IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17618

Rest Assured. The Effects of Sleep on Labor Productivity

Marco Bertoni
Francesca Meli
Lorenzo Rocco

JANUARY 2025

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17618

Rest Assured. The Effects of Sleep on Labor Productivity

Marco Bertoni
University of Padova and IZA

Francesca Meli
University of Padova

Lorenzo Rocco
University of Padova and IZA.

JANUARY 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Rest Assured. The Effects of Sleep on Labor Productivity*

We estimate the effect of sleep on labor productivity addressing the two main challenges in time use research: the unavoidable substitutions among activities implied by the time budget constraint and the endogeneity of the allocation of time. We use complete time diary data to identify the relative effect of sleep vs. non-work activities among employees working the same number of hours, and account for the endogeneity of time use choices by leveraging longitudinal information on productivity in a value-added specification. We show that, when work hours are held constant, substituting sleep with other non-work activities does not affect labor productivity.

JEL Classification: J22, J24

Keywords: sleep, time use, productivity

Corresponding author:
Lorenzo Rocco
Department of Economics and Management "Marco Fanno"
University of Padova
Via del Santo 33
35123 Padova
Italy
E-mail: lorenzo.rocco@unipd.it

* The authors thank Osea Giuntella, Timothy Neal, and audiences at seminars in Alicante, and Padova for useful feedback. They declare that they have no material or financial interest that relate to the research presented in this paper. The paper uses publicly available data from the American Time Use Survey and the Current Population Survey. See link for information on data access: https://www.bls.gov/tus/data.htm.

1 Introduction

Sleep is a biological need. Its importance has been extensively proved by a breadth of medical research reporting large impacts of sleep deprivation on physical health, mental health conditions, as well as cognitive functioning [Cappuccio et al., 2010, Ferrie et al., 2011, Hale et al., 2020].

Consistently, a recent stream of economic research shows that dedicating insufficient time to sleep, at night or in afternoon naps, hampers financial market returns [Kamstra et al., 2000], academic performance [Giuntella et al., 2024, Jagnani, 2024], and labor market outcomes [Gibson and Shrader, 2018, Giuntella and Mazzonna, 2019, Costa-Font and Fleche, 2020, Bessone et al., 2021, Kajitani, 2021, Costa-Font et al., 2024]. This body of research acknowledges that sleep duration is a choice that is affected by economic incentives [Biddle and Hamermesh, 1990], and attempts to address the resulting endogeneity concerns. Common empirical strategies use the variation in sleep duration originating from changes in natural light at time-zone boundaries, in sunset timing, or during daylight-saving time days; the presence of babies disrupting mothers' night sleep; randomly assigned information, encouragements, reminders, and monetary incentives to favor sleep.

Nevertheless, even if sleep time was randomly allocated among individuals, identification is made difficult by the presence of a binding time budget constraint: sleeping more hours necessarily comes at the expense of hours dedicated to alternative uses of time. Due to this inescapable fact, two people with different sleeping hours cannot spend the same amount of time in all other activities. Hence, the observed effect of one more hour of sleep on economic outcomes may be partly or fully due to the shorter time devoted to work, leisure, or other uses of time, a simultaneity concern that complicates interpretation. The conclusion is that whether changing sleep duration *per se* impacts economic outcomes - the parameter of interest of the body of research presented above - is a Fundamentally Unidentified Question [Angrist and Pischke, 2009].

What can be identified, at best, are relative effects of the substitution in the allocation of time among pairs of activities [Fiorini and Keane, 2014, Keane et al., 2022], and there are as many relative effects as the alternative activities that can be considered.[1]

Ignoring this matter has important consequences. So long as different substitute activities impact individual outcomes differently, whether increasing sleep is beneficial or not depends on what is actually replaced. For example, consider labor productivity, and suppose that working longer is more detrimental to productivity than sleep, while other non-work activities are more beneficial than sleep. Then, the

[1]Complementarity across time uses is also possible. For example, the "productive sleep hypothesis" [Gibson and Shrader, 2018] suggests a synergy between sleep and work.

effect of increasing sleep will be positive if one works less to sleep more, and negative if he or she reduces non-work activity.

Empirically, conditional on data availability, a way to isolate the impact of a specific time trade-off is to control for time spent in all the activities but one. This enables to estimate the effect of an increase in the time allocated in the activity of interest relative to the omitted time use category, which necessarily decreases, holding all the other activities fixed.

However, this approach introduces a second empirical complication. Time allocation is an individual choice which responds to many factors, observed and unobserved. The need to control for the complete set of activities extends endogeneity concerns to each time use, and identification of the impacts of the trade-offs of multiple uses of time would require exploiting different and independent sources of exogenous variation, one for each alternative activity. Furthermore, the sources of exogenous variation in sleep duration used in the literature would be hardly useful for this purpose. For example, Nguyen et al. [2024] shows that daylight duration generates changes in the allocation of time of children and adolescents across multiple activities that are substitutes to sleep. This violates the exclusion restriction necessary to isolate the impact of sleep against a *specific* alternative use of time.

Fortunately, the empirical research on the impact of time use has acknowledged these issues and proposed solutions to jointly address both challenges, in order to claim a causal interpretation of relative time use effects [Fiorini and Keane, 2014, Caetano et al., 2019, Keane et al., 2022, Caetano et al., 2024]. These methods generally combine selection on observables assumptions with longitudinal information on productivity, which permits the estimation of value-added (VA) models.[2] Specifically, our VA specification compares the productivity *between* workers sharing a comprehensive set of individual observable characteristics, reporting to work for the same number of hours, and having the same lagged labor productivity, but whose sleeping duration (and necessarily the time devoted to other non-work-related activities) differ. Under a set of assumptions, controlling for lagged productivity serves as a sufficient statistic for all prior inputs in the individual production function [Todd and Wolpin, 2003]. To deal with the potential endogeneity of lagged productivity, we also provide IV estimates of the VA model, using twice-lagged expected labor earnings as instrumental variable.

We find that trading off sleep with other non-work related activities does not exert any economically relevant effect on earnings. Our results are robust to a battery of specification tests, and we find little evidence of non-linear effects or heterogeneity with respect to several individual characteristics.

[2]Alternative methods exploit bunching of the time use of interest at zero [Caetano et al., 2019, 2024], which is not a feasible application in the estimation of sleep effects.

This paper's contribution is threefold. First, it brings for the first time the techniques developed in the literature on the effects of time use into the literature on the economic impacts of sleep. Second, by using complete time diary data of full-time workers from a representative sample of the US population, it estimates the effect of sleep vs. non-work activities on individual labor productivity – proxied by weekly labor earnings - keeping constant the hours of work. Third, by holding work time fixed, it isolates the productivty component from the labor supply component in labor earnings.[3] To the best of our knowledge, this is the first study to estimate the impact of sleep on individual productivity while addressing these challenges.

The rest of the paper unfolds as follows. Section 2 describes the data. Section 3 discusses the estimation strategy, going more in detail through the involved and how we address them. Section 4 collects baseline results, specification tests and robustness checks. Finally, Section 5 concludes.

2 Data

2.1 Relevant datasets

We use data from the American Time Use Survey (ATUS) linked with the Current Population Survey (CPS).

Since 2003, ATUS collects information on how Americans allocate their time. Every month, around 60,000 households that have terminated their CPS interviews are selected to preserve national representativeness; from each, one family member aged 15+ is randomly selected to complete a time diary for a 24-hour period. Selected respondents indicate the activities they engage in, how much time they spend in each, and other details.[4] By design, interviews are evenly administered across months and weeks of the year, but with a considerable weekend oversampling.[5] Selected respondents complete the time-use survey only once - thus, an individual time-use panel is not available.

[3] Weekly earnings mechanically decrease if work hours decline, for a given wage rate and productivity. Hence, if additional sleeping was achieved by reducing work hours, weekly earnings could decline even if the effect of more sleep was that of increasing worker productivity. Holding constant hours of work addresses this concern.

[4] Activities' duration is precisely calculated in minutes from freely indicated starting and ending times (i.e., without fixed time slots). Activities are hierarchically classified: the main, and highest, level of aggregation of the coding-structure consists of 17 groups, further disaggregated in an intermediate and final level of 100 and more than 400 groups, respectively. Such features allows for a rather precise characterisation of individual time allocation. Other details include where the activity took place, who, if anyone, respondents were with when engaging in the reported activities.

[5] Days from Monday to Friday represent 10% of weekly diaries each, while Saturdays and Sundays 25% each, leading to a 50-50% partitioning of interviews between workdays and the weekend. This 50-50% allocation is functional to the purpose of showing differences in time allocation between the two parts of the week.

The linkage between ATUS and CPS allows us to gain access to a large set of information on respondents' demographics and socio-economic status, their employment status, job characteristics, earnings and income. The CPS sample structure is characterized by a rotation pattern: each household is in the sample for four successive months, then out for eight months, and then in the sample again for the final four months. During the "months-in-sample" (MIS), household members complete the "basic monthly survey", providing the main information. Importantly, in the fourth and the eighth MIS households participate in the "earner study", where adult individuals who are not self-employed are asked to report information on usual weekly earnings, that constitute the basis for our measurement of labor productivity.

2.2 Variables definition

2.2.1 Individual labor productivity

Following Gibson and Shrader [2018] and Costa-Font et al. [2024], we capture labor productivity – our outcome variable - using usual weekly earnings as recorded in ATUS.[6] This is defined for all employed respondents who have positive labor income and are not self-employed. They are asked to report usual weekly earnings at the main job, including overtime pay, commissions, and tips usually received, before taxes and other deductions. We also recover respondents' past weekly earnings value from their first CPS earner study (i.e., during their MIS4), occurring 14 to 17 months before the time-use survey. This lagged earnings measure will represent the lagged outcome in the production function.

2.2.2 Time use variables

Given our focus on labor productivity, we only consider time allocation during the workweek, i.e., days from Monday to Friday, when most workers are at work. Activities in ATUS are recorded in minutes per day, which we transform into hours per workweek. This allows to match the frequency of the outcome and ease interpretation, in line with the literature [Gibson and Shrader, 2018, Kajitani, 2021, Costa-Font et al., 2024].

To minimize arbitrariness in the time use classification, we only define three main activities: sleep, work and other non-work-related activities. Sleep accounts for total sleep duration, including nighttime sleep and naps. Work duration is defined as total

[6]Given the short time span between the CPS MIS8 and the ATUS interviews (2 to 5 months), earnings in ATUS differ from earnings in CPS MIS8 only for 40% of respondents who experienced a job change in between. For the others, ATUS earnings coincide with CPS MIS8 earnings.

time spent in all work and work-related activities, including commuting.[7] Finally, the remaining time is attributed to a residual category including all other non-work (and non-sleep) activities.[8] This third category includes the total time devoted to personal and others' care, housework and any type of leisure.

