects on Overall Index
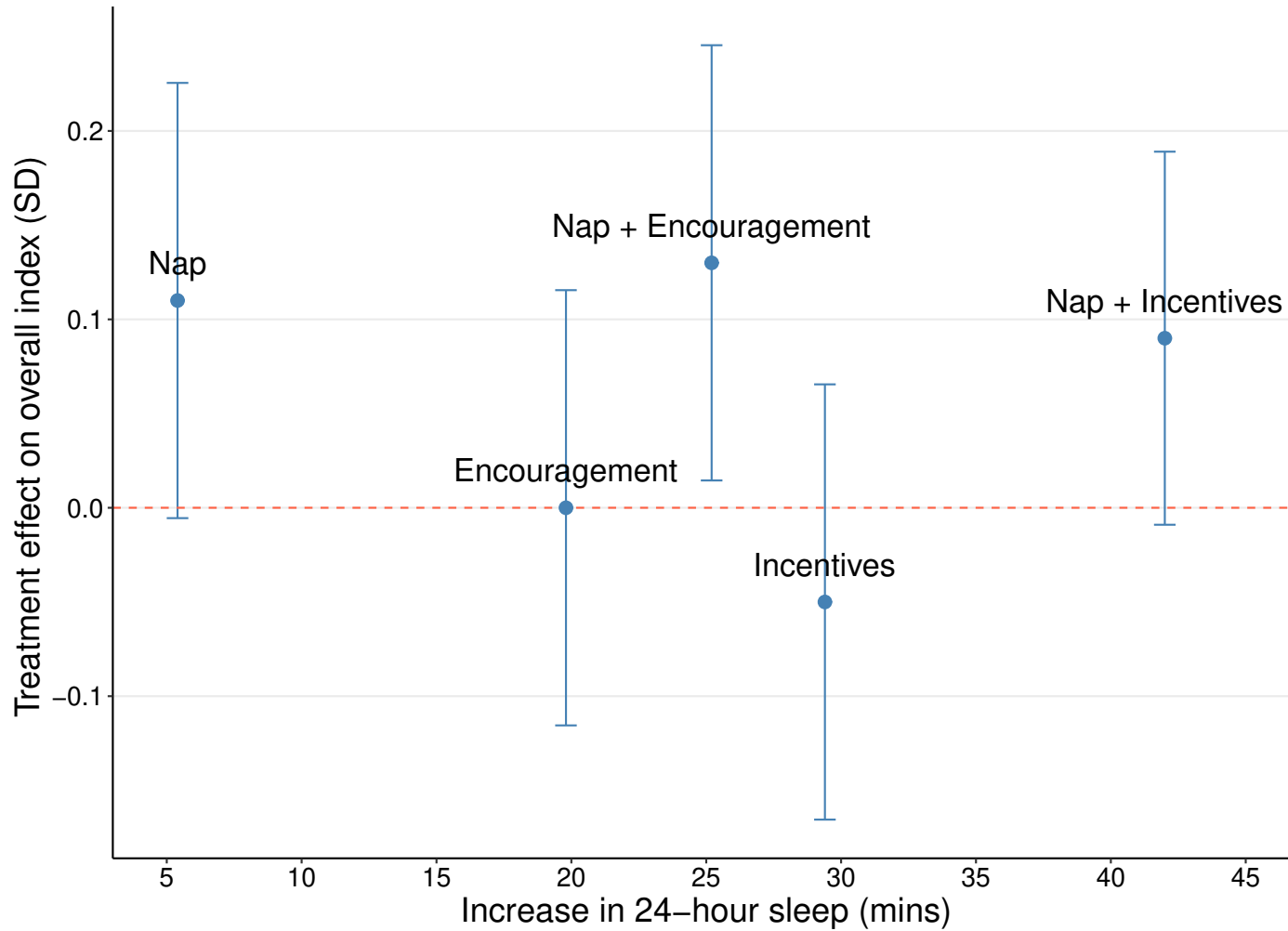
Notes: This figure shows the relationship between the estimated minutes of increased sleep within a 24 hour period due to the treatments (x-axis) and the standardized treatment effects on the overall index for the five different treatment arms (compared to participants not in the nap nor night-sleep treatments) (y-axis). Each point is labeled with the name of the corresponding treatment arm. We abbreviate the "Sleep Devices+Encouragement" to "Encouragement" and "Sleep Devices+Incentives" to "Incentives". Blue bars plot 90% confidence intervals for the treatment effect on the overall index.

B   Online Only Supplementary Tables

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

	Averages						p-values				
							Test of Equality of:				
	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 shows baseline sample characteristics and tests for differences between the fully disaggregated randomized experimental arms at baseline.

- Headers above the columns indicate combinations of night-sleep treatment groups (first row) and nap treatment status (second row).
- Columns 1 to 6 show baseline means and standard errors (in parentheses) for each treatment arm.
- Columns 7 to 11 show *p*-values of *t*-tests between column 1 and each of the other columns, respectively.
- Panel A shows demographics. Panel B shows sleep-related summary statistics. The table continues on the next page.

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

	Averages						p-values				
									Test of Equality of:		
	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. Well-Being, Cognition, Preferences											
Baseline Well-being	-0.00	0.04	0.09	-0.02	-0.00	0.03	0.59	0.17	0.81	1.00	0.65
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)					
Baseline Cognition	0.01	0.10	0.03	0.09	0.03	0.08	0.39	0.87	0.47	0.84	0.53
	(0.08)	(0.07)	(0.08)	(0.07)	(0.07)	(0.08)					
Baseline Preferences	0.00	0.02	0.08	0.12	-0.01	0.13	0.80	0.25	0.08	0.86	0.05
	(0.04)	(0.06)	(0.04)	(0.05)	(0.04)	(0.05)					
Panel D. Baseline Work											
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)					
Joint Orthogonality Test							0.96	0.34	0.50	0.74	0.38
Number of Participants	77	75	74	75	75	76					

Notes: This table is a continuation of the table on the previous page. It shows baseline sample characteristics and tests for differences between the fully disaggregated randomized experimental arms at baseline.

- The row and column structure is as described in part 1 of the table.

- Panel C shows normalized indices for well-being, cognition, and preferences. Panel D shows key work-related variables at baseline. Productivity is the output in the data-entry task per hour typing and earnings refers to data-entry related earnings.

