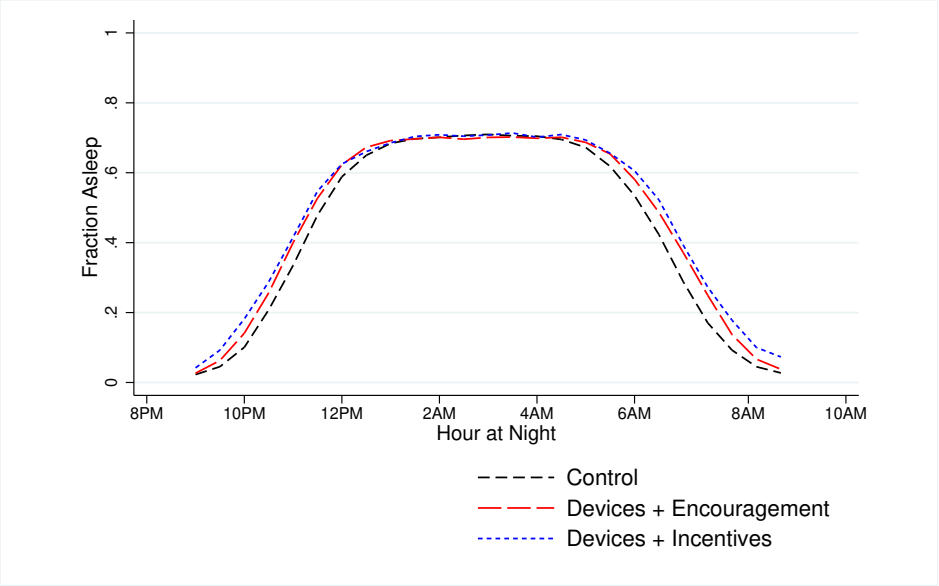
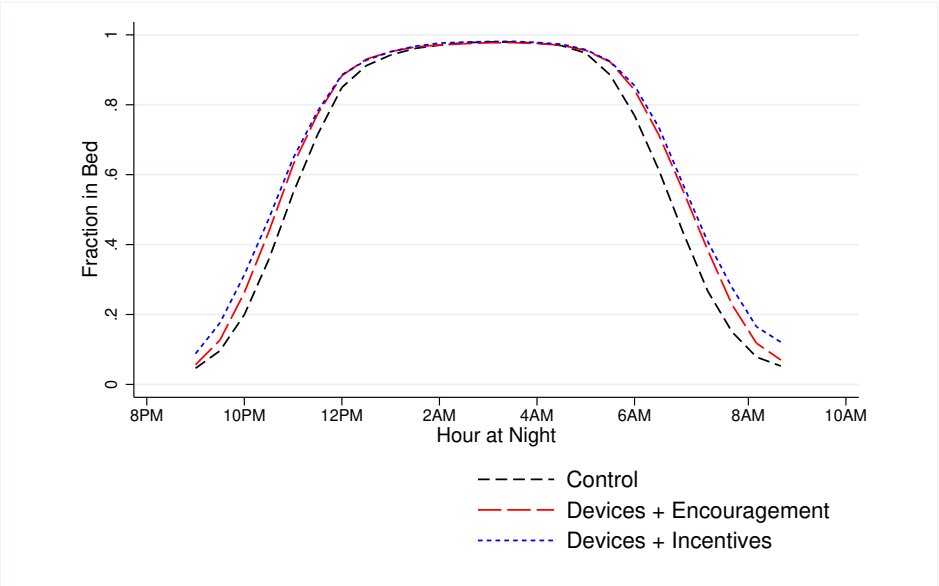
A Online Only Supplementary Figures

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.

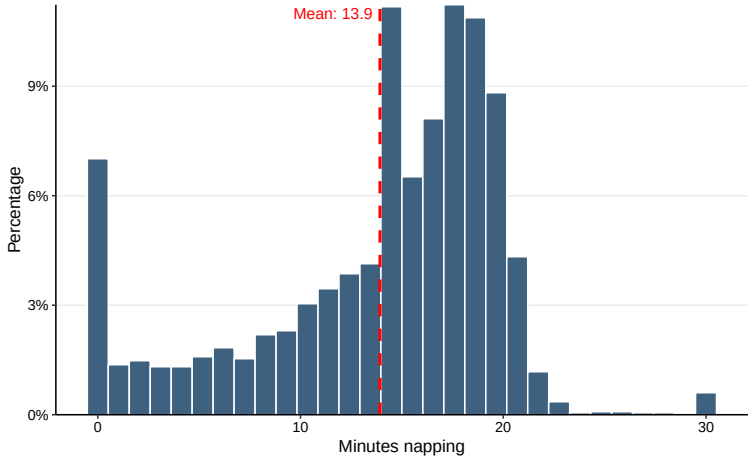
FIGURE A.II: Measuring and Increasing Sleep

(a) **Actigraph Watch**

(b) **Devices to Improve Night Sleep Environments**

(c) **Nap Station**

(d) **Distribution of Nap Duration**

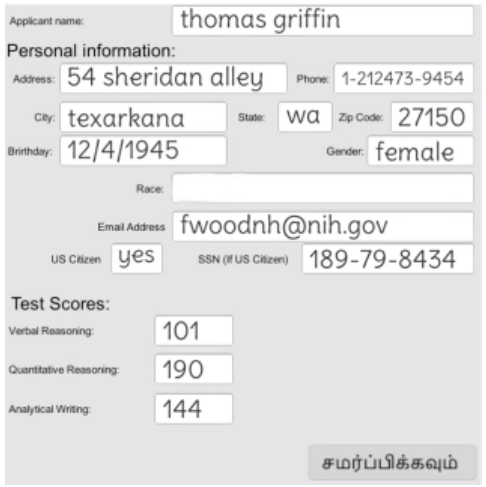


Notes: This figure illustrates devices and treatments used to measure and increase study participants' sleep.

- 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.