2.2.3 Control variables

Our models also include a rich set of control variables that could influence labor productivity and possibly correlate with time use. This set of covariates includes demographic and socioeconomic characteristics (gender, age, race, marital status, number of people in the household, presence of children, education level, primary occupation); location characteristics (state and whether the place of residence is a metropolitan area); attributes of the day of the interview (year, month, day of the week); and the precise number of minutes of daylight during the day of the interview.[9]

2.3 Sample selection

To obtain a suitable sample for the analysis we apply several selection criteria, that are described in detail in Appendix A.

Our final estimation sample counts 12,398 observations for full-time employees interviewed between 2003 and 2019, reporting a workday time diary, with non-missing current and lagged earnings records, who did not change job between CPS MIS4 and ATUS interviews, and whose sleep time, work time, and earnings variables are not outliers in sample distributions. As in Gibson and Shrader [2018], we limit the sample to employees because earnings are not reported in ATUS by the self-employed,

[7] We depart from the BLS high-level classification of "Work & Work-Related Activities" (group 05) in two directions. First, we only include its sub-groups "Working" (0501), "Work-Related Activities" (0502) and "Work and Work-Related Activities, n.e.c" (0599), while we exclude "Other Income-Generating Activities" (0503) and "Job Search and Interviewing" (0504). This is because, as they are conceived in ATUS, "Other Income-Generating Activities" indicate activities which are not part of the main job, but rather conducted in parallel to it or under informal agreement (e.g., renting an accommodation, selling handcrafts or babysitting), whereas "Job Search and Interviewing" is not part of the regular job, but undertaken to change it. Since the outcome represents earnings at the main job, we deem it is more appropriate to separate them from the other sub-groups. Second, although "Travelling" is a distinct main group (group 18), it includes a subgroup "Travel Related to Work" (1805), which we consider more a work-related activity rather than a form of leisure, housework or any activity other than work. Thus, we code it as work.

[8] Since it is possible that total recorded time in the time-diary does not sum exactly to 1440 minutes, we make sure to cover all the time in a day and attribute the unrecorded time to the residual category non-work activities. Hence, we assume that bed and wake time are correctly recorded as well as all sleep and work spells.

[9] To obtain sunlight duration we used Gibson and Shrader [2018] replication files and Meeus [1991] astronomical calculator. Thanks to FIPS code information reported in ATUS, individuals can be geocoded up to the county level. The coordinates of the location centroids in combination with information on the day of the interview were used as inputs in the algorithm to obtain sunrise time, sunset time and daylight duration in the location and day of the interview for each individual.

and to those working full-time because ATUS only collects a single time diary per respondent, and the labor supply of those with vertical part-time contracts could be misclassified if interviewed on a convenient day. Moreover, the number of males working part-time is very small.

Table 1 reports summary statistics for earnings, time use, and individual characteristics in the pooled estimation sample and by gender, while comparable estimates for the initial sample are in Table B1.[10] We see that males, white and educated workers are over-represented with respect to the initial sample as a result of the selection of full-time employees with a continuous earnings history. Consistently, individuals in our estimation sample also have higher earnings, work more hours and sleep less.

3 Methods

This Section presents the parameters of interest and the methodological challenges put forward in Section 1, that motivate our empirical strategy.

3.1 Total vs. partial sleep effects and the role of the time budget constraint

To fix ideas, suppose that productivity is a function of all possible time uses $k = 1, \ldots, K$

$$Y_i = f(T_{i1}, \ldots, T_{iK}) \tag{1}$$

and let $k = 1$ denote sleep. The marginal effect of sleep on productivity, the parameter we are after, is $f_1(.)$, the partial derivative of Y_i with respect to T_1. The time budget constraint implies that $\sum_k T_k = \bar{T}$, whose differentiation yields $\sum_k dT_k = 0$. Without loss of generality, let the substitution patterns between sleep and any time use k be denoted as $dT_k = \alpha_k dT_1$, with no restrictions imposed on the signs and magnitudes of α_k for $k = 2, \ldots, K$ and, by definition, $\alpha_1 = 1$. The time budget constraint implies that any variation in sleep must be accompanied by variations in one or more uses of time, that is, $\sum_k \alpha_k = 0$ and $\sum_{k \neq 1} \alpha_k = -1$.

This inescapable fact implies that we can observe the total derivative of Y with respect to T_1, but not the partial derivative, f_1 [Fiorini and Keane, 2014]. Specifically, a time reallocation compensating an increase in sleep of dT_1 implies a change in productivity equal to

$$dY_i = \sum_k f_k(T_{i1}, \ldots, T_{iK}) \alpha_k dT_1 = (f_1 + \sum_{k \neq 1} \alpha_k f_k) dT_1 \tag{2}$$

[10] Table B2 in Appendix B describes the occupational composition of the sample and Table B3 the distribution across days, months and years.

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled sample		**Male sample**		**Female sample**	
	Mean	SD	Mean	SD	Mean	SD
Weekly earnings ($/week)						
Current value (ATUS)	1,116.64	648.93	1,254.52	689.09	967.04	565.56
Lagged value (CPS MIS 4)	1,043.62	619.37	1,175.35	665.26	900.68	529.57
Time use (hours/week)						
Sleep	37.74	6.21	37.47	6.18	38.04	6.23
Work	46.04	7.59	47.74	7.58	44.20	7.15
Other non-work activities	36.21	9.00	34.79	8.86	37.76	8.89
Individual characteristics						
Female (%)	0.48	0.50	0.00	0.00	1.00	0.00
Age (years)	44.75	9.91	44.37	9.74	45.16	10.08
White (%)	0.84	0.37	0.86	0.35	0.81	0.39
Black (%)	0.10	0.31	0.08	0.27	0.13	0.34
Asian (%)	0.04	0.21	0.05	0.22	0.04	0.19
Other race (%)	0.02	0.13	0.02	0.12	0.02	0.13
No high school diploma (%)	0.04	0.21	0.06	0.23	0.03	0.17
High school diploma (%)	0.38	0.49	0.38	0.49	0.37	0.48
College degree or higher (%)	0.58	0.49	0.56	0.50	0.60	0.49
Married (%)	0.63	0.48	0.70	0.46	0.56	0.50
Widowed (%)	0.02	0.14	0.01	0.10	0.03	0.18
Divorced/separated (%)	0.17	0.38	0.13	0.34	0.22	0.42
Never married (%)	0.17	0.38	0.16	0.37	0.18	0.39
Household size	2.36	0.81	2.41	0.81	2.31	0.80
Any children present (%)	0.53	0.50	0.55	0.50	0.51	0.50
Employee	1.00	0.00	1.00	0.00	1.00	0.00
Self-employed	0.00	0.00	0.00	0.00	0.00	0.00
Unemployed	0.00	0.00	0.00	0.00	0.00	0.00
Not in labor force	0.00	0.00	0.00	0.00	0.00	0.00
Observations	12,398		6,452		5,946	

Notes: Summary statistics refer to the estimation sample. Appendix A describes sample selection criteria.

This is a combination of all partial derivatives f_k and does not permit the identification of f_1. If exactly one time use changes to compensate a sleep increase, say $k = 2$, we have that $\alpha_2 = -1$ and $\alpha_k = 0$ for all $k = 3, \ldots, K$. In this case, Equation (2) simplifies to

$$dY_i = (f_1 - f_2)dT_1 \tag{3}$$

While f_1 remains unidentified, we can identify the effect of sleep "relative" to the effect of time use $k = 2$: a quantity that can be interesting [Fiorini and Keane, 2014]. Empirically, identification of this parameter is achieved by comparing individuals with the same values for each $k = 3, \ldots, K$ but who differ in T_1 and T_2. Although this holds in general, in what follows we only distinguish between sleep ($k = 1$), non-work activities ($k = 2$) and work ($k = 3$), and focus on the relative effect of sleep versus non-work activities holding work time fixed.

Non-work activities are undoubtedly heterogeneous, encompassing a wide array of pursuits and actions. However, the most relevant distinction in time use lies between work and non-work activities, making it both reasonable and practical to group all non-work activities under a single category. This approach serves as a useful simplification, enabling a more manageable and streamlined analysis without compromising the overall validity of the results. In Section 4.6, we will refine this classification to partly address the nuances within non-work activities.

3.2 Complications arising from measuring labor productivity with labor earnings

When labor productivity is measured by labor earnings, things get a little more involved. Under competitive markets and constant marginal productivity, labor productivity coincides with the wage rate. In this case, labor earnings are

$$Y_i T_{i3} = f(T_{i1}, T_{i2}, T_{i3}) T_{i3} \tag{4}$$

so that the change in log earnings $E_i = log(Y_i T_{i3})$ caused by a change in sleep is

$$dE_i = \frac{1}{Y_i T_{i3}} \times [(f_1 + \alpha_2 f_2 + \alpha_3 f_3) T_{i3} + \alpha_3 f(T_{i1}, T_{i2}, T_{i3})] dT_1. \tag{5}$$

This can be rewritten as

$$dE_i / dT_1 = dlog Y_i / dT_1 + \alpha_3 / T_3 \tag{6}$$

The possibility to observe only the total effect rather than the partial effect of sleep on earnings assumes relevance here, as the total effect of sleep on log labor earnings corresponds to the total effect of sleep on log productivity only if $\alpha_3 = \frac{dT_3}{dT_1} = 0$. If, instead, an increase in sleep occurs together with a reduction in hours of work, as it seems plausible, then $\alpha_3 < 0$ and the total effect of sleep on labor earnings is a downward biased proxy of the total effect of sleeping on productivity.

To overcome this relevant issue, in our analysis we hold work time fixed, and thus attribute changes in earnings to changes in productivity.

Empirically, this will consist in comparing individuals with the same number of work hours. Holding work-time constant implies $\alpha_3 = 0$ and $\alpha_2 = -1$, so that

$$\frac{dE_i}{dT_1} = \frac{f_1 - f_2}{Y_i} \tag{7}$$

In this aspect, we depart from previous research [Gibson and Shrader, 2018, Kajitani, 2021, Costa-Font et al., 2024] which use empirical models that only include sleep hours and necessarily estimates a combination of the effects of all possible time substitutions, including sleep-to-work, which can be quite heterogeneous both in sign and in magnitude (see Equation (5) and Equation (6)).

3.3 The endogeneity of time use and our identification assumptions

The challenges stemming from the time budget constraint come together with another problem: the endogeneity in the allocation of time. Time use choices reflect a wide array of observable and unobservable drivers, some of which related with labor productivity.