- The Joint Orthogonality Test row refers to p-values from F-tests of a regression of each treatment dummy on all variables present in the balance table (including both part 1 and part 2). 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 (1)	Treatment (2)	1=2 (3)	Control (4)	Treatment (5)	4=5 (6)
Panel A. Demographics						
Female	0.66	0.66	0.91	0.65	0.66	0.62
	(0.04)	(0.03)		(0.03)	(0.03)	
Age	35.84	34.50	0.06	34.94	34.50	0.97
	(0.62)	(0.40)		(0.46)	(0.40)	
Number of Children	1.42	1.32	0.34	1.30	1.32	0.29
	(0.09)	(0.06)		(0.07)	(0.06)	
Years of Education	10.35	10.10	0.39	10.34	10.10	0.26
	(0.23)	(0.17)		(0.19)	(0.17)	
Familiar with Computer	0.30	0.33	0.67	0.35	0.33	0.41
	(0.07)	(0.05)		(0.06)	(0.05)	
Unemployed	0.95	0.94	0.54	0.95	0.94	0.84
	(0.02)	(0.01)		(0.01)	(0.01)	
Panel B. Baseline Sleep						
Self-Reported Night Sleep (Hrs)	7.22	7.18	0.63	7.24	7.18	0.24
	(0.08)	(0.05)		(0.07)	(0.05)	
Actigraph Night Sleep (Hrs)	5.57	5.58	0.85	5.57	5.58	0.89
	(0.07)	(0.05)		(0.06)	(0.05)	
Actigraph Time in Bed (Hrs)	8.11	8.11	0.94	8.09	8.11	0.66
	(0.07)	(0.05)		(0.06)	(0.05)	
Sleep Efficiency	0.69	0.70	0.75	0.70	0.70	0.77
	(0.01)	(0.00)		(0.01)	(0.00)	
Number of Sleep Devices Owned	2.52	2.53	0.97	2.54	2.53	0.87
	(0.13)	(0.09)		(0.11)	(0.09)	
Number of Participants	152	300		226	226	

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

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

- Panel A shows demographics. Panel B shows sleep-related summary statistics. The table continues on the next page.

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

	Night-Sleep Treatments			Nap Treatments		
	Control (1)	Treatment (2)	1=2 (3)	Control (4)	Treatment (5)	4=5 (6)
Panel C. Well-Being, Cognition, and Preferences						
Baseline Well-Being	-0.01	0.04	0.25	0.04	0.04	0.34
	(0.04)	(0.02)		(0.03)	(0.02)	
Baseline Cognition	0.05	0.06	0.86	0.05	0.06	0.76
	(0.05)	(0.04)		(0.04)	(0.04)	
Baseline Preferences	0.06	0.05	0.93	0.03	0.05	0.22
	(0.03)	(0.02)		(0.03)	(0.02)	
Panel D. Baseline Work						
Typing Time (Hrs)	4.49	4.47	0.86	4.48	4.47	0.99
	(0.05)	(0.06)		(0.04)	(0.06)	
Time in Office (Hrs)	7.95	7.89	0.36	7.92	7.89	0.70
	(0.06)	(0.04)		(0.05)	(0.04)	
Productivity	2373.51	2459.66	0.59	2558.52	2459.66	0.09
	(127.65)	(93.42)		(118.05)	(93.42)	
Earnings	234.63	234.05	0.96	243.64	234.05	0.05
	(8.52)	(5.93)		(7.72)	(5.93)	
Attendance	0.95	0.94	0.06	0.94	0.94	0.28
	(0.01)	(0.00)		(0.00)	(0.00)	
Joint Orthogonality Test			0.61			0.67
Number of Participants	152	300		226	226	

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

- The row and column structure is as described in part 1 of the table.

- Panel C shows normalized indices for well-being, cognition, and preferences. Panel D shows key work-related variables at baseline. Productivity is the output in the data-entry task per hour typing and earnings refers to data-entry related earnings.

- The Joint Orthogonality Test row refers to p-values from F-tests of a regression of each treatment dummy on all variables present in the balance table (including both part 1 and part 2). 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 and Activities 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
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	p25	Median	p75	N	Mean	p25	Median	p75	N	Mean	p25	Median	p75	N	Mean	p25	Median	p75	N
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, the control group mean for each outcome, and a benchmark describing how responsive participants' productivity was to a change in incentives. 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. Where required, signs are flipped such that higher values indicate 'better' outcomes. The sleep science experts predicted the ITT estimate 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*** (0.05)	0.46*** (0.08)	0.46*** (0.08)	0.38*** (0.12)	0.37*** (0.11)	-0.14 (0.42)	-0.18 (0.65)	-0.17 (0.65)	-1.29 (0.80)	-0.59 (0.79)					
X		-0.04 (0.11)	0.15* (0.09)	-0.11 (0.11)	-0.11 (0.10)		0.15 (0.73)	0.13 (0.81)	-0.62 (0.78)	0.71 (0.82)		0.09 (0.26)	-0.40 (0.43)	-0.10 (0.18)	-0.40 (0.21)
Night-Sleep Treatments*X		0.03 (0.10)	0.04 (0.10)	0.12 (0.14)	0.17 (0.12)		0.09 (0.82)	0.06 (0.86)	1.56 (0.97)	0.68 (0.95)					
Nap Treatment											13.84*** (0.30)	13.62*** (0.42)	13.34*** (0.42)	13.21*** (0.63)	13.32*** (0.55)
Nap Treatment*X												0.48 (0.58)	0.98* (0.58)	0.85 (0.70)	0.78 (0.65)
DV Control Group Mean	5.60	5.60	5.60	5.60	5.60	69.83	69.83	69.83	69.83	69.83	0	0	0	0	0
Participants	451	451	451	451	451	451	451	451	451	451	452	452	452	452	452

Notes: This table shows 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: (i) cols 1-5: night sleep duration (in hours) as measured by the actigraph; (ii) cols 6-10: night sleep efficiency (sleep duration/time in bed) as measured by the actigraph; cols 11-15: Nap duration (in minutes) as measured by the actigraph. Cols 1, 6, and 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 (work and break days jointly). Regressions are at the participant-day level and standard errors (in parentheses) are clustered at the participant level.