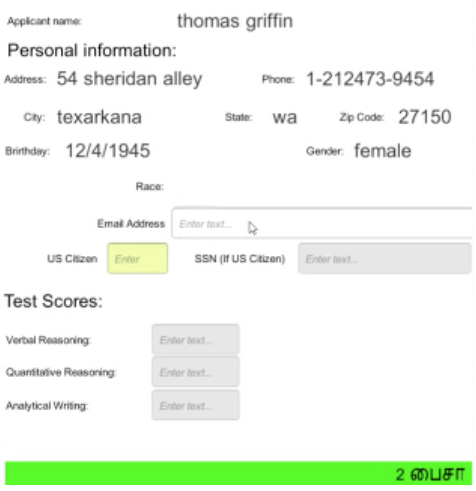
FIGURE A.III: Data-Entry Interface with Salient and Non-Salient Piece Rates

(a) **Left side**

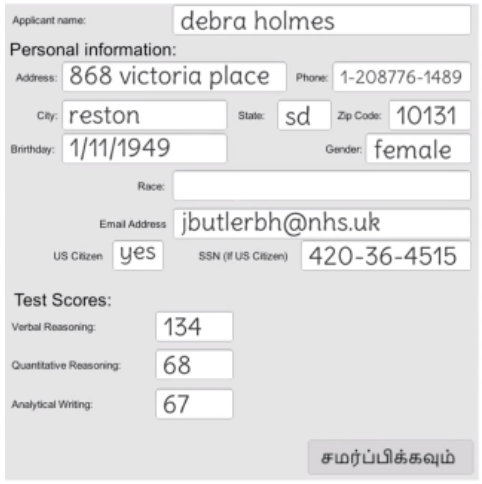


(b) **Right side (salient, low)**

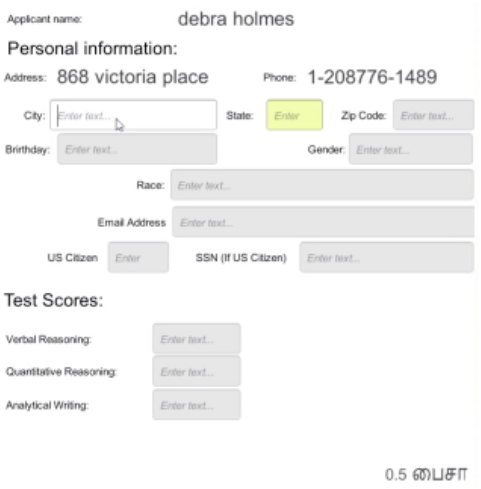
(c) **Right side (salient, high)**



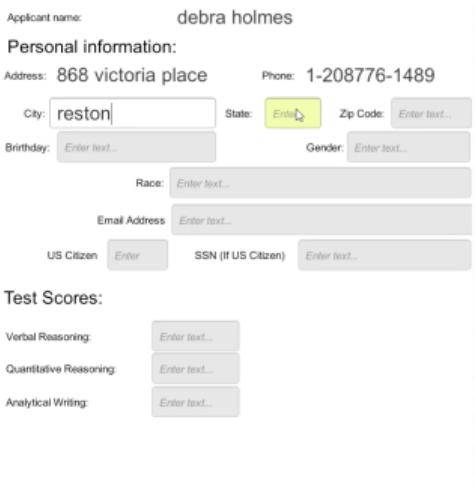
(d) **Left side**



(e) **Right side (non-salient, start)**



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

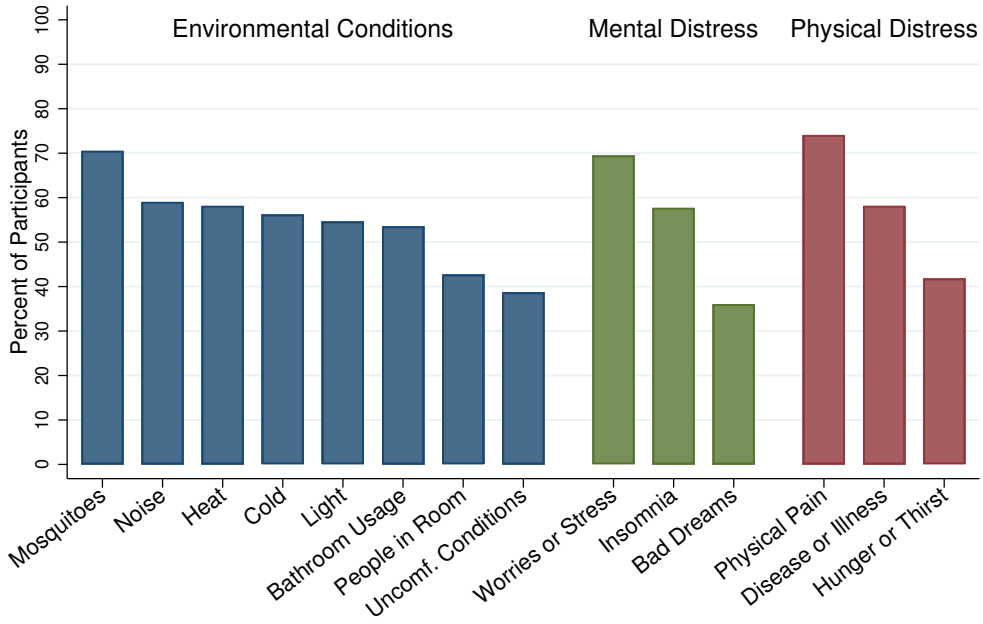


Notes: This figure shows screen shots of the data-entry task interface used by participants.

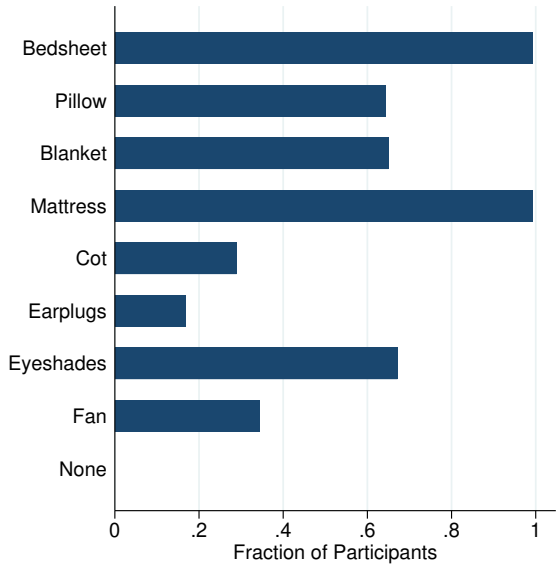
- Panels (a) and (d) show the left side of the screen, which contains the (fictional) data to be transcribed by study participants. The remaining panels show versions of the right side of the screen, where participants were supposed to enter the data.
- Panels (b) and (c) show the right side of the screen 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