To jointly address both challenges, we follow Todd and Wolpin [2003] and recent advances in the literature on the impact of time allocation on human capital production [Fiorini and Keane, 2014, Keane et al., 2022], which propose a value-added (VA) approach to identification. Accordingly, we refine the specification of labor earnings as

$$E_{it} = \sum_{\tau=1}^{t} (\beta_{t-\tau+1} I_{i\tau}) + \gamma_t \mu_{i0} + \varepsilon_{it}. \tag{8}$$

In Equation (8), E_{it} is the natural logarithm of individual earnings in period t, $I_{i\tau}, \tau = 1, \ldots, t$ is a vector of contemporaneous and past inputs, μ_{i0} are initial endowments and ε_{it} is a random shock. In our case, $I_{i\tau}$ includes all time use and a set of individual observed and unobserved traits affecting earnings.

This specification is quite demanding and the estimation of its parameters requires to observe the complete vector of inputs at all periods from birth to time $\tau = t$, a rather taxing requirement. Furthermore, we are only interested in the impact of sleep at time $\tau = t$ on contemporaneous earnings, while the parameters relating to the impact of other contemporaneous inputs or previous inputs dating before $\tau = t$ are not of primary importance in our analysis.

Fortunately, Todd and Wolpin [2003, 2007] show that, for identification of the impact of inputs at time $\tau = t$, Equation (8) can be conveniently simplified under rather mild conditions. Provided that $\beta_\tau = \lambda\beta_{\tau-1}$ and $\gamma_\tau = \lambda\gamma_{\tau-1}, \forall\tau$ - a "common depreciation rate" assumption[11] - Equation (8) can be rewritten as

$$E_{it} = \beta_1 I_{it} + \lambda E_{it-1} + \nu_{it} \tag{9}$$

If the "common depreciation rate" assumption holds, then the lagged outcome

[11] This assumption states that the effects of all observed and unobserved inputs, including initial endowments, geometrically decline at the same rate λ, meaning that the effect of past inputs declines as inputs get more remote.

represents a sufficient statistic to proxy for the complete history of all past observed and unobserved inputs, including the unobserved initial endowments. Thus, the value added approach drastically reduces the information required to estimate our parameters of interest.

A remaining concern is that Equation (9) assumes that all contemporaneous inputs in vector I_{it} are included in the specification of the production function, but due to data availability limitations the omission of unobservable contemporaneous inputs remains possible.

We specify the vector I_{it} as $I_{it} = (T_{it}, X_{it}, Z_{it})$, where T_{it} is the vector of time uses (sleep, work, non-work); X_{it} are other observed inputs; and Z_{it} are unobserved inputs. We assume that among full-time workers that are comparable in all their observable characteristics as well as past productivity, differences in how they allocate their time across the three considered activities are as good as random. Using the time budget constraint to write $NonWork_{it} = \bar{T} - Sleep_{it} - Work_{it}$, we state this assumption as follows:

$$\nu_{it} \perp (Sleep_{it}, Work_{it}, \bar{T} - Sleep_{it} - Work_{it}) | E_{it-1}, X_{it} \tag{10}$$

Under this condition, we express Equation (9) by specifying the linear model

$$E_{it} = \beta Sleep_{it} + \gamma Work_{it} + \delta X_{it} + \lambda E_{it-1} + \nu_{it} \tag{11}$$

that identifies the the effect of sleep relative to non-work activities through the parameter β.

The conditional independence of the joint distribution of time uses in Equation (10) implies that $Work_{it}$ can be treated as an exogenous control in Equation (11) for the purpose of isolating the trade-off between sleep and non-work time.

3.4 Estimation

We take Equation (11) to the data by estimating with Ordinary Least Squares (OLS) the following flexible regression specification:

$$\begin{aligned} E_{it} = &\beta Sleep_{it} + \sum_{v=3}^{19} \gamma_{vg} \times Work_{itvg} + \\ &+ \sum_{v=3}^{19} \delta_{vg} \times Work_{itvg} \times Y_{it-1} + \\ &+ \sum_{v=3}^{19} \gamma_{vg} \times Work_{itvg} \times X_{it} + \sum_{s=2}^{S} \alpha_{sg} State_{is} + \lambda E_{it-1} + \nu_{it} \end{aligned} \tag{12}$$

where E_{it} is the natural logarithm of usual weekly earnings for individual i at period t, E_{it-1} is the natural logarithm of lagged weekly earnings (measured 14 to 17 months earlier), $Sleep_{it}$ is hours of sleep, and we flexibly control for work time by including in Equation (11) a vector of dummies for being in the $v-th$ ventile of the gender-specific distribution of weekly working hours, $Work_{itvg}, v = 3, \ldots, 19$.[12]

The parameter β is the coefficient of interest, identifying the effect of an increase in sleep relative to a decrease in non-work activities (the omitted time use category), holding the number of work hours fixed.

The vector of observable characteristics X_{it} is listed in Section 2.2.3 and includes demographic and socioeconomic characteristics. To allow for flexibility in the mapping between E_{it} and both E_{it-1} and X_{it}, we allow the coefficients related with these variables, δ and γ, to vary by ventile of the gender-specific work time distribution.

Finally, $State_{it}$ is a set of state fixed effects and ν_{it} is an idiosyncratic individual error term.

To account for gender heterogeneity in the labor market conditions, working times, as well as sleep patterns and biological needs, in the analysis we allow for gender heterogeneity by using gender-specific coefficients on all control variables and state fixed effects when estimating pooled effects, and by splitting the sample by gender when investigating gender differences.

4 Results

4.1 Main results

The effects of sleep (relative to non-work activities) on labor productivity are precisely estimated zeros as reported in Table 2. We estimate that increasing sleep by 1 standard deviation - equivalent to 6.2 hours per week,[13] or 17 percent of mean sleep, a substantial time use change - would affect earnings by a tiny -0.25%. This result holds both in the pooled sample and after splitting by gender. The large value of the R-squared, ranging between 0.5 and 0.7, remarks that the set of controls and fixed effects included in the models captures a substantial share of the heterogeneity in current earnings.

[12] Table B4 in Appendix B shows that this functional form permits a finely-grid partitioning of the support of work hours, while at the same time leaving enough variation in sleep to permit identification. We defined gender-specific work hours ventiles during the sample selection, and the two extremes are trimmed out when defining the estimation sample. Details are described in Appendix A

[13] This is the variation in weekly sleep duration across individuals in our sample. Knutson et al. [2007] finds that the within-individual standard deviation of sleep duration per day is 1.26 hours, or 6.3 hours on a weekly basis, very similar to our benchmark.

Table 2: Main results

	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled sample		**Male sample**		**Female sample**	
	ln(earn)	**ln(earn)**	**ln(earn)**	**ln(earn)**	**ln(earn)**	**ln(earn)**
Sleep	-0.00078	-0.00040	-0.00090	-0.00029	-0.00066	-0.00051
	(0.00054)	(0.00058)	(0.00075)	(0.00083)	(0.00076)	(0.00082)
Observations	12,398	12,398	6,452	6,452	5,946	5,946
Controls	No	Yes	No	Yes	No	Yes
R^2	0.54	0.71	0.54	0.70	0.49	0.68
Sleep mean (SD)	37.74 (6.21)		37.47 (6.18)		38.04 (6.23)	

Notes: each column reports the coefficient on sleep duration obtained from the OLS estimation of Equation (12). All models include controls for gender-specific work-hour ventile dummies, lagged productivity, and their interactions. Models in even columns also include the controls in vector X, described in section Section 2.2.3, interacted with gender-specific work-hour ventile dummies, and a set of state fixed effects. In the pooled sample, all coefficients are allowed to be gender-specific. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Specification tests

The main results in Table 2 hold when we adopt different specifications. Given the stability of the results, we present the estimates in Appendix B.

First, Table Table B5 in Appendix B shows that results are unchanged when - unlike in our baseline - we use survey weights to reduce potential issues due to differences in the sampling and response rates across subpopulations and days of the week in ATUS. This finding also tones down potential concerns about the limited representativeness of our sample, as in that case the impact use of survey weights should make a starker difference.

Second, in Table Table B6 in Appendix B we allow for correlation in the error terms across individuals living in the same state and working in the same occupation at the CPS MIS 4 interview. The precision of our estimates is unghanged.

Next, Table Table B7 in Appendix B probes the robustness of the results when we change the specification adopted to control for work hours in Equation (12) and partition its support in 10, 20 (our baseline choice), 30, 50 or even 100 quantiles and use quantile dummies. In the last column of the Table we instead control for work hours linearly and drop the interactions of work time with the other controls, replicating a specification common in the literature [Fiorini and Keane, 2014, Keane et al., 2022]. Changing the binning of work-time has no relevant effect on the estimates. Instead, the estimated effect of sleep becomes slightly larger when we use the less flexible linear specification.

Finally, in Table B8 in Appendix B we compare our baseline regression, where we interact work-hour ventile dummies with all controls as well as with lagged earnings, with to two slightly different versions: the former excludes the interactions between work-hour ventile dummies and lagged earnings, while the latter adds additional interactions between work-hour ventile dummies and a linear trend in work hours. Their inclusion serves to control for potential residual variability of work hours within ventiles (see Table B4). The results are comparable to our baseline.

4.3 Robustness checks

In this section we consider three concerns regarding our identification strategy, namely: a possible violation of the conditional independence assumption, the possible endogeneity of lagged earnings and the consequences of meausurement error in our measure of sleep.

4.3.1 Selection on unobservables

Our identification strategy relies on a selection on observables assumption (see Equation (10)), but selection on (contemporaneous) unobservables can still confound identification. We assess the relevance of this matter by using the methods developed by Oster [2019]. Specifically, we leverage changes in the effect of sleep and in the R-squared of the regression that result from the inclusion of the controls in vector X to gauge the sensitivity of our estimated effects to the presence of selection on unobservables.

In particular, we estimate how large would the estimated effects of sleep be if unobservable inputs were as relevant as observables in affecting our estimates. In the parlance of Oster [2019], this amounts to estimating the level of β - the treatment effect - obtained under the assumption that δ - the degree of proportional selection on observables and unobservables - is equal to one. Assuming that the maximum attainable R-squared if all unobservables were included in the regression was equal to 1.3 times the R-squared of the model that includes all observed covariates, as suggested by Oster [2019], we obtain that the impact of sleep would change only marginally, from -0.0004 to 0.0008 in the pooled sample, from -0.0005 to 0.001 for males, and from -0.0003 to -0.0002 for females.

We also assess how relevant should selection on unobservables be in order to lead us to estimate a positive and large coefficient of sleep on productivity - for instance at least as large as the one estimated by Gibson and Shrader [2018], equal to 0.11. This amounts to estimate the level of δ such that $\beta = 0.11$. Under the same assumption on the maximum R-squared discussed above, for the pooled sample we estimate that δ would need to be larger than 2, a very large value according to the standards adopted to gauge proportional selection by Oster [2019].