Table A.VI: Treatment Effects on Sleep, Pooling Night-Sleep Treatments and Including Bounds

	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)
Lee Lower Bound	0.41	0.59	-0.29	-0.00	0.40
Lee Upper Bound	0.48	0.69	0.22	-0.00	0.48
Confidence Interval	[0.31, 0.57]	[0.48, 0.80]	[-1.11, 1.03]	[-0.01, 0.01]	[0.31, 0.57]
Devices+Encouragement	0.33***	0.51***	-0.44	-0.00	0.33***
	(0.06)	(0.06)	(0.47)	(0.01)	(0.06)
Lee Lower Bound	0.31	0.49	-0.60	-0.00	0.31
Lee Upper Bound	0.37	0.57	-0.11	-0.00	0.37
Confidence Interval	[0.20, 0.48]	[0.37, 0.69]	[-1.52, 0.80]	[-0.01, 0.01]	[0.20, 0.48]
Devices+Incentives	0.55***	0.76***	0.22	-0.00	0.55***
	(0.06)	(0.07)	(0.49)	(0.01)	(0.06)
Lee Lower Bound	0.51	0.70	0.02	-0.00	0.50
Lee Upper Bound	0.59	0.80	0.54	-0.00	0.59
Confidence Interval	[0.40, 0.70]	[0.57, 0.93]	[-0.93, 1.48]	[-0.01, 0.01]	[0.39, 0.70]
Nap Treatment	-0.08*	-0.17***	0.27	0.24***	0.13**
	(0.05)	(0.05)	(0.40)	(0.00)	(0.05)
Lee Lower Bound	-0.13	-0.22	-0.16	0.24	0.08
Lee Upper Bound	-0.04	-0.11	0.52	0.24	0.17
Confidence Interval	[-0.22, 0.05]	[-0.32, -0.01]	[-0.91, 1.28]	[0.23, 0.25]	[-0.02, 0.27]
Control Mean	5.61	8.07	69.86	0.00	5.61
Control SD	1.20	1.37	11.28	0.00	1.20
Participant-Nights	8454	9036	8454	7191	8035
Participants	451	452	451	450	451

Notes: Compliance with wearing the actigraphs was high (94%) and not statistically different across treatment groups. This table accounts for imperfect compliance of actigraph wearing as well as occasional missing sleep measurements (e.g. due to batteries running out of charge) by showing Lee (2009) bounds and associated 95% confidence intervals for the treatment effect of the night-sleep and nap interventions on sleep outcomes.

- Night sleep, time in bed, nap sleep and 24-hour sleep (cols 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 in the office and night sleep.

- The first bold row (Night-Sleep Treatments) and the fourth bold row (Nap Treatment) are generated from the same regression. They show estimates of the pooled treatment effect for the two night-sleep treatments and the treatment effect for the Nap Treatment, respectively. The regression does not include a separate indicator for the group that receives both the Night-Sleep and Nap Treatments.

- The second and third bold (indented) rows show the relative impacts of each of the night-sleep treatments: the Devices + Encouragement and the Devices + Incentives Treatment. These estimates are generated from a separate regression which controls for Nap Treatment status as well as the controls listed below.

- The first row for each treatment shows the OLS estimates of equation (1) for each group, controlling for the average baseline measure of the dependent variable, age, sex, and day-in-study and date fixed effects. Standard errors, presented in parentheses below each point estimate, are clustered at the participant level.

- The rows named 'Lower Bound' and 'Upper Bound' are the estimates for the Lower and Upper Lee Bounds for the corresponding point estimate. The rows named 'Confidence Interval' show the interval that covers the identified set with 95% confidence.

Table A.VII: Main Treatment Effects, Fully-Disaggregated and Including Multiple Hypothesis Testing Corrections

	OVERALL	WORK				WELL-BEING		
	Index (1)	Earnings (2)	Productivity (3)	Labor Supply (4)	Output (5)	Index (6)	Physical (7)	Psychological (8)
Devices+Encouragement Only	**0.00** {1.00}	**-0.07** {0.03} [0.11]	-0.02 {0.43} [0.45]	-0.06 {0.10} [0.20]	-0.07 {0.02} [0.06]	**0.13** {0.01} [0.04]	0.16 {0.01} [0.02]	0.02 {0.79} [0.80]
Devices+Incentives Only	**-0.05** {0.48}	**-0.07** {0.05} [0.22]	-0.02 {0.57} [0.57]	-0.08 {0.05} [0.10]	-0.08 {0.02} [0.05]	**0.05** {0.31} [0.69]	0.08 {0.21} [0.40]	-0.02 {0.80} [0.81]
Nap Only	**0.11** {0.10}	**-0.07** {0.10} [0.28]	-0.01 {0.65} [0.66]	-0.07 {0.07} [0.16]	-0.05 {0.18} [0.27]	**0.18** {0.00} [0.01]	0.16 {0.01} [0.03]	0.19 {0.05} [0.06]
Devices+Encouragement and Nap	**0.13** {0.06}	**-0.07** {0.07} [0.21]	0.04 {0.15} [0.20]	-0.15 {0.00} [0.00]	-0.04 {0.19} [0.20]	**0.13** {0.01} [0.04]	0.12 {0.06} [0.12]	0.11 {0.21} [0.22]
Devices+Incentives and Nap	**0.09** {0.15}	**-0.07** {0.05} [0.10]	0.04 {0.10} [0.11]	-0.17 {0.00} [0.00]	-0.07 {0.04} [0.07]	**0.11** {0.03} [0.10]	0.12 {0.06} [0.12]	0.07 {0.46} [0.47]
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)
Devices+Encouragement Only	**-0.00** {0.95} [0.96]	-0.00 {0.98} [0.97]	0.04 {0.82} [0.97]	**-0.09** {0.16} [0.31]	-0.05 {0.68} [0.68]	-0.10 {0.27} [0.48]	-0.15 {0.24} [0.48]
Devices+Incentives Only	**-0.03** {0.74} [0.75]	0.04 {0.54} [0.54]	-0.12 {0.54} [0.54]	**-0.04** {0.52} [0.75]	0.11 {0.36} [0.59]	-0.16 {0.05} [0.16]	-0.06 {0.68} [0.68]
Nap Only	**0.09** {0.23} [0.42]	0.07 {0.30} [0.38]	0.15 {0.37} [0.38]	**-0.01** {0.85} [0.85]	0.12 {0.34} [0.56]	-0.09 {0.33} [0.56]	-0.01 {0.92} [0.92]
Devices+Encouragement and Nap	**0.05** {0.54} [0.55]	0.13 {0.08} [0.17]	0.03 {0.88} [0.88]	**0.09** {0.17} [0.32]	0.23 {0.05} [0.16]	0.03 {0.79} [0.84]	0.03 {0.83} [0.84]
Devices+Incentives and Nap	**0.14** {0.04} [0.10]	0.09 {0.18} [0.18]	0.32 {0.04} [0.08]	**0.01** {0.87} [0.87]	0.11 {0.38} [0.62]	-0.10 {0.28} [0.61]	0.01 {0.96} [0.96]
Participants	452	452	429	452	452	415	415