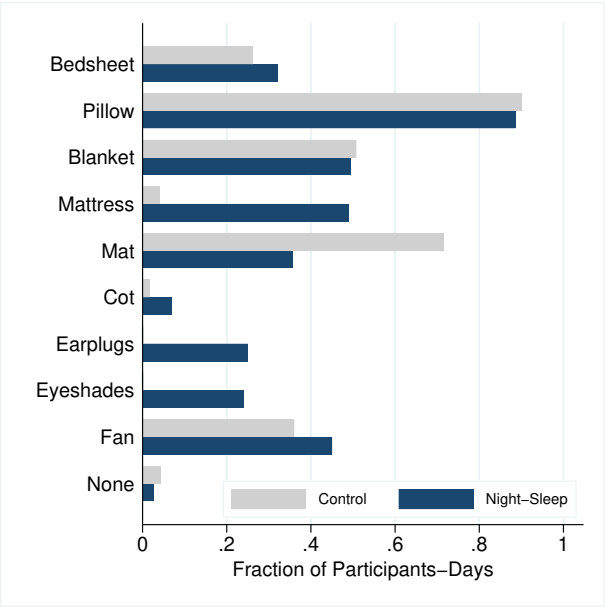
(a) **Factors Interfering with Study Participants' Sleep**



(b) **Take-up of Night-Sleep Devices**



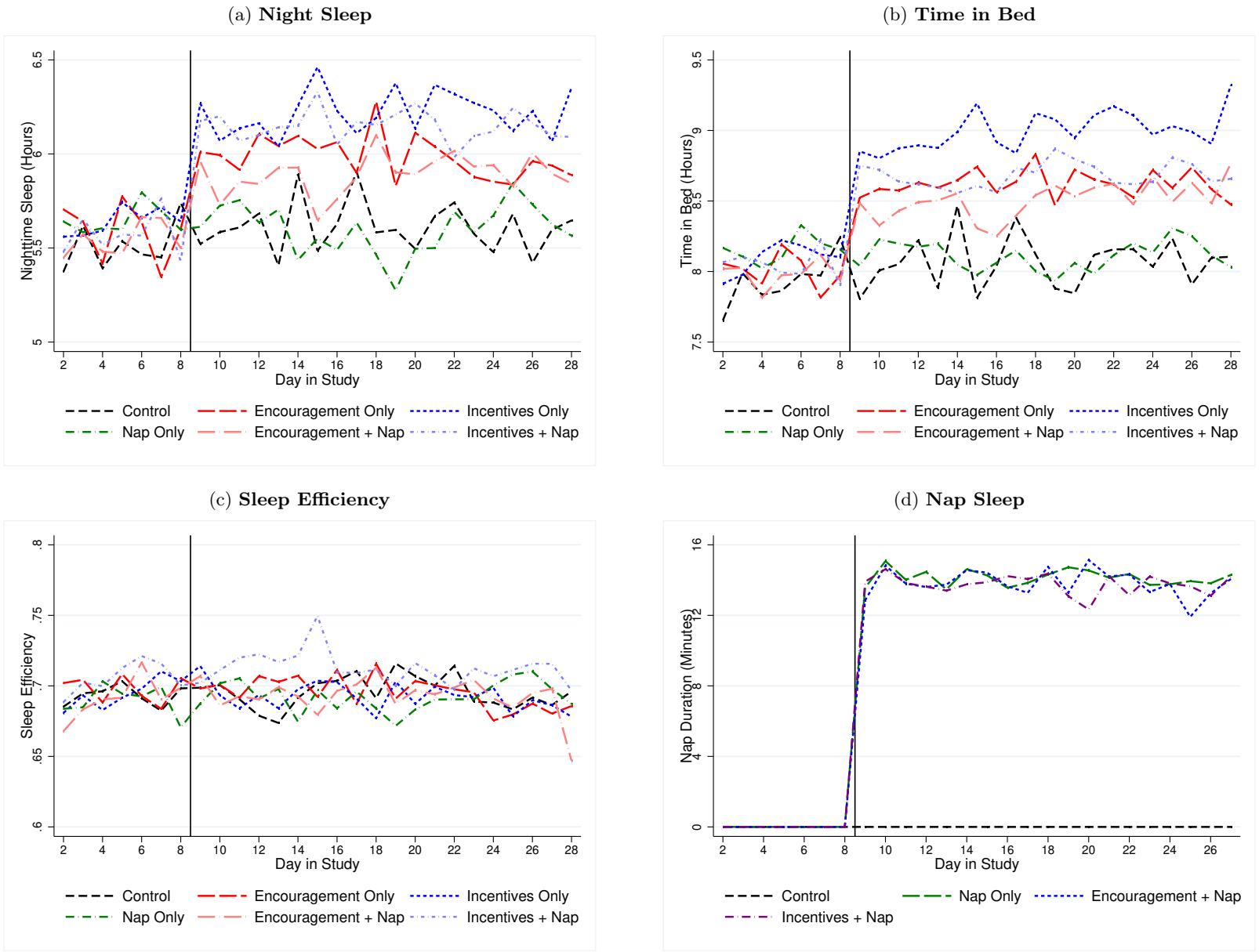
(c) **Self-reported Usage**



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.