4.3.2 Endogenous lagged productivity and IV estimation

According to the structural interpretation of Equation (9), as in Todd and Wolpin [2003], the error term is $\nu_{it} = \varepsilon_{it} - \lambda\varepsilon_{it-1}$, implying that E_{it-1} is correlated with ν_{it}. We tackle this issue by using the 2-year lagged mean earnings by gender, year, state and occupation from American Community Survey (ACS) microdata as an instrumental variable (IV) for E_{it-1}.[14] The instrument relies on the persistence in the local wage structure by gender and occupation to generate variation in earnings that is predictive of individual earnings, but at the same time does not depend on individual traits - besides those determining location and occupation choices.

We facilitate the estimation of the IV regression by removing the interactions between lagged earnings and the work-hour ventile dummies (as done in the first column of Table B8), so that we only have to instrument one endogenous variable with one instrument.

Column (1) of Table 3 reports the OLS estimates obtained with this specification in the pooled sample and after splitting by gender. These results are very similar to those in Table 2. Column (2) of the same Table reports positive and significant first-stage coefficients of lagged mean earnings on current earnings. The Kleibergen and Paap [2006], first-stage F statistics are well above 10, suggesting that the instruments are not weak. Finally, the IV estimates are in Column (3) of Table 3, and still report small effects of sleep on productivity, that - as in our baseline - are statistically indistinguishable from zero.

4.3.3 Measurement error

Time use as reported in ATUS eventually refers to a given day and not to the usual time allocation. The evidence provided by epidemiological studies such as Jonasdottir et al. [2021], Willoughby et al. [2023] suggests that there is non-negligible intra-individual variability in sleep time across consecutive working days. As a result, a single sleep measure might not properly reflect individuals' usual sleep schedule. This fact raises concerns about measurement error in sleep.

If variation in sleep across days is random, then measurement error is classical and leads to attenuation bias. In the absence of suitable instrumental variables in ATUS (such as alternative measures of respondents' time use allocation reported by

[14]The definition of earnings and occupation categories reported in the ACS match well those in ATUS and the CPS. On the one hand, labor earnings is defined as respondent's total pre-tax wage and salary income (i.e., money received as an employee), including wages, salaries, commissions, cash bonuses, tips, and other money income received from an employer, while excluding payments-in-kind or reimbursements for business expenses. On the other hand, a harmonized occupation coding scheme, based on the SOC 2010 and offered by IPUMS, allows us to match CPS and ACS data over the entire sample period.

Table 3: OLS and IV estimates

	(1) OLS ln(earn)	(2) First stage ln(earn) lag	(3) IV ln(earn)
		Pooled sample	
Sleep	-0.00053	-0.0022**	0.0010
	(0.00061)	(0.00085)	(0.00089)
ln(earn) lag	0.40***		1.13***
	(0.030)		(0.15)
ln(ACS mean earnings) lag - IV		0.30***	
		(0.048)	
Observations	12,398	12,398	12,398
K-P F stat		38.2	
N clusters		1,716	
		Male sample	
Sleep	-0.000085	-0.0023*	0.0015
	(0.00089)	(0.0012)	(0.0014)
ln(earn) lag	0.44***		1.17***
	(0.034)		(0.23)
ln(ACS mean earnings) lag - IV		0.27***	
		(0.064)	
Observations	6,452	6,452	6,452
K-P F stat		17.7	
N clusters		918	
		Female sample	
Sleep	-0.00098	-0.0020*	0.00048
	(0.00082)	(0.0012)	(0.0011)
ln(earn) lag	0.37***		1.08***
	(0.046)		(0.19)
ln(ACS mean earnings) lag - IV		0.33***	
		(0.071)	
Observations	5,946	5,946	5,946
K-P F stat		22.1	
N clusters		798	

Notes: with respect to the even columns of Table 2, the models estimated in this Table exclude the interactions between lagged earnings and work-hour ventile dummies. The instrumental variable for own ln(earn) lag used in Columns (2) and (3) is the 2-year lagged ln of mean earnings by gender, year, state and occupation from American Community Survey (ACS) microdata. Standard errors clustered by state-gender-MIS 4 occupation are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

a household member), we use simulations to illustrate that the consequences of measurement error are not substantive in our application.

We proceed as follows:

1. For each individual, we draw a value of "usual" night sleep duration - $UsualSleep_i$ - from a normal distribution with parameters taken from the evidence reported by epidemiological studies on sleep duration across individuals in the population, gathered using wearable devices [Jonasdottir et al., 2021, Willoughby et al., 2023].

2. For each individual we then draw a realisation of sleep time - $ActualSleep_i$ - from a normal distribution centered at the individual-specific usual night sleep duration $UsualSleep_i$, drawn in step 1, and with standard deviation taken from the intra-individual variability in sleep reported in the same studies.

3. For each individual, we specify individual productivity as $Y_i = \alpha + \beta UsualSleep_i + u_i$, where α is a constant - equal to average natural logarithm of weekly earnings in the ATUS sample; $UsualSleep_i$ is the individual-specific usual sleep time drawn in step 1; u_i is a normally distributed error term with zero mean and standard deviation equal to the one observed for the natural logarithm of weekly earnings in the ATUS sample; and β is set equal to the effect found by Gibson and Shrader [2018]. Since they use total sleep hours over 7 days, while in this simulation we focus on sleep hours per night, we divide their effect by 7.

4. We regress Y_i on $ActualSleep_i$ and save the resulting estimate of β.

The empirical distribution of the sleep effects β estimated in 1000 iterations of this procedure are reported in Figure B6 and Figure B7. The former takes the parameters of the sleep distributions from Jonasdottir et al. [2021] and the latter from Willoughby et al. [2023]. The results reveal an attenuation factor around 43%, implying that even in the absence of measurement error our estimated effects would remain very small.

4.4 Non-linear effects

While our main specification in Equation (12) assumes that the effect of sleep is linear across its support, it may be that this effect is in fact non-linear. For instance, trading hours of sleep with activities other than work may be beneficial for productivity only if workers get enough sleep to function properly or if they do not oversleep. Given that eventually we do not find much evidence of non-linear responses, we again report results in Appendix B.

In Table B10 we first use a quadratic instead of linear functional form for sleep duration. We find evidence that, in the pooled sample and for males, increasing sleep has a positive effect on productivity up to a bliss point at 37,5 hours of sleep a week (i.e., around 7 hours and a half per night), and then declines. For females, the pattern is qualitatively similar but the curvature is less accentuates, the bliss point comes earlier, and the overall impacts are too small to be significantly estimated. Despite the non-linear behaviour, the effects remain very small. For instance, considering the pooled specification, we find that increasing sleep by 1SD (6.2 hours) from the mean (37.7 hours) decreases productivity by 0.71%, while decreasing sleep by 1SD from the mean decreases productivity by 0.16%. However, the effects are not statistically different from zero or from each other.

We confirm these inverted U-shaped patterns even when we adopt a semiparametric approach and replace the linear specification of sleep with dummies for being in different fifths of the sleep distribution.[15]

Overall, these results bring some qualitative support to the medical insights suggesting that both excessively short and long sleep duration impairs health and cognitive functioning, which may reflect into productivity.[16] Nevertheless, the differences in productivity that we estimate for different sleep durations are again very small, confirming our main finding that the substitution between sleep and non-work related activities is not a key driver of productivity.

4.5 Heterogeneity analysis

This section investigates whether the relative effect of sleep varies along the earnings distribution, the levels of other time uses, and individual characteristics. Overall, we find little evidence of heterogeneous effects. Results are collected in Appendix B.

First, Table B11 reports the impacts of sleep on the 10th, 25th, 50th, 75th and 90th quantiles on the earnings distribution, estimated via unconditional quantile regressions [Firpo et al., 2009]. While differences in the effects at different quantiles are not significant, a positive gradient in the effect is visible, pointing to more sleep being harmful at the bottom of the distribution and beneficial at the top. Again, however, effects are rather small. For instance, if everyone slept one additional hour a week, at the expense of one hour of non-work activities, the gap in earnings between the 10th and the 90th quantiles of the earnings distribution would only widen by 0.46%.

Similarly, in Table B12 we estimate a positive gradient in the effect of sleep on mean earnings as we move along the fifths of the lagged earnings distribution. Still, differences in the effects are small in size and not statistically significant.

[15] Table B9 and Figure B5 illustrate the fifths of the sleep distribution.

[16] We also test non-linearities using specific thresholds for sleep duration used in medical research: one for sleeping less than 6 hours per night and another for sleeping between 7 and 9 hours per night. In all the three samples, coefficients associated with the relative dummies are negative in the former case and positive in the latter, in line with medical insights. Yet, in either case they are not statistically different from zero.

Next, we investigate whether the consequences of different sleep durations become evident only for those who work for many hours a day. Table B13 reports the estimates of a model where we linearly introduce sleep and work durations - instead of the ventiles for work used in our main specification - as well as an interaction term between the two. It turns out that the estimate of the coefficient related with this interaction term is close to zero in magnitude and not statistically significant, indicating that the production function does not exhibit these sort of complementarities. We obtain similar results when we use a more flexible specification that uses work ventiles, in Table B14.[17]

Finally, we investigate whether the effect of sleep differs depending on individual characteristics. In Table B15, Table B16, and Table B17 we respectively assess heterogeneous effects by age groups, the presence of children in the household, and occupations. However, we don't find evidence of heterogeneous effects depending on any of these characteristics.

4.6 Time use trade-offs

Our empirical exercise differs from other studies on the impacts of sleep on labor productivity mostly because we hold work time fixed and estimate the trade-off between sleep and other non-work related activities. As stressed by Fiorini and Keane [2014] and Keane et al. [2022], the net impact of sleep crucially depends on the activities that replace it, and no meaningful interpretation can be given to unconditional effects. In what follows, we report two pieces of evidence that clarify the importance of specifying what time use trade-off one looks at.

Table 4: Time uses trade-offs

	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled sample		**Male sample**		**Female sample**	
	Work	Other	Work	Other	Work	Other
Sleep	-0.19***	-0.81***	-0.23***	-0.77***	-0.14***	-0.86***
	(0.011)	(0.011)	(0.015)	(0.015)	(0.016)	(0.016)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,398	12,398	6,452	6,452	5,946	5,946
R^2	0.14	0.39	0.11	0.35	0.08	0.40

Notes: We use the same specification as in the even columns of Table 2, but the dependent variables are hours of work and hours spent in other non-work activities and - given the presence of the time budget constraint - we do not control for other uses of time but sleep. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[17] In principle, a possibility to investigate this type of heterogeneity could be to compare between full- and part-time workers. However, ATUS only collects one time diary per respondent, and the labor supply of vertical part-time workers may be misclassified if surveyed on a convenient day, leading to bias. Additionally, very few males work part-time. For these reasons, this analysis isn't pursued. See Section 2.3

Preliminarily, Table 4 reports the substitution patterns between sleep and the two other uses of time that we consider - work and non-work activities. We estimate these patterns with a specification akin to Equation (12), where the dependent variables are either work (in odd columns) or non-work (in even columns) hours and - given the presence of the time budget constraint - we do not include other uses of time besides sleep. We find that, when sleep duration increases, work hours decrease less than other activities do. One additional hour of weekly sleep is substituted by a decrease in work hours of roughly 12 minutes and a decrease in non-work activities of 48 minutes - with men decreasing work time slightly more than women do.