Notes: This table shows the treatment effects of the five fully-disaggregated treatment arms on the overall index as well as the four families of outcomes.

- This table is exactly the same as Table III except for that it also includes p-values that account for multiple hypothesis testing.
- Below each coefficient, we report (i) the corresponding standard errors in parentheses ($\cdot$), robust to heteroscedasticity and clustered at the participant-level when applicable, (ii) the unadjusted p-value in curly brackets {$\cdot$}, and (iii) the Westfall-Young FWER-adjusted p-value in square brackets [$\cdot$], as described in Section E.

Table A.VIII: Main Treatment Effects, Pooling Night-Sleep Treatments and Including Multiple Hypothesis Testing Corrections

	OVERALL	WORK				WELL-BEING		
	Index (1)	Earnings (2)	Productivity (3)	Labor Supply (4)	Output (5)	Index (6)	Physical (7)	Psychological (8)
Night-Sleep Only	**-0.02** (0.06) {0.69}	**-0.07** (0.03) {0.02} [0.10]	-0.02 (0.02) {0.43} [0.44]	-0.07 (0.03) {0.04} [0.09]	-0.08 (0.03) {0.01} [0.02]	**0.09** (0.04) {0.04} [0.11]	0.12 (0.05) {0.03} [0.07]	0.00 (0.08) {0.99} [0.99]
Nap Only	**0.11** (0.07) {0.10}	**-0.07** (0.04) {0.10} [0.29]	-0.01 (0.03) {0.65} [0.67]	-0.07 (0.04) {0.07} [0.16]	-0.05 (0.04) {0.17} [0.29]	**0.18** (0.06) {0.00} [0.01]	0.16 (0.07) {0.01} [0.03]	0.19 (0.10) {0.05} [0.05]
Night-Sleep and Nap	**0.11** (0.06) {0.05}	**-0.07** (0.03) {0.04} [0.11]	0.04 (0.02) {0.08} [0.09]	-0.16 (0.04) {0.00} [0.00]	-0.05 (0.03) {0.06} [0.09]	**0.12** (0.04) {0.01} [0.02]	0.12 (0.05) {} [0.07]	0.09 (0.08) {0.25} [0.25]
Participants	451	451	451	451	451	452	452	452
p-values NS vs. Nap	{0.02}	{0.79} [0.79]	{0.87} [0.94]	{0.94} [0.94]	{0.38} [0.68]	{0.06} [0.25]	{0.41} [0.67]	{0.02} [0.05]
p-values NS vs. Both	{0.01}	{0.87} [0.87]	{0.00} [0.01]	{0.00} [0.01]	{0.29} [0.31]	{0.35} [0.61]	{0.97} [1.00]	{0.16} [0.31]
p-values Nap vs. Both	{0.98}	{0.88} [0.91]	{0.04} [0.09]	{0.01} [0.01]	{0.88} [0.88]	{0.22} [0.62]	{0.39} [0.67]	{0.23} [0.41]

	COGNITION			PREFERENCES			
	Index (9)	Lab Tasks (10)	Work Task (11)	Index (12)	Time (13)	Social (14)	Risk (15)
Night-Sleep Only	**-0.02** (0.06) {0.81} [0.81]	0.02 (0.06) {0.75} [0.78]	-0.04 (0.15) {0.79} [0.78]	**-0.07** (0.06) {0.24} [0.42]	0.03 (0.11) {0.80} [0.81]	-0.13 (0.08) {0.08} [0.24]	-0.10 (0.11) {0.37} [0.61]
Nap Only	**0.09** (0.07) {0.23} [0.41]	0.07 (0.07) {0.30} [0.37]	0.15 (0.17) {0.36} [0.37]	**-0.01** (0.07) {0.86} [0.85]	0.12 (0.12) {0.34} [0.57]	-0.09 (0.09) {0.33} [0.57]	-0.01 (0.14) {0.92} [0.92]
Night-Sleep and Nap	**0.09** (0.06) {0.13} [0.25]	0.11 (0.06) {0.08} [0.16]	0.18 (0.14) {0.22} [0.22]	**0.05** (0.06) {0.38} [0.38]	0.17 (0.11) {0.12} [0.34]	-0.04 (0.08) {0.67} [0.89]	0.02 (0.11) {0.89} [0.89]
Participants	452	452	429	452	452	415	415
p-values NS vs. Nap	{0.13} [0.36]	{0.35} [0.59]	{0.23} [0.40]	{0.32} [0.55]	{0.33} [0.69]	{0.55} [0.57]	{0.45} [0.57]
p-values NS vs. Both	{0.06} [0.17]	{0.06} [0.13]	{0.11} [0.22]	{0.01} [0.07]	{0.06} [0.18]	{0.15} [0.21]	{0.20} [0.21]
p-values Nap vs. Both	{0.91} [0.91]	{0.53} [0.78]	{0.86} [0.98]	{0.26} [0.62]	{0.58} [0.93]	{0.53} [0.80]	{0.80} [0.80]

Notes: This table shows the treatment effects on the overall index as well as the four families of outcomes: work, well-being, cognition, and preferences.

- This table is identical to Table III, except for that (i) it pools the two Night-Sleep Treatments and (ii) it provides *p*-values to test for the equality of treatment effects between the different treatment cells.