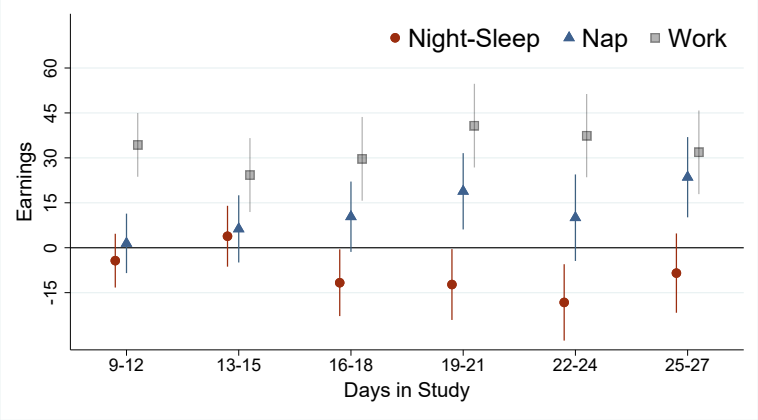
FIGURE A.V: Fully-Disaggregated Impacts on Nighttime and Nap Sleep



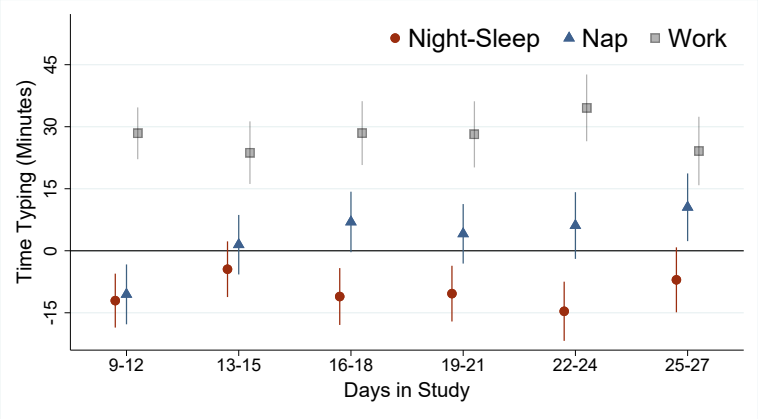
Notes: This figure shows the average of key sleep-related variables for different treatment arms by day in study of the RCT. The different panels show exactly the same variables as in Figure V but the different treatment groups are fully disaggregated. All outcomes are actigraph measures. In panels (a) 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

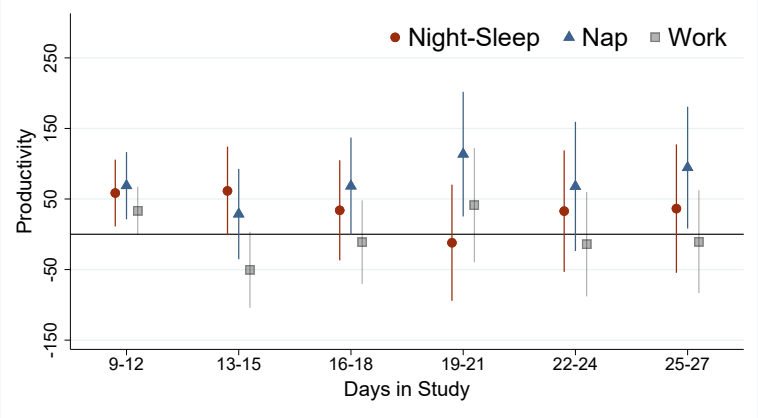
(a) **Earnings**



(b) **Hours Typing**



(c) **Productivity**

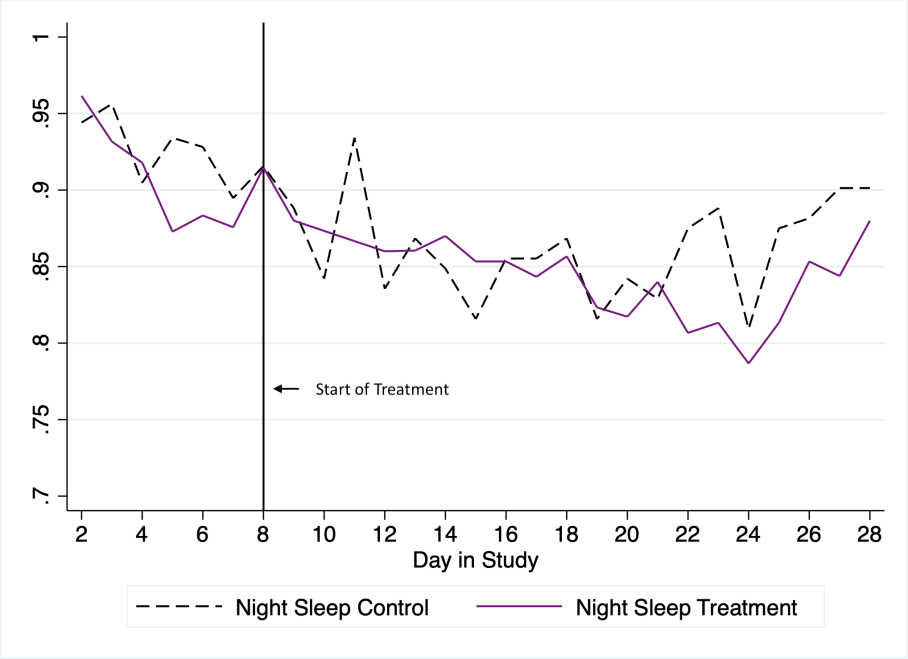


Notes: This figure shows the dynamics of the nap and night-sleep treatment effects throughout the treatment period.

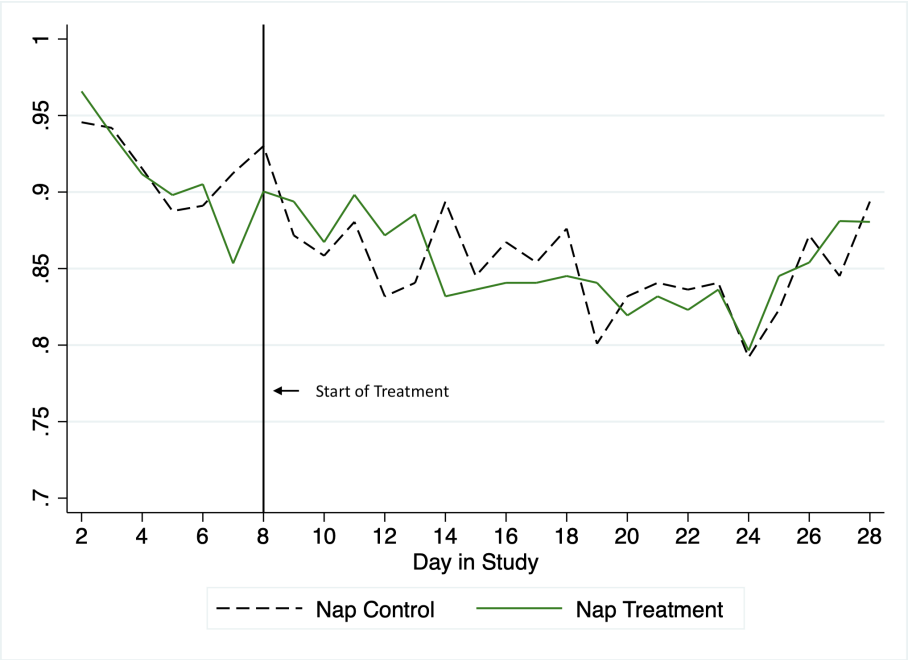
- The blue series (triangles) shows the nap treatment effect in comparison to participants on Break days (i.e., not assigned to the nap treatment and not allowed to work during the nap break that day). The grey series (squares) compares participants on Work days (i.e., participants not assigned to nap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