Table 5 reports the impacts of sleep on earnings when we: *i)* do not control for other time uses - in Column (1) - allowing for unspecified substitution patterns; *ii)* fix work hours - in Column (2), which replicates our baseline - thereby estimating the trade-off between sleep and non-work activities; *iii)* fix hours spent in non-work activities - in Column (3) - thus estimating the trade-off between sleep and work hours.

Table 5: Unconditional sleep effects, sleep vs. work, sleep vs. other activities

	(1) Unconditional	**(2) Sleep vs. other**	**(3) Sleep vs. work**
		Pooled sample	
Sleep	-0.0016***	-0.00040	-0.0053***
	(0.00051)	(0.00058)	(0.00073)
Observations	12,398	12,398	12,398
R^2	0.62	0.71	0.72
		Male sample	
Sleep	-0.0013*	-0.00029	-0.0053***
	(0.00072)	(0.00083)	(0.00099)
Observations	6,452	6,452	6,452
R^2	0.61	0.70	0.70
		Female sample	
Sleep	-0.0019***	-0.00051	-0.0052***
	(0.00074)	(0.00082)	(0.0011)
Observations	5,946	5,946	5,946
R^2	0.58	0.68	0.70

Note: Column (2) replicates the specification in the even columns of Table 2. Column (1) only controls for hours of sleep, and for no other time use, while Column (3) substitutes gender-specific work-hour ventile dummies with gender-specific dummies for the ventiles of the distribution of non-work related activities. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We find a negative unconditional effect, that results from a weighted average of the zero substitution effect between sleep and non-work activities and the negative and significant substitution effect between work and sleep. These findings support the prediction of Section 3 that leaving work-time free to adjust in the model specification causes a downward bias in the estimates of the marginal effect of sleep.

Next, in Table 6, we verify whether the relative effect of sleep varies depending on the specific non-work activity that replaces it. We split non-work time in housework and care vs. leisure, and estimate the effect of sleep relative to either of the two, while controlling for quantiles of the joint distribution of work-by-the-other-non-work-time.[18]

Table 6: Heterogeneity by non-work time sub-group

	(1) Sleep vs. other	(2) Sleep vs. housework	(3) Sleep vs. leisure	(4) Sleep vs. rel. leisure
	Pooled sample			
Sleep	-0.00040	-0.0024**	0.0000026	0.00055
	(0.00058)	(0.00095)	(0.00089)	(0.00096)
Observations	12,398	12,398	12,398	12,397
R^2	0.71	0.82	0.82	0.82
Mean (SD) alt. time use	36.21 (9.00)	20.51 (9.14)	15.70 (8.74)	10.72 (7.87)
	Male sample			
Sleep	-0.00029	-0.0017	0.00029	0.0012
	(0.00083)	(0.0013)	(0.0012)	(0.0013)
Observations	6,452	6,452	6,452	6,452
R^2	0.70	0.81	0.81	0.81
Mean (SD) alt. time use	34.79 (8.86)	18.28 (8.49)	16.51 (8.77)	11.55 (8.12)
	Female sample			
Sleep	-0.00051	-0.0033**	-0.00034	-0.00020
	(0.00082)	(0.0014)	(0.0013)	(0.0014)
Observations	5,946	5,946	5,946	5,945
R^2	0.68	0.82	0.81	0.80
Mean (SD) alt. time use	37.76 (8.89)	22.94 (9.19)	14.82 (8.62)	9.81 (7.49)

Notes: we use the same specification as in the even columns of Table 2, and further control for the specific non-work time sub-group. Specifically, columns (2)-(4) isolate time trade-offs by controlling for deciles of work-time by fifths of non-work time dummies. Column (1) replicates the specification in the even columns of Table 2. Column (2) fixes hours spent in leisure, yielding estimates of the effect of sleep relative to caring and household activities. Column (3) fixes hours spent in caring and household activities, yielding estimates of the effect of sleep relative to leisure. Column (4) fixes hours in caring and household activities and non-relaxing forms of leisure, yielding estimates of the effect of sleep relative to relaxing forms of leisure. Robust standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We find that sleep is equivalent to leisure, and that for females increasing sleep at the expense of housework damages productivity. Although perhaps surprising, this

[18]Specifically, we create combinations of work deciles and fifths of each of the non-work time subcategory. This allows to isolate the specific time trade-off of interest while comparing individuals with the same composition of work and remaining-non-work time. The choice of deciles and fifths is necessary to have sufficient observations in each cell, avoiding arbitrary decisions on how to aggregate small groups.

finding is in line with Gupta et al. [2003], who suggest that housework has a more negative effect on wages for women than for men due to the different types and the different degree of flexibility of the activities they engage in. Indeed, on top of the total duration, while women tend to engage in fixed activities that need to be done at certain times (e.g., childcare and daily chores), men tend to engage in activities that can be more freely allocated (e.g., gardening). If so, with hours of work fixed, for women an increase in sleep further compresses the time available to complete these inflexible tasks, increasing pressure and decreasing concentration at work, thus damaging productivity.

5 Conclusions

This paper approaches the question of whether sleep affects labor productivity, addressing the main empirical challenges highlighted by the research on the effects of time use [Fiorini and Keane, 2014, Keane et al., 2022].

The presence of the time budget constraint implies that estimating the impact of increasing a single use of time in isolation is impossible, as other uses of time would necessarily change. We overcome this issue by focusing on estimating the impact of substituting sleep with non-work related activities, holding work hours fixed.

We further borrow from the economics of time use by addressing the endogeneity of the overall time allocation with a value-added (VA) specification. As argued by Todd and Wolpin [2003], the VA model controls for the lagged outcome as a sufficient statistics for all past inputs in the individual production function. The inclusion of a comprehensive set of contemporaneous observable controls alleviates residual concerns of omitted variables, as further confirmed by the results of the Oster [2019] test.

Using data from the ATUS and CPS surveys of American full-time workers, we find that - once we hold working hours fixed - trading one hour of sleep for one hour of non-work time does not affect earnings. This result is robust to a series of specification tests, and further estimates do not reveal evidence of non-linear or heterogeneous effects.

This finding does not mean that sleep is unnecessary. The medical and biological literature has long clarified that this is not the case. Rather, it means that, as regards labor productivity, the benefits of additional sleep can be offset by the lost benefits of other non-work activities that must be given up. For instance, if one more hour of sleep comes at the cost of one less hour of physical exercise, team building with the co-workers, or duties which would increase stress if not done, the effects on labor productivity are hard to predict and can be very small, as our findings show.

Our result are not exempt from limitations, that call for future research. For instance, our analysis is conducted on a sample of full-time employees, and results may not be generalizable to part-time or self-employed workers. Furthermore, our results concern sleep duration, but individuals may improve productivity by optimizing the timing or improving the quality of their sleep, for given duration. Unfortunately, our data do not have information on sleep quality, and credible variation in sleep quality and timing is also hard to exploit. More generally, for each human activity, including sleep, what matters is not only the amount of time devoted to it, but also the intensity with which it is pursued.

References

J. D. Angrist and J.-S. Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2009.

P. Bessone, G. Rao, F. Schilbach, H. Schofield, and M. Toma. The economic consequences of increasing sleep among the urban poor. *The Quarterly Journal of Economics*, 136(3):1887–1941, 2021.

J. E. Biddle and D. S. Hamermesh. Sleep and the allocation of time. *Journal of Political Economy*, 98(5, Part 1):922–943, 1990.

C. Caetano, G. Caetano, and E. Nielsen. Are children spending too much time on enrichment activities? *Economics of Education Review*, 98:102503, 2024.

G. Caetano, J. Kinsler, and H. Teng. Towards causal estimates of children's time allocation on skill development. *Journal of Applied Econometrics*, 34(4):588–605, 2019.

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

J. Costa-Font and S. Fleche. Child sleep and mother labour market outcomes. *Journal of Health Economics*, 69:102258, 2020.

J. Costa-Font, S. Fleche, and R. Pagan. The labour market returns to sleep. *Journal of Health Economics*, 93:102840, 2024.

J. E. Ferrie, M. Kumari, P. Salo, A. Singh-Manoux, and M. Kivimäki. Sleep epidemiology—a rapidly growing field, 2011.

M. Fiorini and M. P. Keane. How the allocation of children's time affects cognitive and noncognitive development. *Journal of Labor Economics*, 32(4):787–836, 2014.

S. Firpo, N. M. Fortin, and T. Lemieux. Unconditional quantile regressions. *Econometrica*, 77(3):953–973, 2009.

M. Gibson and J. Shrader. Time use and labor productivity: The returns to sleep. *Review of Economics and Statistics*, 100(5):783–798, 2018.

O. Giuntella and F. Mazzonna. Sunset time and the economic effects of social jetlag: evidence from us time zone borders. *Journal of health economics*, 65:210–226, 2019.

O. Giuntella, S. Saccardo, and S. Sadoff. Sleep: Educational impact and habit formation. Technical report, National Bureau of Economic Research, 2024.

N. D. Gupta, J. Bonke, and N. Smith. Timing and flexibility of housework and men's and women's wages. 2003.

L. Hale, W. Troxel, and D. J. Buysse. Sleep health: an opportunity for public health to address health equity. *Annual review of public health*, 41(1):81–99, 2020.

M. Jagnani. Children's sleep and human capital production. *Review of Economics and Statistics*, 106(4):983–996, 2024.

S. S. Jonasdottir, K. Minor, and S. Lehmann. Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study. *Sleep*, 44(2):zsaa169, 2021.

S. Kajitani. The return of sleep. *Economics & Human Biology*, 41:100986, 2021.

M. J. Kamstra, L. A. Kramer, and M. D. Levi. Losing sleep at the market: The daylight saving anomaly. *American Economic Review*, 90(4):1005–1011, 2000.

M. Keane, S. Krutikova, and T. Neal. Child work and cognitive development: Results from four low to middle income countries. *Quantitative Economics*, 13(2):425–465, 2022.

F. Kleibergen and R. Paap. Generalized reduced rank tests using the singular value decomposition. *Journal of econometrics*, 133(1):97–126, 2006.

K. L. Knutson, P. J. Rathouz, L. L. Yan, K. Liu, and D. S. Lauderdale. Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the cardia study. *Sleep*, 30(6):793–796, 2007.

J. Meeus. *Astronomical Algorithms*. Willmann-Bell, 1991.

H. T. Nguyen, S. R. Zubrick, and F. Mitrou. Daylight duration and time allocation of children and adolescents. *Economics & Human Biology*, 55:101435, 2024.