- Below each coefficient, we report (i) the corresponding standard errors in parentheses (·), robust to heteroscedasticity and clustered at the participant-level when applicable, (ii) the unadjusted *p*-value in curly brackets {·}, and (iii) the Westfall-Young FWER-adjusted *p*-value in square brackets [·], as described in Section E.

- The three rows at the bottom of each panel show the p-values from tests of equality across (i) the Nap only vs. the Night-Sleep only coefficients, (ii) the Night-Sleep and Nap vs. the Night-Sleep only coefficients, and (iii) the Night-Sleep and Nap vs. the Nap only coefficients. Unadjusted *p*-value are in curly brackets {·} and Westfall-Young FWER-adjusted *p*-value are in square brackets [·].

Table A.IX: Main Treatment Effects: Breakdown Break vs. Work

	OVERALL	WORK				WELL-BEING		
	Index (1)	Earnings (2)	Productivity (3)	Labor Supply (4)	Output (5)	Index (6)	Physical (7)	Psychological (8)
Night-Sleep Treatments	**-0.01** (0.04) {0.79}	**-0.04** (0.02) {0.08} [0.30]	0.02 (0.02) {0.30} [0.32]	-0.08 (0.02) {0.00} [0.00]	-0.04 (0.02) {0.05} [0.09]	**0.01** (0.03) {0.66} [0.96]	0.04 (0.04) {0.34} [0.36]	-0.05 (0.06) {0.36} [0.36]
Nap Treatment vs. Pooled Control	**0.12** (0.04) {0.00}	**-0.02** (0.02) {0.23} [0.24]	0.04 (0.02) {0.03} [0.06]	-0.09 (0.02) {0.00} [0.00]	-0.01 (0.02) {0.77} [0.78]	**0.08** (0.03) {0.01} [0.02]	0.05 (0.03) {0.12} [0.24]	0.12 (0.05) {0.02} [0.04]
Nap Treatment vs. Break Day	**0.15** (0.04) {0.00}	**0.05** (0.02) {0.02} [0.05]	0.04 (0.02) {0.03} [0.06]	0.01 (0.02) {0.60} [0.60]	0.05 (0.02) {0.01} [0.02]			
Nap Treatment vs. Work Day	**0.08** (0.04) {0.03}	**-0.10** (0.02) {0.00} [0.00]	0.03 (0.02) {0.04} [0.05]	-0.20 (0.02) {0.00} [0.00]	-0.07 (0.02) {0.00} [0.00]			
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** (0.05) {0.93} [0.96]	0.03 (0.04) {0.51} [0.76]	-0.01 (0.11) {0.95} [0.95]	**-0.00** (0.04) {0.95} [0.96]	0.04 (0.07) {0.58} [0.65]	-0.04 (0.06) {0.45} [0.65]	-0.04 (0.08) {0.65} [0.65]
Nap Treatment vs. Pooled Control	**0.10** (0.05) {0.03} [0.08]	0.08 (0.04) {0.03} [0.06]	0.20 (0.11) {0.07} [0.07]	**0.07** (0.04) {0.05} [0.10]	0.13 (0.06) {0.04} [0.11]	0.03 (0.05) {0.52} [0.51]	0.07 (0.08) {0.33} [0.51]
Participants	452	452	429	452	452	415	415

Notes: This table shows the treatment effects of the night-sleep intervention and the nap intervention on the overall index as well as the four families of outcomes: work, well-being, cognition, and preferences. This table closely follows Table IV, with the exception that for the work outcomes, it splits the nap treatment into two comparisons: Nap vs. Break Days and Nap vs. Work Days. These comparisons are well-defined for the work outcomes, which are measured every day and estimated using Equation 1. See Section 4.1 for more details.

- The first two rows of the table are from the same regressions as Table IV, displaying coefficients of (pooled) Night-Sleep and Nap Treatments compared to the Control Group that receives no sleep-related treatments. The outcomes correspond to those in Tables III and IV, described in more detail in the corresponding table notes and in Section 3.2.

- Rows 3 and 4 in cols 1 through 5 are based on *separate* regressions. "Nap Treatment vs. Break Day" in row 3 reports the "Nap Treatment" coefficient of a regression that is identical to the regression in rows 1 and 2 but also includes a "Work Day" dummy. It thus compares the nap treatment group to the control group workers on days when they were randomized to have "Break Days". Similarly, "Nap Treatment vs. Work Day" in row 4 reports the "Nap Treatment" coefficient of a regression that also includes "Break Day" dummy, thus comparing the nap treatment group to control group workers on days when they were randomized to have "Work Days".

- Below each coefficient, we report (i) the corresponding standard errors in parentheses (·), robust to heteroscedasticity and clustered at the participant-level when applicable, (ii) the unadjusted *p*-value in curly brackets {·}, and (iii) the Westfall-Young FWER-adjusted *p*-value in square brackets [·], as described in Section E.

Table A.X: Impacts on Time Allocation

	Labor Supply				Sleep	
	Time at Office (1)	Arrival Time (2)	Leave Time (3)	Work Breaks (4)	Get Up Time (5)	Bed Time (6)
Night-Sleep Treatments	-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 Treatment	0.06	-0.00	0.05**	-0.01	-0.05	0.14***
	(0.04)	(0.03)	(0.03)	(0.01)	(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.

- As in Table IV, this table pools the two Night-Sleep Treatments and does not include include a separate indicator for the group that receives both the Night-Sleep and Nap Treatments.

- 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 24-hour military time, i.e. 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, the day type (long vs. short), and the fraction of high piece rates received in the typing task on that day.

Table A.XI: Treatment Effects on Psychological and Physical Well-being

	Panel A: Standardized Psychological Well-being 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

	Panel B: Standardized Physical Well-being Components				
	Biking (6)	Illness (7)	Pain (8)	Daily Act. (9)	BP (10)
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 shows treatment effects of the pooled night-sleep and the nap treatments on psychological and physical well-being.

- All dependent variables are normalized with respect to the Control Group's mean and standard deviation. When two or more outcomes make up a dependent variable (as in the case of biking and blood pressure), the outcomes are first normalized independently, then averaged, and then the average is also normalized.

- As in Table IV, this table pools the two Night-Sleep Treatments and does not include include a separate indicator for the group that receives both the Night-Sleep and Nap Treatments.