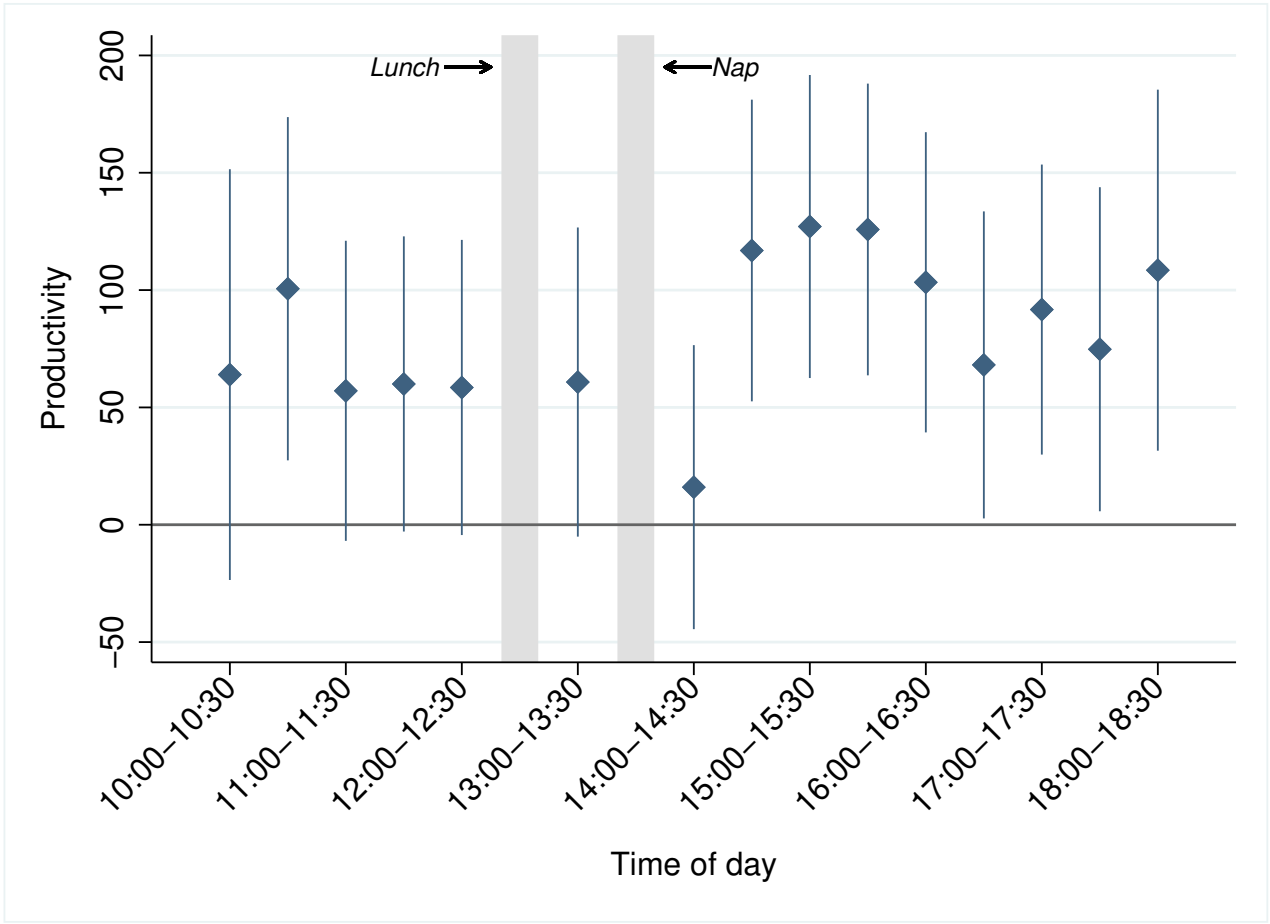
(b) **Fraction of Participants Present by Nap Group**