E. Oster. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204, 2019.

P. E. Todd and K. I. Wolpin. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485):F3–F33, 2003.

P. E. Todd and K. I. Wolpin. The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital*, 1(1):91–136, 2007.

A. R. Willoughby, I. Alikhani, M. Karsikas, X. Y. Chua, and M. W. Chee. Country differences in nocturnal sleep variability: Observations from a large-scale, long-term sleep wearable study. *Sleep Medicine*, 110:155–165, 2023.

Appendices

Appendix A Sample selection

To obtain a suitable sample for the analysis we apply several selection criteria.

We started from a sample of working-age individuals (25 to 65 years old)[19], interviewed in the years between 2003 and 2019, from the beginning of ATUS until COVID-19 outbreak. This initial sample counts 146,949 observations.

Second, given that sleep and work hours likely differ between weekdays and weekends, we exclude all time diaries completed during Saturdays or Sundays. Despite the considerable loss in sample size of about 50%, as implied by the ATUS sampling scheme[20], since only one time-diary is available for each individual we prefer not to infer work-week time allocation from weekends for 50% of the sample. For the same reason, we also drop information for respondents filling time diaries during holidays. This leaves us with 71,790 observations.

Among these, we select only those satisfying two conditions:

1. reporting to work full-time in ATUS and in CPS MIS 4 (when lagged earnings are measured);
2. with non-missing earnings values in either of the three interviews of CPS MIS 4, CPS MIS 8 and ATUS, and with no job change in the overall period.

Condition 1 is needed to avoid potential miss-classification of the labor supply of vertical part-time workers. The availability of one time diary only per respondent, combined with the impossibility to distinguish between horizontal and vertical part-time workers, could lead to severe errors in the measurement of usual work time if the latter are surveyed on a convenient day (see Section 2.3). Due to this reason, as also done in the literature [Gibson and Shrader, 2018], we retain full-time workers only, which count to 31,697.

On top of the need to have lagged and current earning information, condition 2 aims at limiting the possibility that some workers had an unemployment spell or were not receiving earnings between the two main time records. After this operation, the sample counts 16,002 observations.

We then exclude observations with earnings above the 1st and below the 99th percentile of the earnings distribution to avoid extreme potential misreporting, losing 157 more observations.

[19] We set 65 as the upper limit in line with the literature [Gibson and Shrader, 2018, Costa-Font et al., 2024]. However, in order to reassure against potentially different behaviours or labor fatigue around the retirement age, in Table B18 we also check robustness of results with two alternative cutoffs at 62 and 60 years.

[20] See footnote 5.

Finally, we exclude all time diaries of full-time workers reporting zero hours of sleep or work, and also drop time diaries with reported work duration falling in the upper or bottom 5% tails of the gender-specific distribution and with reported sleep duration below the 1% or above the 99%. Figure B1 and Figure B2 in Appendix B represent the distribution of time uses and the thresholds set in this trimming phase.

The resulting final sample includes 12,398 observations.

Appendix B Additional tables and figures

Table B1: Summary statistics in the initial sample

	Pooled sample		Male sample		Female sample	
	mean	sd	mean	sd	mean	sd
Weekly earnings ($/week)						
Current value (ATUS)	957.60	664.89	1129.63	706.95	797.93	579.28
Lagged value (CPS MIS 4)	928.36	634.77	1095.48	679.38	774.67	547.47
Time use (hours/week)						
Sleep	40.59	10.22	39.92	10.18	41.14	10.22
Work	28.91	23.81	34.99	23.55	23.97	22.85
Other activities	50.49	22.09	45.09	21.50	54.89	21.58
Individual characteristics						
Female (%)	0.55	0.50	0.00	0.00	1.00	0.00
Age (years)	44.37	10.91	44.63	10.75	44.16	11.03
White (%)	0.80	0.40	0.82	0.38	0.79	0.41
Black (%)	0.14	0.34	0.12	0.32	0.15	0.36
Asian (%)	0.04	0.19	0.04	0.19	0.04	0.20
Other race (%)	0.02	0.15	0.02	0.15	0.02	0.15
Married (%)	0.59	0.49	0.62	0.49	0.57	0.49
Widowed (%)	0.03	0.16	0.01	0.11	0.04	0.19
Divorced/separated (%)	0.18	0.39	0.16	0.37	0.20	0.40
Never married (%)	0.20	0.40	0.21	0.41	0.18	0.39
Household size	2.36	0.81	2.32	0.84	2.39	0.79
Any children present (%)	0.52	0.50	0.48	0.50	0.55	0.50
No high school diploma (%)	0.09	0.28	0.09	0.28	0.08	0.28
High school diploma (%)	0.43	0.49	0.43	0.50	0.42	0.49
College degree or higher (%)	0.49	0.50	0.48	0.50	0.50	0.50
Employee	0.68	0.47	0.73	0.44	0.64	0.48
Self-employed	0.08	0.28	0.11	0.31	0.06	0.24
Unemployed	0.04	0.20	0.04	0.20	0.04	0.20
Not in labor force	0.19	0.40	0.12	0.32	0.25	0.44
Observations	71,790		32,209		39,581	

Table B2: Occupational composition of the sample
(Census Occupation Classification)

	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled sample		Male sample		Female sample	
	Mean	SD	Mean	SD	Mean	SD
Management	0.13	0.34	0.15	0.35	0.12	0.32
Business and financial operations	0.06	0.24	0.05	0.22	0.07	0.26
Computer and mathematical science	0.05	0.22	0.08	0.26	0.02	0.15
Architecture and engineering	0.03	0.18	0.06	0.23	0.01	0.09
Life. physical. and social science	0.02	0.12	0.02	0.12	0.02	0.12
Community and social service	0.02	0.16	0.02	0.13	0.03	0.18
Legal	0.02	0.14	0.02	0.14	0.02	0.13
Education. training. and library	0.10	0.29	0.05	0.21	0.15	0.36
Arts. design. entertainment. sports. and media	0.02	0.13	0.02	0.13	0.02	0.12
Healthcare practitioner and technical	0.06	0.24	0.03	0.17	0.09	0.29
Healthcare support	0.01	0.11	0.00	0.04	0.02	0.16
Protective service	0.02	0.15	0.03	0.18	0.01	0.10
Food preparation and serving related	0.01	0.12	0.01	0.11	0.02	0.13
Building and grounds cleaning and maintenance	0.03	0.16	0.03	0.17	0.02	0.14
Personal care and service	0.01	0.09	0.01	0.07	0.01	0.10
Sales and related	0.07	0.25	0.08	0.27	0.06	0.23
Office and administrative support	0.15	0.36	0.05	0.22	0.26	0.44
Farming. fishing. and forestry	0.00	0.06	0.01	0.07	0.00	0.04
Construction and extraction	0.04	0.19	0.07	0.26	0.00	0.03
Installation. maintenance. and repair	0.04	0.20	0.07	0.26	0.00	0.05
Production	0.07	0.26	0.10	0.29	0.04	0.20
Transportation and material moving	0.04	0.19	0.06	0.25	0.01	0.09
Observations	12,398		6,452		5,946	

Table B3: Time composition of the sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled sample		**Male sample**		**Female sample**	
	Mean	SD	Mean	SD	Mean	SD
Day of the week						
Monday	0.20	0.40	0.20	0.40	0.19	0.39
Tuesday	0.20	0.40	0.21	0.40	0.20	0.40
Wednesday	0.21	0.41	0.21	0.40	0.21	0.41
Thursday	0.20	0.40	0.21	0.40	0.20	0.40
Friday	0.19	0.39	0.18	0.39	0.19	0.40
Month						
January	0.10	0.30	0.09	0.29	0.10	0.31
February	0.08	0.27	0.08	0.27	0.08	0.27
March	0.09	0.29	0.09	0.28	0.09	0.29
April	0.09	0.29	0.09	0.29	0.09	0.29
May	0.09	0.29	0.09	0.28	0.09	0.29
June	0.08	0.27	0.08	0.27	0.08	0.26
July	0.07	0.26	0.07	0.26	0.07	0.26
August	0.08	0.27	0.09	0.28	0.08	0.27
September	0.08	0.27	0.08	0.27	0.08	0.27
October	0.09	0.28	0.09	0.28	0.09	0.28
November	0.08	0.27	0.08	0.27	0.07	0.26
December	0.08	0.27	0.08	0.27	0.08	0.27
Year						
Year 2003	0.07	0.26	0.07	0.26	0.07	0.26
Year 2004	0.07	0.26	0.07	0.25	0.07	0.26
Year 2005	0.06	0.24	0.06	0.24	0.07	0.25
Year 2006	0.07	0.26	0.07	0.25	0.07	0.26
Year 2007	0.07	0.25	0.07	0.25	0.06	0.25
Year 2008	0.06	0.24	0.07	0.25	0.06	0.24
Year 2009	0.07	0.25	0.07	0.25	0.07	0.25
Year 2010	0.06	0.24	0.06	0.24	0.06	0.25
Year 2011	0.06	0.24	0.06	0.24	0.06	0.24
Year 2012	0.06	0.24	0.06	0.24	0.06	0.24
Year 2013	0.06	0.23	0.06	0.23	0.06	0.23
Year 2014	0.05	0.23	0.05	0.22	0.06	0.23
Year 2015	0.05	0.21	0.05	0.21	0.05	0.21
Year 2016	0.05	0.22	0.06	0.23	0.05	0.21
Year 2017	0.04	0.20	0.04	0.21	0.04	0.20
Year 2018	0.05	0.21	0.05	0.21	0.05	0.22
Year 2019	0.05	0.21	0.05	0.22	0.04	0.20
Minutes of sunlight	729.17	112.10	729.86	112.76	728.43	111.39
Observations	12,398		6,452		5,946	

Table B4: Work and sleep hours by work ventiles

Work hours Ventile	(1) Males work Mean	(2) Males work SD	(3) Males sleep Mean	(4) Males sleep SD	(5) Females work Mean	(6) Females work SD	(7) Females sleep Mean	(8) Females sleep SD
2	33.91	2.54	38.99	7.20	28.12	3.57	38.88	7.32
3	38.36	0.62	37.96	6.22	35.40	0.98	39.38	6.46
4	40.00	0.30	38.83	6.23	37.75	0.45	38.58	6.07
5	41.21	0.34	38.45	6.15	39.09	0.35	37.91	6.06
6	42.32	0.23	38.94	5.92	40.14	0.22	38.42	6.13
7	43.35	0.35	38.43	5.97	41.02	0.23	38.54	5.53
8	44.33	0.23	38.72	6.16	41.82	0.23	38.35	6.01
9	45.29	0.35	37.54	5.93	42.46	0.09	38.62	6.26
10	46.39	0.22	37.52	6.30	43.10	0.25	38.76	6.02
11	47.34	0.24	38.23	6.15	44.08	0.34	38.43	5.59
12	48.57	0.44	37.80	5.55	45.11	0.21	37.97	5.76
13	49.84	0.24	36.83	5.88	46.02	0.23	38.22	6.04
14	51.06	0.44	36.67	6.11	47.07	0.38	38.63	6.46
15	52.64	0.48	36.51	5.99	48.52	0.47	37.68	6.14
16	54.49	0.58	36.73	5.82	50.08	0.43	37.06	6.49
17	56.83	0.70	36.12	6.00	52.19	0.72	36.83	6.09
18	59.62	0.98	35.47	5.93	54.84	0.87	36.61	6.60
19	64.03	1.63	34.79	5.76	59.24	1.84	35.89	6.13
Observations	6,452		6,452		5,946		5,946	