- The psychological well-being (Panel A) outcomes are: self-reported depression from the PHQ-9 survey (sightly modified to facilitate comprehension, col 1); self-reported happiness on a scale from 1 to 5 (col 2); responses to the "Cantril Scale" ladder of life possibility (col 3); self-reported life satisfaction on a scale from 1 to 10 (col 4); self-reported stress on a scale from 1 to 6 (col 5). Additional details of the outcome measures are located in Section C.3.

- The physical well-being (Panel B) outcomes are: the average of distance and maximum speed recorded during the endline biking task (col 6); number of self-reported illnesses over the last seven days (col 7); self-reported pain on a scale from 1 to 10 (col 8); self-reported extent to which health has limited daily activities (col 9); the average of winsorized systolic and winsorized diastolic blood pressure (col 10). Additional details of the outcome measures are located in Section C.4.

- When required, outcomes are flipped so that a positive value aligns with what would be considered a "better" outcome.

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

	Inhibitory Control			Memory	Attention			
	(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	-0.02
	(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	0.17***	0.13***	0.14***	0.00
	(0.04)	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Participants	449	449	449	449	452	452	452	452

Notes: This table considers the treatment effect of the night-sleep and nap interventions on the three laboratory measures of cognition: inhibitory control (Hearts & Flowers), memory (Corsi Block Span), and attention (PVT). For more detail on these tasks, see Dean et al. (2019).

- As in Table IV, this table pools the two night-sleep treatments and does not include include a separate indicator for the group that receives both the night-sleep and nap treatments.

- 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). Where required, signs are flipped when needed to ensure that for all variables higher indicates better performance.

- The outcomes in cols 1-3 are all related to inhibitory control, measured by the Hearts and Flowers task. The outcome variable in col 1 is the payment participants earn for completing the task, where the payment is a weighted average of the fraction of correct entries and (faster) reaction times. Cols 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 col 4 is the participants' earnings for completing the Corsi block span task, which measures memory. Payment is a function of the maximum number of blocks the participant can correctly recall.

- Cols 5-8 show outcomes related to the Psychomotor Vigilance Task, a frequently used proxy for vigilance (a form of attention). The outcome variables in col 5 is the overall payment for the PVT task. The payment is a function of three performance metrics in cols 6-8. Col 6 shows treatment effects for the inverse reaction time (reaction time captures how fast participants react to each stimulus). Col 7 depicts minor lapses (significant delays between when the signal appears and the participant acts). The outcome variable in col 8 is the number of false starts (when the participant acts before the signal is displayed). Signs are flipped for cols 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, 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. Standard errors are clustered at the participant level.

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

	Overall		Morning		Afternoon	
	Output (1)	Minutes (2)	Output (3)	Minutes (4)	Output (5)	Minutes (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 NS 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 presents 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 and in Appendix C.5.2.

- Each column shows the attention parameter (Gabaix, 2019) described in detail in Appendix C.5.2. This parameter varies from 0 to 1: a value of 0 means that participants do not 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.

- The two dependent variables from which the attention parameters are calculated are output (correct characters entered) and minutes (minutes actively typing), each of which are captured at the 30-minute incentive-session level.

- Cols 1 and 2 include the entire day. Cols 3 and 4 only use observations from the morning (i.e., pre-naps). Cols 5 and 6 include observations from the afternoon (i.e., after the nap).

- We consider attention for three groups: (row 1) the Night-Sleep Group, pooling across both night-sleep interventions, (row 2) the Nap Treatment Group, and (row 3) the Control Group, consisting of individuals assigned to neither any of the Night-Sleep Treatment Groups nor to the Nap Treatment Group.

- Rows 4 and 5 of the table show *p*-values of tests of differences between the coefficients estimated between the night-sleep treatment group and control group, and between the nap treatment and control groups, respectively.

Table A.XIV: 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 two measures of time preferences.

- **Panel A: Savings.** This task is described briefly in Section 3.2 subsection "Savings", and in greater detail in Appendix C.6.2, along with a description of the related "defaults task."

  - The dependent variable in col 1 captures daily deposits (which is equivalent to winsorizing daily net savings at Rs. 0) at the study office. Col 2 shows daily net savings (difference between deposits and withdrawals). Cols 3 and 4 show daily interest accrued on the participants' savings, with col 3 excluding individuals who were assigned a zero interest rate and col 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 cols 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 cols 3-6 is the OLS present bias parameter, the percentage decrease in effort chosen on "work-days". In cols 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 cols 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.XV: 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 treatments on risk and social preferences. These tasks are described in greater detail in Section 3.2 and Section C.6.1.

- As in Table IV, this table pools the two Night-Sleep Treatments and does not include include a separate indicator for the group that receives both the Night-Sleep and Nap Treatments.

- 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. Each column shows the OLS estimates of equation (2). Standard errors in parentheses are robust to heteroscedasticity.

- Risk preferences components use the point at which the participant switched from the risky to safe choice in the risk aversion task (col 1) and the point at which the participant switched from the risky to safe choice in the loss aversion task (col 2) as the dependant variable.

- Social preferences components include the amount of money the sender sent in the dictator game (col 3), the amount of money the sender sent in the ultimatum game (col 4), the amount of money the sender sent in the trust game (col 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 (col 6), and the average amount of money the recipient would send back to the sender in the trust game (col 7).

- These risk preferences regression specifications differ compared to those underlying the indices in Table III and IV. In particular, following the standard in the literature, non-monotonic observations are excluded in the component risk preferences regressions. However, these observations are retained in the indices to simplify and to avoid dropping a large number of observations from the index due to a single task. This difference accounts for the differences in the number of participants across columns.

Table A.XVI: 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 shows the effect of the night-sleep and nap treatments 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 Table IV).

- Col 1 displays the treatment effect for the whole sample as a point of reference.
- In cols 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:
    - Col 2: Sleep duration as measured by the actigraph during baseline.
    - Col 3: Sleep efficiency (sleep duration/time in bed) as measured by the actigraph during baseline.
    - Col 4: The overall index itself during baseline.
    - Col 5: Whether the participant reported napping at least once a day before the beginning of the study.
    - Col 6: Whether the participant is female.
    - Col 7: The participants' age.
- Regressions control for gender, age, and the baseline outcome variable. Standard errors in parentheses are robust to heteroscedasticity.