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

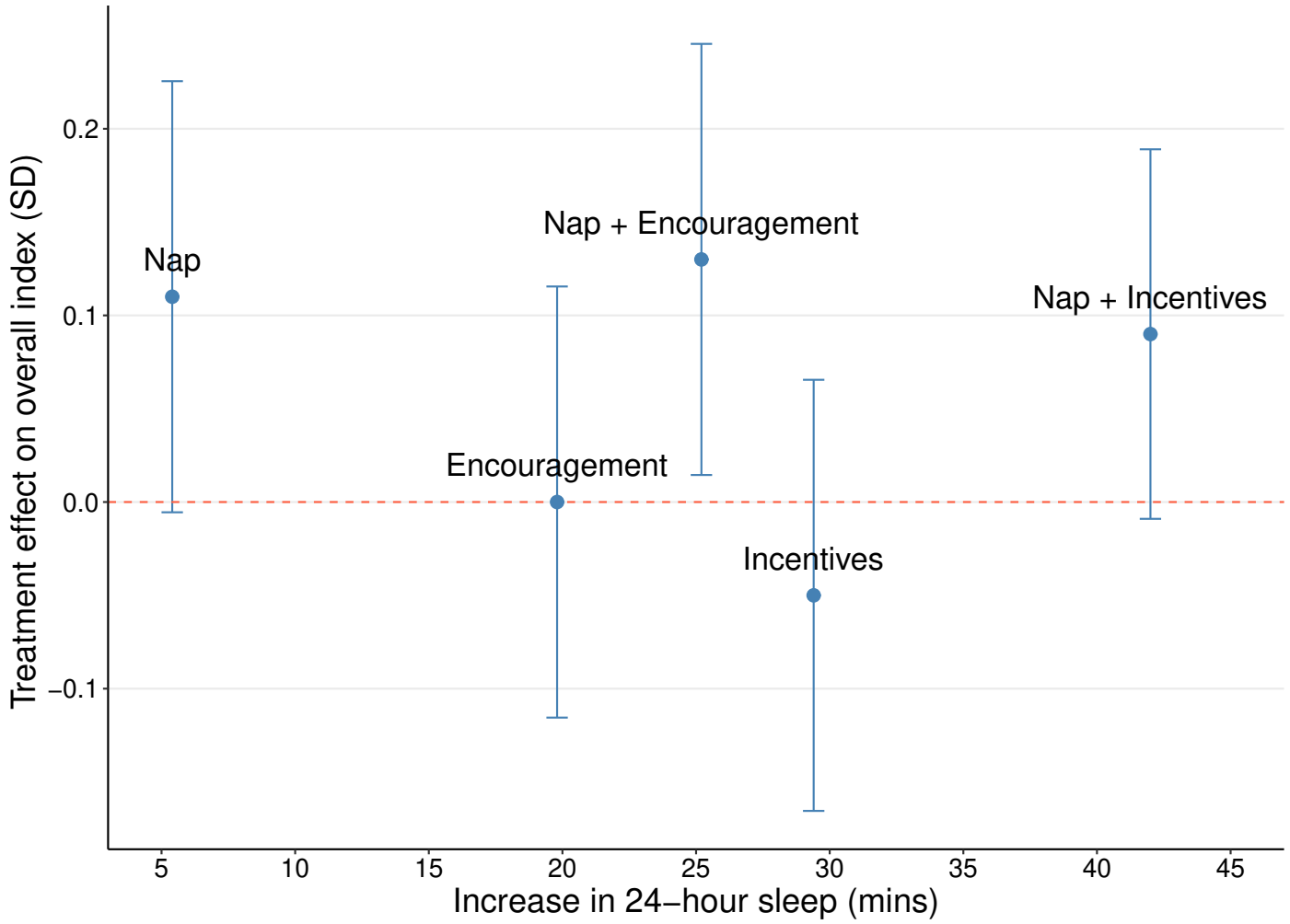
- In panel (a), the solid purple line represents participants in the two night-sleep treatment groups while the dashed black line represents the corresponding control group.
- In panel (b) the solid green line represents the nap group while the dashed black line represents the corresponding (no-nap) 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	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 C. Wellbeing, Cognition, Preferences											
Baseline Wellbeing	-0.00	0.03	0.09	-0.03	-0.01	0.02	0.67	0.19	0.70	0.93	0.72
	(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	0.49	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	Treatment	1=2	Control	Treatment	4=5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Demographics						
Female	0.66	0.66	0.91	0.65	0.66	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	Treatment	1=2	Control	Treatment	4=5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel C. Wellbeing, Cognition, and Preferences						
Baseline Wellbeing	-0.01	0.03	0.26	0.04	0.03	0.30
	(0.03)	(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**				
	(1)	X=Sleep Duration (2)	X=Sleep Efficiency (3)	X=Baseline Naps (4)	X=Female (5)	(6)	X=Sleep Duration (7)	X=Sleep Efficiency (8)	X=Baseline Naps (9)	X=Female (10)	(11)	X=Sleep Duration (12)	X=Sleep Efficiency (13)	X=Baseline Naps (14)	X=Female (15)
Night-Sleep Treatments	0.48***	0.46***	0.46***	0.38***	0.37***	-0.14	-0.18	-0.17	-1.29	-0.59					
	(0.05)	(0.08)	(0.08)	(0.12)	(0.11)	(0.42)	(0.65)	(0.65)	(0.80)	(0.79)					
X		-0.04	0.15*	-0.11	-0.11		0.15	0.13	-0.62	0.71		0.21	-0.14	-0.05	-0.34
		(0.11)	(0.09)	(0.11)	(0.10)		(0.73)	(0.81)	(0.78)	(0.82)		(0.15)	(0.14)	(0.17)	(0.20)
Night-Sleep Treatments*X		0.03	0.04	0.12	0.17		0.09	0.06	1.56	0.68					
		(0.10)	(0.10)	(0.14)	(0.12)		(0.82)	(0.86)	(0.97)	(0.95)					
Nap Treatment											13.84***	13.62***	13.34***	13.24***	13.30***
											(0.30)	(0.42)	(0.42)	(0.63)	(0.55)
Nap Treatment*X												0.46	0.98*	0.81	0.81
												(0.58)	(0.58)	(0.71)	(0.64)
DV Control Group Mean	5.60	5.60	5.60	5.60	5.60	69.83	69.83	69.83	69.83	69.83					
Participants	451	451	451	451	451	451	451	451	451	451	452	452	452	452	452

Notes: This table 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	0.48	-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]	[0.36, 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

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	**Earnings**	Productivity	Labor Supply	Output	**Index**	Physical	Mental
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Devices+Encouragement Only	**0.00**	**-0.07**	-0.02	-0.06	-0.07	**0.13**	0.16	0.02
	{1.00}	{0.03}	{0.43}	{0.10}	{0.02}	{0.01}	{0.01}	{0.79}
		[0.11]	[0.44]	[0.19]	[0.06]	[0.04]	[0.02]	[0.79]
Devices+Incentives Only	**-0.05**	**-0.07**	-0.02	-0.08	-0.08	**0.05**	0.08	-0.02
	{0.49}	{0.05}	{0.57}	{0.05}	{0.02}	{0.34}	{0.23}	{0.80}
		[0.21]	[0.57]	[0.10]	[0.05]	[0.69]	[0.39]	[0.80]
Nap Only	**0.11**	**-0.07**	-0.01	-0.07	-0.05	**0.18**	0.16	0.19
	{0.11}	{0.10}	{0.65}	{0.07}	{0.18}	{0.00}	{0.01}	{0.05}
		[0.29]	[0.66]	[0.17]	[0.28]	[0.01]	[0.02]	[0.05]
Devices+Encouragement and Nap	**0.13**	**-0.07**	0.04	-0.15	-0.04	**0.13**	0.11	0.11
	{0.05}	{0.07}	{0.15}	{0.00}	{0.19}	{0.01}	{0.07}	{0.21}
		[0.22]	[0.20]	[0.00]	[0.20]	[0.03]	[0.10]	[0.22]
Devices+Incentives and Nap	**0.08**	**-0.07**	0.04	-0.17	-0.07	**0.10**	0.11	0.07
	{0.19}	{0.05}	{0.10}	{0.00}	{0.04}	{0.04}	{0.09}	{0.46}
		[0.11]	[0.12]	[0.00]	[0.08]	[0.11]	[0.12]	[0.45]
Participants	451	451	451	451	451	452	452	452