Table B5: Robustness: WLS estimates using survey weights

	(1) **Pooled sample** ln(earn)	(2) **Male sample** ln(earn)	(3) **Female sample** ln(earn)
Sleep	-0.00031 (0.00063)	-0.00072 (0.00089)	0.00020 (0.00089)
Observations	12,398	6,452	5,946
R^2	0.73	0.72	0.71

Notes: we use the same specification as in the even columns of Table 2. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B6: Robustness: clustering standard errors

	(1) **Pooled sample** ln(earn)	(2) **Male sample** ln(earn)	(3) **Female sample** ln(earn)
Sleep	-0.00040 (0.00056)	-0.00029 (0.00087)	-0.00051 (0.00079)
Observations	12,398	6,452	5,946
R^2	0.71	0.70	0.68
N clusters	1,052	918	798

Notes: we use the same specification as in the even columns of Table 2. Standard errors clustered by state-by-occupation in MIS 4 in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B7: Robustness: different functional forms for work hours

	(1) **10th**	(2) **20th (baseline)**	(3) **30th**	(4) **50th**	(5) **100th**	(6) **Linear**
			Pooled sample			
Sleep	-0.00082	-0.00040	0.000050	-0.00026	-0.00013	-0.00076
	(0.00054)	(0.00058)	(0.00063)	(0.00078)	(0.0013)	(0.00052)
Observations	12,398	12,398	12,398	12,398	12,398	12,398
R^2	0.67	0.71	0.75	0.82	0.88	0.62
			Male sample			
Sleep	-0.00059	-0.00029	0.00023	-0.00066	0.00015	-0.00030
	(0.00077)	(0.00083)	(0.00087)	(0.0011)	(0.0018)	(0.00073)
Observations	6,452	6,452	6,452	6,452	6,452	6,452
R^2	0.66	0.70	0.74	0.81	0.88	0.61
			Female sample			
Sleep	-0.0011	-0.00051	-0.00014	0.00017	-0.00043	-0.0013*
	(0.00077)	(0.00082)	(0.00092)	(0.0011)	(0.0018)	(0.00074)
Observations	5,946	5,946	5,946	5,946	5,946	5,946
R^2	0.64	0.68	0.73	0.81	0.87	0.58

Notes: we use the same specification as in the even columns of Table 2, but a different functional form for work hours. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B8: Robustness: alternative specifications for the control variables

	(1) Lag not int	(2) Work int
	Pooled sample	
Sleep	-0.00053	-0.00031
	(0.00060)	(0.00058)
Observations	12,398	12,398
R^2	0.69	0.71
	Male sample	
Sleep	-0.000085	-0.000077
	(0.00084)	(0.00082)
Observations	6,452	6,452
R^2	0.69	0.70
	Female sample	
Sleep	-0.00098	-0.00055
	(0.00085)	(0.00083)
Observations	5,946	5,946
R^2	0.66	0.68

Notes: we use the same specification as in the even columns of Table 2, but a different functional form for the controls. The first column excludes the interactions between work-hour ventile dummies and lagged earnings, while the second adds interactions between work-hour ventile dummies and a linear trend in work hours. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B9: Distribution of sleep by fifths of the sleep distribution

Sleep fifth	(1) Male sample Mean	(2) Male sample SD	(3) Female sample Mean	(4) Female sample SD
1	29.16	3.17	29.23	3.30
2	34.65	0.77	35.07	0.92
3	37.55	0.79	38.68	1.08
4	40.87	1.13	41.71	0.76
5	46.67	3.00	46.63	3.11
Observations	6,452		5,946	

Table B10: Non-linear effects

	(1) **Pooled sample** ln(earn)	(2) **Male sample** ln(earn)	(3) **Female sample** ln(earn)
Sleep	0.0081*	0.015**	0.00093
	(0.0046)	(0.0065)	(0.0065)
Sleep^2	-0.00011*	-0.00020**	-0.000019
	(0.000061)	(0.000087)	(0.000085)
Observations	12398	6452	5946
R^2	0.71	0.70	0.68
Bliss point	36.81	37.5	23.25
Sleep - 1^{st} fifth	-0.0017	-0.0055	0.0019
	(0.010)	(0.014)	(0.015)
Sleep - 2^{nd} fifth	0.0094	0.00062	0.018
	(0.011)	(0.015)	(0.015)
Sleep - 3^{rd} fifth	-	-	-
Sleep - 4^{th} fifth	-0.0046	-0.0062	-0.0043
	(0.011)	(0.014)	(0.017)
Sleep - 5^{th} fifth	-0.014	-0.014	-0.013
	(0.011)	(0.016)	(0.015)
Observations	12,398	6,452	5,946
R^2	0.71	0.70	0.68

Notes: we use the same specification as in the even columns of Table 2, but adopt non-linear functional forms for sleep duration. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B11: Heterogeneous effects on different quantiles of the earnings distribution

	(1) **Pooled sample** ln(earn)	(2) **Male sample** ln(earn)	(3) **Female sample** ln(earn)
Sleep - 10^{th} earnings quantile	-0.0022*	-0.0043**	-0.00100
	(0.0013)	(0.0019)	(0.0018)
Sleep - 25^{th} earnings quantile	-0.0015	-0.0010	-0.00086
	(0.0010)	(0.0015)	(0.0014)
Sleep - 50^{th} earnings quantile	-0.00076	-0.0011	-0.0033***
	(0.00092)	(0.0013)	(0.0013)
Sleep - 75^{th} earnings quantile	0.0021*	0.0027*	0.0011
	(0.0011)	(0.0016)	(0.0015)
Sleep - 90^{th} earnings quantile	0.0024	0.0023	0.00038
	(0.0017)	(0.0021)	(0.0023)
Observations	12,398	6,452	5,946

Notes: we use the same specification as in the even columns of Table 2, but estimate impacts on different quantiles of the earnings distribution, using unconditional quantile regressions. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B12: Heterogeneous effects by fifths of the distribution of lagged earnings

	(1) **Pooled sample** ln(earn)	(2) **Male sample** ln(earn)	(3) **Female sample** ln(earn)
Sleep	-0.00022	-0.000097	-0.00042
	(0.00061)	(0.00087)	(0.00084)
1^{st} lagged earnings fifth $\times$ Sleep	-0.0060***	-0.0059***	-0.0061***
	(0.00049)	(0.00072)	(0.00067)
2^{nd} lagged earnings fifth $\times$ Sleep	-0.0030***	-0.0027***	-0.0032***
	(0.00031)	(0.00045)	(0.00044)
3^{rd} lagged earnings fifth $\times$ Sleep	-	-	-
4^{th} lagged earnings fifth $\times$ Sleep	0.0030***	0.0028***	0.0032***
	(0.00031)	(0.00046)	(0.00044)
5^{th} lagged earnings fifth $\times$ Sleep	0.0077***	0.0074***	0.0080***
	(0.00052)	(0.00077)	(0.00071)
Observations	12,398	6,452	5,946
R^2	0.72	0.71	0.70

Notes: we use the same specification as in the even columns of Table 2, and add interaction terms between sleep and dummies for belonging to different fifths of the distribution of lagged earnings. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B13: Heterogeneity by hours of work

	(1) **Pooled sample** ln(earn)	(2) **Male sample** ln(earn)	(3) **Female sample** ln(earn)
Sleep	0.0023	-0.00074	0.0058
	(0.0036)	(0.0056)	(0.0050)
Work	0.015***	0.017**	0.014*
	(0.0052)	(0.0074)	(0.0075)
Sleep × Work	-0.000057	0.000011	-0.00014
	(0.000078)	(0.00012)	(0.00011)
Observations	12,398	6,452	5,946
R^2	0.71	0.70	0.68

Notes: we use the same specification as in the even columns of Table 2, but linearly introduce sleep and work durations - instead of the ventiles - as well as an interaction term between the two. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B14: Heterogeneity by work hour ventile

	(1) Male sample ln(earn)	(2) Female sample ln(earn)
2^{nd} work ventile $\times$ Sleep	-0.0016	0.00084
	(0.0041)	(0.0032)
3^{rd} work ventile$\times$ Sleep	-0.0043	0.00036
	(0.0031)	(0.0034)
4^{th} work ventile $\times$ Sleep	0.00056	-0.0017
	(0.0033)	(0.0032)
5^{th} work ventile $\times$ Sleep	0.00038	-0.00042
	(0.0030)	(0.0032)
6^{th} work ventile $\times$ Sleep	0.0018	-0.00039
	(0.0036)	(0.0035)
7^{th} work ventile $\times$ Sleep	-0.0077***	0.0020
	(0.0028)	(0.0034)
8^{th} work ventile $\times$ Sleep	0.0069	0.0070**
	(0.0042)	(0.0033)
9^{th} work ventile $\times$ Sleep	0.0059	0.0011
	(0.0037)	(0.0041)
10^{th} work ventile $\times$ Sleep	-0.0022	-0.0035
	(0.0034)	(0.0036)
11^{th} work ventile $\times$ Sleep	0.0017	-0.0064*
	(0.0030)	(0.0037)
12^{th} work ventile $\times$ Sleep	-0.0044	-0.0022
	(0.0038)	(0.0038)
13^{th} work ventile $\times$ Sleep	0.0053	0.00073
	(0.0033)	(0.0028)
14^{th} work ventile $\times$ Sleep	-0.0038	-0.000037
	(0.0031)	(0.0031)
15^{th} work ventile $\times$ Sleep	0.0015	0.00036
	(0.0030)	(0.0032)
16^{th} work ventile $\times$ Sleep	0.0031	0.0013
	(0.0040)	(0.0038)
17^{th} work ventile $\times$ Sleep	0.0013	-0.000027
	(0.0031)	(0.0039)
18^{th} work ventile $\times$ Sleep	-0.0056	-0.0045
	(0.0037)	(0.0037)
19^{th} work ventile $\times$ Sleep	-0.000083	-0.0023
	(0.0037)	(0.0033)
Observations	6,452	5,946
R^2	0.70	0.68