Table A.XVII: Effects Per Hour of Night Sleep and Naps: Instrumental Variable Estimates

	OVERALL	WORK				WELL-BEING		
	Index (1)	Earnings (2)	Productivity (3)	Labor Supply (4)	Output (5)	Index (6)	Physical (7)	Psychological (8)
Night Sleep (Hrs.)	**-0.04**	**-0.06**	0.03	-0.15	-0.08	**0.03**	0.09	-0.08
	(0.08)	(0.04)	(0.03)	(0.05)	(0.04)	(0.06)	(0.08)	(0.11)
	{0.62}	{0.15}	{0.37}	{0.00}	{0.05}	{0.58}	{0.27}	{0.44}
		[0.61]	[0.37]	[0.01]	[0.10]	[0.92]	[0.44]	[0.44]
Nap Sleep (Hrs.)	**0.47**	**-0.12**	0.17	-0.43	-0.05	**0.29**	0.19	0.43
	(0.16)	(0.09)	(0.07)	(0.12)	(0.09)	(0.12)	(0.15)	(0.21)
	{0.00}	{0.22}	{0.02}	{0.00}	{0.60}	{0.01}	{0.20}	{0.05}
		[0.22]	[0.04]	[0.00]	[0.60]	[0.05]	[0.20]	[0.09]
Unadjusted *p*-value NS vs. Nap	{0.00}	{0.58}	{0.07}	{0.01}	{0.68}	{0.04}	{0.51}	{0.02}
FWER-corrected *p*-value NS vs. Nap		[0.58]	[0.15]	[0.04]	[0.68]	[0.08]	[0.51]	[0.05]
Participants	449	449	449	449	449	450	450	450

	COGNITION			PREFERENCES			
	Index (9)	Lab Tasks (10)	Work Task (11)	Index (12)	Time (13)	Social (14)	Risk (15)
Night Sleep (Hrs.)	**0.01**	0.04	0.05	**-0.06**	0.09	-0.16	-0.06
	(0.08)	(0.08)	(0.20)	(0.11)	(0.20)	(0.12)	(0.17)
	{0.92}	{0.64}	{0.78}	{0.60}	{0.64}	{0.16}	{0.71}
	[0.92]	[0.78]	[0.78]	[0.92]	[0.71]	[0.49]	[0.71]
Nap Sleep (Hrs.)	**0.44**	0.37	0.81	**0.31**	0.58	0.09	0.28
	(0.18)	(0.16)	(0.44)	(0.16)	(0.28)	(0.23)	(0.32)
	{0.02}	{0.02}	{0.06}	{0.06}	{0.04}	{0.70}	{0.38}
	[0.05]	[0.04]	[0.06]	[0.11]	[0.11]	[0.70]	[0.70]
Unadjusted *p*-value NS vs. Nap	{0.03}	{0.04}	{0.11}	{0.04}	{0.09}	{0.28}	{0.30}
FWER-corrected *p*-value NS vs. Nap	[0.08]	[0.08]	[0.11]	[0.08]	[0.28]	[0.30]	[0.30]
Participants	450	450	429	450	450	415	415

Notes: This table estimates the treatment effect of nighttime and nap sleep using an instrumental variables specification.

- Hours of nighttime sleep and nap sleep are each instrumented simultaneously by the five binary variables indicating the treatment assignment cells (Nap Only, Devices+Encouragement Only, Devices+Incentives Only, Devices+Encouragement and Nap, and Devices+Encouragement and Nap). Regressions control for gender, age, and the baseline outcome variable.

- The outcome variables include the overall index as well as the four families of outcomes: work, well-being, cognition, and preferences. All outcome variables are the same as in Tables IV. 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. *p*-values in curly brackets are not adjusted for multiple hypothesis testing. *p*-values in square brackets are adjusted for FWER using the procedure in Hochberg (1988). The final rows show *p*-values for a test of equality between the night-sleep and nap coefficients.

Table A.XVIII: Comparing Different Multiple-Hypothesis Testing Corrections for Table IV

	Unadjusted (1)	FWER			FDR		
		WY (2)	Hochberg (3)	Holm (4)	BH (5)	BY (6)	BKY (7)
Panel A: Night-Sleep vs. Control							
Work Earnings	0.08	0.30	0.32	0.32	0.32	0.67	0.32
Well-Being	0.66	0.96	0.95	1.00	0.95	1.00	0.95
Cognition	0.93	0.96	0.95	1.00	0.95	1.00	0.95
Preferences	0.95	0.96	0.95	1.00	0.95	1.00	0.95
Panel B: Nap vs. Control							
Work Earnings	0.23	0.24	0.23	0.23	0.23	0.48	0.17
Well-Being	0.01	0.02	0.02	0.02	0.02	0.05	0.02
Cognition	0.03	0.08	0.08	0.08	0.05	0.11	0.04
Preferences	0.05	0.10	0.09	0.09	0.06	0.13	0.05
Panel C: Nap vs. Night-Sleep							
Work Earnings	0.62	0.63	0.62	0.62	0.62	1.00	0.62
Well-Being	0.14	0.30	0.30	0.48	0.20	0.42	0.20
Cognition	0.12	0.30	0.30	0.48	0.20	0.42	0.20
Preferences	0.15	0.30	0.30	0.48	0.20	0.42	0.20

Notes: This table reports adjusted p-values (or q-values), corresponding to Table IV, which correct for multiple hypothesis testing (MHT).

- In Panel A, we report the adjustments for the comparison of the night-sleep treatments with the control group.
- In Panel B, we report adjustments for the comparison of the nap treatment with the control group.
- In Panel C, we report adjustments for the comparison of the nap treatment with the night-sleep treatments.