	COGNITION			**PREFERENCES**			
	Index	Lab Tasks	Work Task	**Index**	Time	Social	Risk
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Devices+Encouragement Only	**-0.00**	-0.00	0.04	**-0.09**	-0.05	-0.10	-0.15
	{0.95}	{0.98}	{0.82}	{0.16}	{0.68}	{0.27}	{0.24}
	[0.96]	[0.98]	[0.97]	[0.31]	[0.69]	[0.48]	[0.48]
Devices+Incentives Only	**-0.03**	0.04	-0.12	**-0.04**	0.11	-0.16	-0.06
	{0.74}	{0.54}	{0.54}	{0.52}	{0.36}	{0.05}	{0.68}
	[0.75]	[0.54]	[0.54]	[0.75]	[0.59]	[0.17]	[0.68]
Nap Only	**0.09**	0.07	0.15	**-0.01**	0.12	-0.09	-0.01
	{0.23}	{0.30}	{0.37}	{0.85}	{0.34}	{0.33}	{0.92}
	[0.41]	[0.38]	[0.38]	[0.85]	[0.56]	[0.56]	[0.92]
Devices+Encouragement and Nap	**0.05**	0.13	0.03	**0.09**	0.23	0.03	0.03
	{0.54}	{0.08}	{0.88}	{0.17}	{0.05}	{0.79}	{0.83}
	[0.55]	[0.17]	[0.88]	[0.33]	[0.16]	[0.84]	[0.84]
Devices+Incentives and Nap	**0.14**	0.09	0.32	**0.01**	0.11	-0.10	0.01
	{0.04}	{0.18}	{0.04}	{0.87}	{0.38}	{0.28}	{0.96}
	[0.11]	[0.19]	[0.08]	[0.86]	[0.62]	[0.62]	[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	**Earnings**	Productivity	Labor Supply	Output	**Index**	Physical	Mental
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Night-Sleep Only	**-0.02**	**-0.07**	-0.02	-0.07	-0.08	**0.09**	0.12	0.00
	(0.06)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)	(0.06)	(0.08)
	{0.69}	{0.02} [0.10]	{0.43} [0.44]	{0.04} [0.09]	{0.01} [0.02]	{0.05} [0.13]	{0.03} [0.07]	{0.99} [0.99]
Nap Only	**0.11**	**-0.07**	-0.01	-0.07	-0.05	**0.18**	0.16	0.19
	(0.07)	(0.04)	(0.03)	(0.04)	(0.04)	(0.06)	(0.07)	(0.10)
	{0.11}	{0.10} [0.28]	{0.65} [0.66]	{0.07} [0.16]	{0.17} [0.28]	{0.00} [0.01]	{0.01} [0.03]	{0.05} [0.05]
Night-Sleep and Nap	**0.11**	**-0.07**	0.04	-0.16	-0.05	**0.12**	0.11	0.09
	(0.06)	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)	(0.06)	(0.08)
	{0.06}	{0.04} [0.11]	{0.08} [0.08]	{0.00} [0.00]	{0.06} [0.08]	{0.01} [0.03]	{0.04} [0.09]	{0.25} [0.25]
Participants	451	451	451	451	451	452	452	452
p-values NS vs. Nap	{0.03}	{0.79} [0.79]	{0.87} [0.93]	{0.94} [0.93]	{0.38} [0.69]	{0.07} [0.25]	{0.47} [0.50]	{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} [0.30]	{0.41} [0.67]	{0.84} [0.86]	{0.16} [0.31]
p-values Nap vs. Both	{0.98}	{0.88} [0.91]	{0.04} [0.08]	{0.01} [0.01]	{0.88} [0.88]	{0.21} [0.61]	{0.37} [0.40]	{0.23} [0.40]

	COGNITION			**PREFERENCES**			
	Index	Lab Tasks	Work Task	**Index**	Time	Social	Risk
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Night-Sleep Only	**-0.02**	0.02	-0.04	**-0.07**	0.03	-0.13	-0.10
	(0.06)	(0.06)	(0.15)	(0.06)	(0.11)	(0.08)	(0.11)
	{0.81} [0.81]	{0.75} [0.79]	{0.79} [0.79]	{0.24} [0.42]	{0.80} [0.79]	{0.08} [0.23]	{0.37} [0.60]
Nap Only	**0.09**	0.07	0.15	**-0.01**	0.12	-0.09	-0.01
	(0.07)	(0.07)	(0.17)	(0.07)	(0.12)	(0.09)	(0.14)
	{0.23} [0.40]	{0.30} [0.36]	{0.36} [0.36]	{0.86} [0.86]	{0.34} [0.57]	{0.33} [0.57]	{0.92} [0.91]
Night-Sleep and Nap	**0.09**	0.11	0.18	**0.05**	0.17	-0.04	0.02
	(0.06)	(0.06)	(0.14)	(0.06)	(0.11)	(0.08)	(0.11)
	{0.13} [0.24]	{0.08} [0.15]	{0.22} [0.22]	{0.38} [0.38]	{0.12} [0.34]	{0.67} [0.89]	{0.89} [0.89]
Participants	452	452	429	452	452	415	415
p-values NS vs. Nap	{0.13} [0.36]	{0.35} [0.35]	{0.23} [0.35]	{0.32} [0.54]	{0.33} [0.55]	{0.55} [0.55]	{0.45} [0.55]
p-values NS vs. Both	{0.06} [0.16]	{0.06} [0.10]	{0.11} [0.12]	{0.01} [0.06]	{0.06} [0.16]	{0.15} [0.21]	{0.20} [0.21]
p-values Nap vs. Both	{0.91} [0.91]	{0.53} [0.78]	{0.86} [0.85]	{0.26} [0.61]	{0.58} [0.80]	{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	**Earnings**	Productivity	Labor Supply	Output	**Index**	Physical	Psychological
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(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}	{0.69}	{0.35}	{0.36}
		[0.30]	[0.32]	[0.00]	[0.09]	[0.97]	[0.36]	[0.36]
Nap Treatment vs. Pooled Control	**0.12**	**-0.02**	0.04	-0.09	-0.01	**0.08**	0.05	0.12
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.05)
	{0.00}	{0.23}	{0.03}	{0.00}	{0.77}	{0.01}	{0.18}	{0.02}
		[0.25]	[0.06]	[0.00]	[0.77]	[]	[]	[]
Nap Treatment vs. Break Day	**0.15**	**0.05**	0.04	0.01	0.05			
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)			
	{0.00}	{0.02}	{0.03}	{0.60}	{0.01}			
		[0.05]	[0.05]	[0.61]	[0.02]			
Nap Treatment vs. Work Day	**0.08**	**-0.10**	0.03	-0.20	-0.07			
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)			
	{0.03}	{0.00}	{0.04}	{0.00}	{0.00}			
		[0.00]	[0.05]	[0.00]	[0.00]			
Participants	451	451	451	451	451	452	452	452