In the rows of each panel we show, respectively, the adjusted p-values (and q-values) for the Work family outcome (Earnings), Well-Being, Cognition, and Preferences Indices. The False Discovery Rate (FDR) corrections attempt to control the share of rejections which are false positives. The corresponding q-value indicates the share of rejections with an equal or smaller unadjusted p-value which are expected to be false positives. The Family-Wise Error Rate (FWER) corrections control the probability of at least one rejection being a false positive. The adjusted p-value indicates this probability. We report the following adjustments

- Col 1, Unadjusted: p-values without correcting for MHT. The same as shown in Table IV.
- Col 2, Westfall-Young (WY): p-values using the step-down permutation-based FWER adjustment proposed by Westfall and Young (1993) as described in Anderson (2008). This approach accounts for the actual covariance structure of the data, and is our **preferred** approach, reported in our main tables.
- Col 3, Hochberg: p-values using the step-up FWER adjustment proposed by Hochberg (1988). This adjustment holds under non-negative dependence in the test statistics, and is used in the IV estimates in Table A.XVII.
- Col 4, Holm: p-values using the step-down FWER adjustment proposed by Holm (1979). This adjustment allows for arbitrary dependence across tests.
- Col 5, BH: q-values using the FDR adjustment proposed by Benjamini and Hochberg (1995). This adjustment allows for positive dependence across tests.
- Col 6, BY: q-values using the FDR adjustment proposed by Benjamini and Yekutieli (2001). This adjustment allows for arbitrary correlation across tests.
- Col 7, BKY: q-values using the FDR adjustment proposed by Benjamini, Krieger, and Yekutieli (2006). This adjustment allows for positive correlation across tests and estimates the number of rejections from data.

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.XIV, 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

---

[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.

decisions to be the one that counts. For example, if decision c was selected, the piece rate associated with 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.

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

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

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

74

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

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

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.XIX). 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.XIX: 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 cols 1 and 2 (which is equivalent to winsorizing daily net savings at Rs. 0); lateness in minutes on "short days" (i.e. the maximum between zero and arrival time - 11am) in cols 3 and 4; typing time (measured in hours) in cols 5 and 6; and hours of night sleep (measured by actigraph) in cols 7 and 8. All dependent variables are study-long averages (including the baseline period).

- Cols 1, 3, 5, and 7 have no controls. Cols 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.XIV.

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 B.

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.VII) and excluding missing observations improves statistical power without changing results qualitatively (results in the working version of this paper, (Bessone et al., 2020)).

- **Measures of labor supply.** The main measure of labor supply pre-registered and reported in the paper was time spent actively typing. We additionally pre-registered total time at office as a measure of labor supply. These two 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.

- **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 IV, A.VII, A.VIII and A.IX.

To apply these corrections, we ran simulations to control the Family-Wise Error Rate following the step-down procedure of Westfall and Young (1993), following the steps laid out in Anderson (2008). 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 account for 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 5,000 iterations according to the following sub-steps:

    - Re-randomize the treatment assignments (night sleep and nap). When randomizing, follow the same stratification and standardization procedures as in the RCT.
    - Run the core regressions relevant to the family in question. For instance, for the work-related outcomes run the main productivity, labor supply, and earnings specifications.
    - Save the z-scores computed for each regression coefficient, so the result is 5000 z-scores multiplied by the number of outcomes in our family.

3. Apply the step-down procedure of Westfall and Young (1993) to the simulated test statistics. See Anderson (2008) for more details.

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.XX, and the demographic characteristics across the first two stages can be found in Appendix Table A.XXI.

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.XXII).[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.XX: 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.
- "Percent of Last Stage" indicates the percentage of participants who advanced from the prior stage to the stage of interest. "Percent of Total" is the percentage of participants who advanced to the stage of interest divided by the total participants approached. "Frequency" is the count of participants advancing to each stage.
- "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 to distribute on that day, 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 full nights.

Table A.XXI: 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.

- Each stage of the survey is described in the notes of Table A.XX.

- With the exception of Age and Highest Grade Attended which are average values, 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 to Rs. 40,000; high income - monthly household income Rs. 40,001 or above. 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.XXII: Sleep Survey - Sleep Correlates

	Self-reported Night Sleep		Actigraph Night Sleep			Actigraph 24-Hour Sleep		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
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
	(0.10)	(0.10)		(0.12)	(0.13)		(0.12)	(0.13)
Higher Income	0.06	0.00		0.04	0.03		0.02	0.04
	(0.11)	(0.11)		(0.13)	(0.14)		(0.13)	(0.14)
Female	-0.04	-0.07		0.19**	0.18*		0.14	0.12
	(0.05)	(0.06)		(0.09)	(0.10)		(0.09)	(0.11)
Age 34 - 45	-0.54***	-0.50***		0.01	0.01		0.03	0.00
	(0.06)	(0.06)		(0.11)	(0.11)		(0.11)	(0.12)
Age 46 - 58	-0.58***	-0.52***		-0.55***	-0.56***		-0.49***	-0.53***
	(0.07)	(0.07)		(0.11)	(0.12)		(0.12)	(0.13)
Age 59 - 92	-0.45***	-0.40***		-0.06	-0.08		-0.02	-0.06
	(0.07)	(0.07)		(0.11)	(0.11)		(0.12)	(0.12)
Children (#)	-0.07***	-0.06**		-0.10**	-0.11**		-0.09*	-0.10**
	(0.03)	(0.03)		(0.05)	(0.05)		(0.05)	(0.05)
Some School		-0.11			0.26*			0.21
		(0.11)			(0.13)			(0.13)
College		0.05			0.12			0.02
		(0.12)			(0.15)			(0.16)
Employment		-0.09			-0.01			0.01
		(0.06)			(0.09)			(0.10)
Constant	6.90***	6.99***	4.98***	5.42***	5.28***	5.27***	5.69***	5.60***
	(0.12)	(0.17)	(0.07)	(0.16)	(0.20)	(0.08)	(0.16)	(0.22)
Mean of DV	6.49	6.49	5.45	5.45	5.45	5.69	5.69	5.69
N	3389	3387	1367	1367	1367	1367	1367	1367
Participants	3387	3387	439	439	439	439	439	439

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

- Cols 1 and 2 consider nighttime sleep (in hours) as self-reported by the participant. Cols 3 to 5 examine actigraph measurements of nighttime sleep (in hours). Cols 6 to 8 include total hours of sleep per 24 hour period, summing actigraph measures of night sleep and naps.

- Covariates in the rows include: (1) sleep quartiles based on the participants' self-reported average nighttime sleep during the survey period, (2) income categories derived from the surveyor's assessment of the income level of the participant's neighborhood, (3) a dummy for whether the participant is female (self-reported), (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 School", some college or more ("College"), (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 unique 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.