	COGNITION			PREFERENCES			
	Index	Lab Tasks	Work Task	**Index**	Time	Social	Risk
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
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}
	[0.98]	[0.76]	[0.95]	[0.98]	[0.64]	[0.64]	[0.64]
Nap Treatment vs. Pooled Control	**0.10**	0.08	0.20	**0.07**	0.13	0.03	0.07
	(0.05)	(0.04)	(0.11)	(0.04)	(0.06)	(0.05)	(0.08)
	{0.03}	{0.03}	{0.07}	{0.05}	{0.04}	{0.52}	{0.33}
	[0.08]	[0.07]	[0.07]	[0.15]	[0.20]	[0.51]	[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 $(\cdot)$, 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.X: Impacts on Time Allocation

	Labor Supply				**Sleep**	
	Time at Office	Arrival Time	Leave Time	Work Breaks	Get Up Time	Bed Time
	(1)	(2)	(3)	(4)	(5)	(6)
Night-Sleep Treatments	-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 get out of bed in the morning, measured by the actigraph; (6) time the participants go to bed at night, measured by the actigraph. 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	Happiness	Life Possibility	Life Satisfaction	Stress
	(1)	(2)	(3)	(4)	(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	Illness	Pain	Daily Act.	BP
	(6)	(7)	(8)	(9)	(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	Minutes	Output	Minutes	Output	Minutes
	(1)	(2)	(3)	(4)	(5)	(6)
Night-Sleep Treatments	0.85	0.80	0.83	0.94	0.85	0.80
	(0.03)	(0.13)	(0.05)	(0.54)	(0.04)	(0.12)
Nap Treatment	0.94	0.99	0.85	0.84	0.97	0.96
	(0.04)	(0.16)	(0.05)	(0.61)	(0.06)	(0.14)
Control	0.84	0.77	0.80	0.65	0.86	0.75
	(0.04)	(0.15)	(0.05)	(0.38)	(0.06)	(0.13)
p-value 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	**Earnings**	Productivity	Labor Supply	Output	**Index**	Physical	Mental
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Night Sleep (Hrs.)	**-0.04**	**-0.06**	0.03	-0.15	-0.08	**0.03**	0.08	-0.08
	(0.08)	(0.04)	(0.03)	(0.05)	(0.04)	(0.06)	(0.08)	(0.11)
	{0.59}	{0.15}	{0.37}	{0.00}	{0.05}	{0.66}	{0.31}	{0.44}
		[0.61]	[0.37]	[0.01]	[0.10]	[0.92]	[0.44]	[0.44]
Nap Sleep (Hrs.)	**0.46**	**-0.12**	0.17	-0.43	-0.05	**0.28**	0.16	0.43
	(0.16)	(0.09)	(0.07)	(0.12)	(0.09)	(0.12)	(0.15)	(0.21)
	{0.01}	{0.22}	{0.02}	{0.00}	{0.60}	{0.02}	{0.29}	{0.05}
		[0.22]	[0.04]	[0.00]	[0.60]	[0.07]	[0.29]	[0.09]
Unadjusted p-value NS vs. Nap	{0.00}	{0.58}	{0.07}	{0.01}	{0.68}	{0.05}	{0.62}	{0.02}
FWER-corrected p-value NS vs. Nap		[0.58]	[0.15]	[0.04]	[0.68]	[0.11]	[0.62]	[0.05]
Participants	449	449	449	449	449	450	450	450

	COGNITION			**PREFERENCES**			
	Index	Lab Tasks	Work Task	**Index**	Time	Social	Risk
	(9)	(10)	(11)	(12)	(13)	(14)	(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 Hommel (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

		FWER			**FDR**		
	Unadjusted	WY	Hochberg	Holm	BH	BY	BKY
	(1)	(2)	(3)	(4)	(5)	(6)	(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	0.69	0.97	0.95	1.00	0.95	1.00	0.95
Cognition	0.93	0.98	0.95	1.00	0.95	1.00	0.95
Preferences	0.95	0.98	0.95	1.00	0.95	1.00	0.95
Panel B: Nap vs. Control							
Work Earnings	0.23	0.25	0.23	0.23	0.23	0.48	0.17
Well-Being	0.01	0.03	0.04	0.04	0.04	0.07	0.03
Cognition	0.03	0.08	0.08	0.08	0.05	0.11	0.04
Preferences	0.05	0.15	0.09	0.09	0.06	0.13	0.05
Panel C: Nap vs. Night-Sleep							
Work Earnings	0.62	0.62	0.62	0.62	0.62	1.00	0.62
Well-Being	0.16	0.41	0.32	0.46	0.22	0.45	0.22
Cognition	0.12	0.41	0.32	0.46	0.22	0.45	0.22
Preferences	0.15	0.41	0.32	0.46	0.22	0.45	0.22

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

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

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

Hommel (1988)

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 10,000 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.

Hommel, G, "A stagewise rejective multiple test procedure based on a modified Bonferroni test," *Biometrika*, 1988, *75*, 383–386.

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