Notes: we use the same specification as in the even columns of Table 2, bud add interaction terms between sleep and work hours ventiles. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B15: Heterogeneity by age group

	Pooled sample	**Male sample**	**Female sample**
	ln(earn)	ln(earn)	ln(earn)
Age 25-38 $\times$ Sleep	0.000073	0.00014	0.00000071
	(0.00065)	(0.00092)	(0.00094)
Age 39-51 $\times$ Sleep	-0.00041	-0.00025	-0.00058
	(0.00061)	(0.00086)	(0.00086)
Age 51-64 $\times$ Sleep	-0.00089	-0.00084	-0.00095
	(0.00066)	(0.00094)	(0.00093)
Observations	12,398	6,452	5,946
R^2	0.71	0.70	0.68

Notes: we use the same specification as in the even columns of Table 2, and add interaction terms between sleep and dummies for different age groups. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B16: Heterogeneity by presence of children

	Pooled sample	**Male sample**	**Female sample**
	ln(earn)	ln(earn)	ln(earn)
Without children $\times$ Sleep	-0.00074	-0.00069	-0.00082
	(0.00060)	(0.00087)	(0.00084)
With children $\times$ Sleep	-0.00014	0.000049	-0.00034
	(0.00060)	(0.00085)	(0.00085)
Observations	12,398	6,452	5,946
R^2	0.71	0.69	0.68

Notes: we use the same specification as in the even columns of Table 2, and add interaction terms between sleep and dummies for having young children in the household or not. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B17: Heterogeneity by occupation group

	Pooled sample	**Male sample**	**Female sample**
	ln(earn)	ln(earn)	ln(earn)
Management × Sleep	0.00073	0.0019	-0.00016
	(0.00078)	(0.0012)	(0.0010)
Service & Sales × Sleep	-0.0012	-0.0025*	-0.00069
	(0.00080)	(0.0014)	(0.0010)
Constr & Transp × Sleep	-0.0016	-0.0016	-0.0020
	(0.0010)	(0.0012)	(0.0019)
Observations	12,398	6,452	5,946
R^2	0.71	0.70	0.68

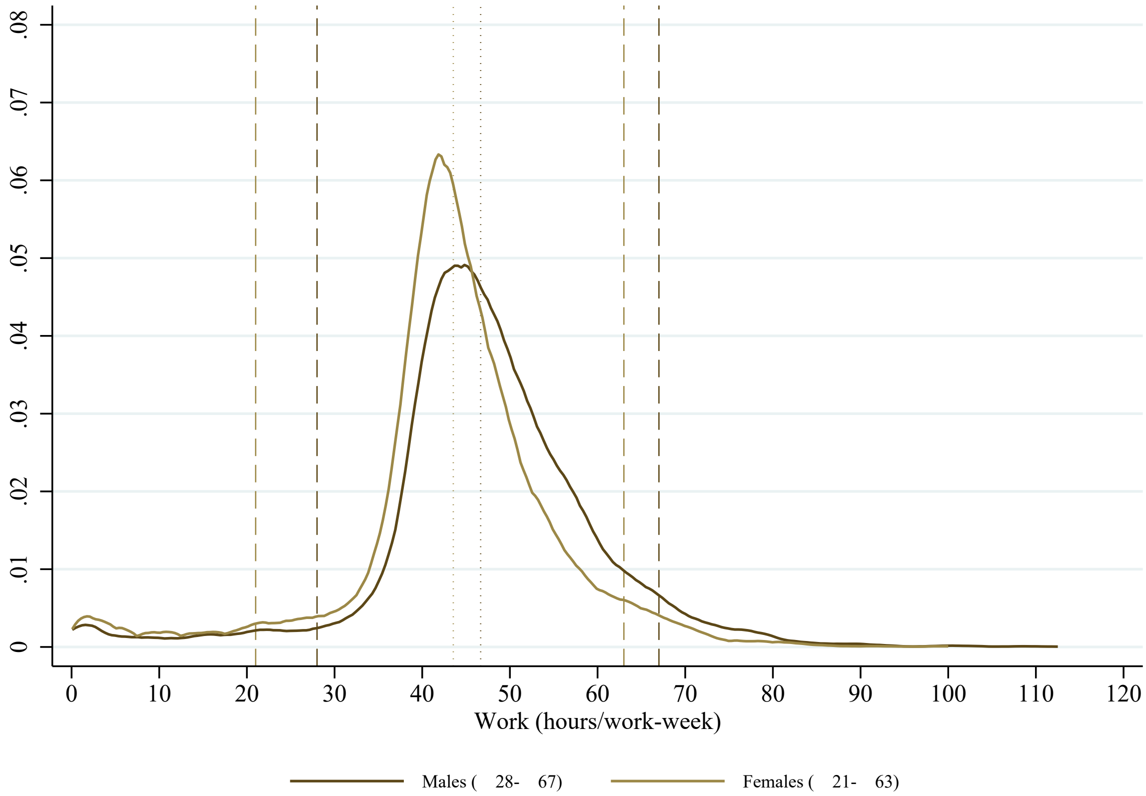
Notes: we use the same specification as in the even columns of Table 2, and add interaction terms between sleep and dummies for belonging to different occupations. We group occupation reported in CPS MIS 4 in three groups, broadly representing highly technical occupations, service-related and first sector occupations. The first category is exactly the first group in the SOC High-level aggregation, mainly referring to Management, Business, and Science Occupations, including highly technical ones such as Engineering and Legal occupations. The second category pairs together the second and third groups: "Service Occupations" and "Sales and Office Occupations". Finally, the third group combines SOC fourth and fifth high-level aggregation groups: "Natural Resources, Construction, and Maintenance Occupations" and "Production, Transportation, and Material Moving Occupations". Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B18: Robustness: alternative age cutoffs

	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled sample		**Male sample**		**Female sample**	
	ln(earn)	ln(earn)	ln(earn)	ln(earn)	ln(earn)	ln(earn)
			Maximum age 62			
Sleep	-0.00088	-0.00053	-0.0010	-0.00050	-0.00072	-0.00057
	(0.00055)	(0.00060)	(0.00076)	(0.00085)	(0.00079)	(0.00084)
Observations	11908	11908	6227	6227	5681	5681
Controls	No	Yes	No	Yes	No	Yes
R^2	0.54	0.71	0.54	0.70	0.49	0.69
			Maximum age 60			
Sleep	-0.00097*	-0.00054	-0.0011	-0.00048	-0.00084	-0.00061
	(0.00056)	(0.00062)	(0.00078)	(0.00088)	(0.00079)	(0.00085)
Observations	11,445	11,445	6,012	6,012	5,433	5,433
Controls	No	Yes	No	Yes	No	Yes
R^2	0.55	0.72	0.53	0.70	0.51	0.70

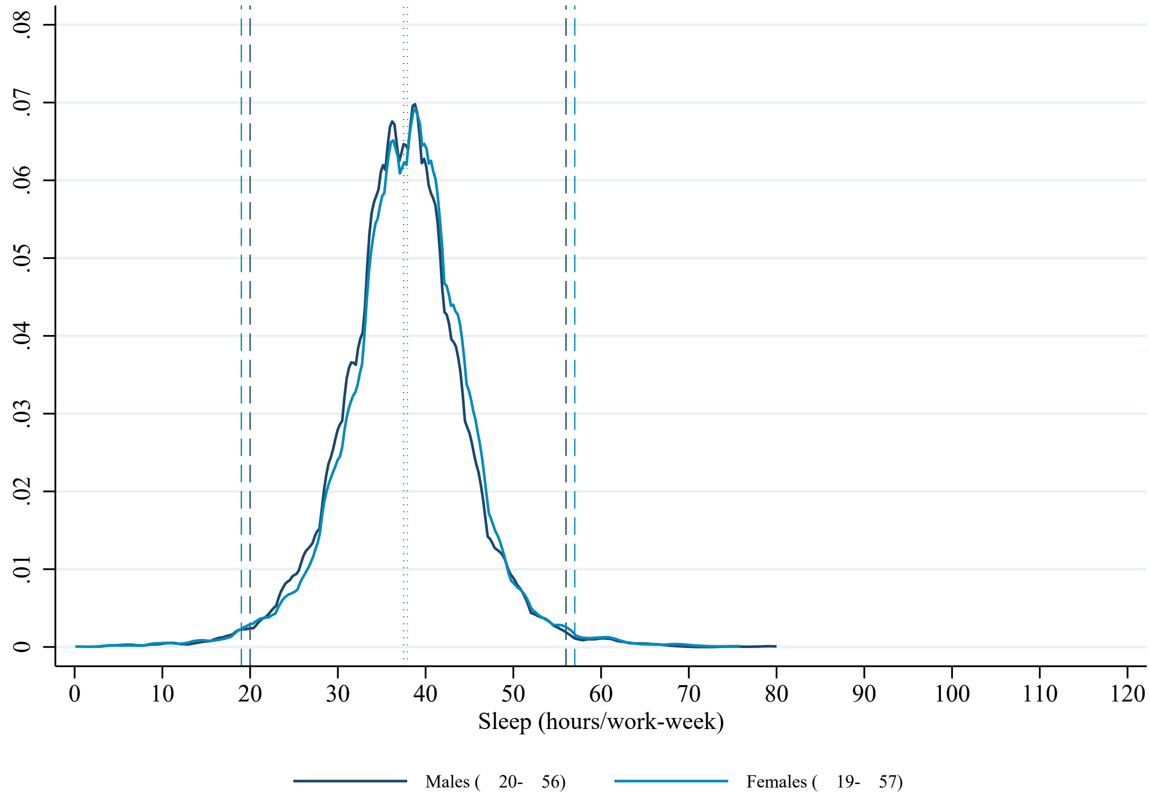
Notes: we use the same specification as in Table 2, but drop from the sample individuals older than 62 (top panel) or 60 (bottom panel). Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure B1: Work distribution by gender



Notes: Vertical dashed lines indicate the 5^{th} and the 95^{th} percentile of the work hours distribution for the selected sample (see Appendix A). Percentiles are specified in parentheses. The darker (lighter) color refers to the males' (females') distribution. Vertical dotted lines represent the medians of the two distributions, corresponding to 46,67 (43,54) hours per working-week for males (females).

Figure B2: Sleep distribution by gender



Notes: Vertical dashed lines indicate the 1^{st} and the 99^{th} percentile of the sleep hours distribution for the selected sample (see Appendix A) falling within the $5^{th} - 95^{th}$ percentile range of the gender-specific work hour distribution shown in Figure B1. Percentiles are specified in parentheses. The darker (lighter) color refers to the males' (females') distribution. Vertical dotted lines represent the medians of the two distributions, corresponding to 37,5 (37,92) hours per working-week for males (females).

Figure B3: Earnings distribution by gender

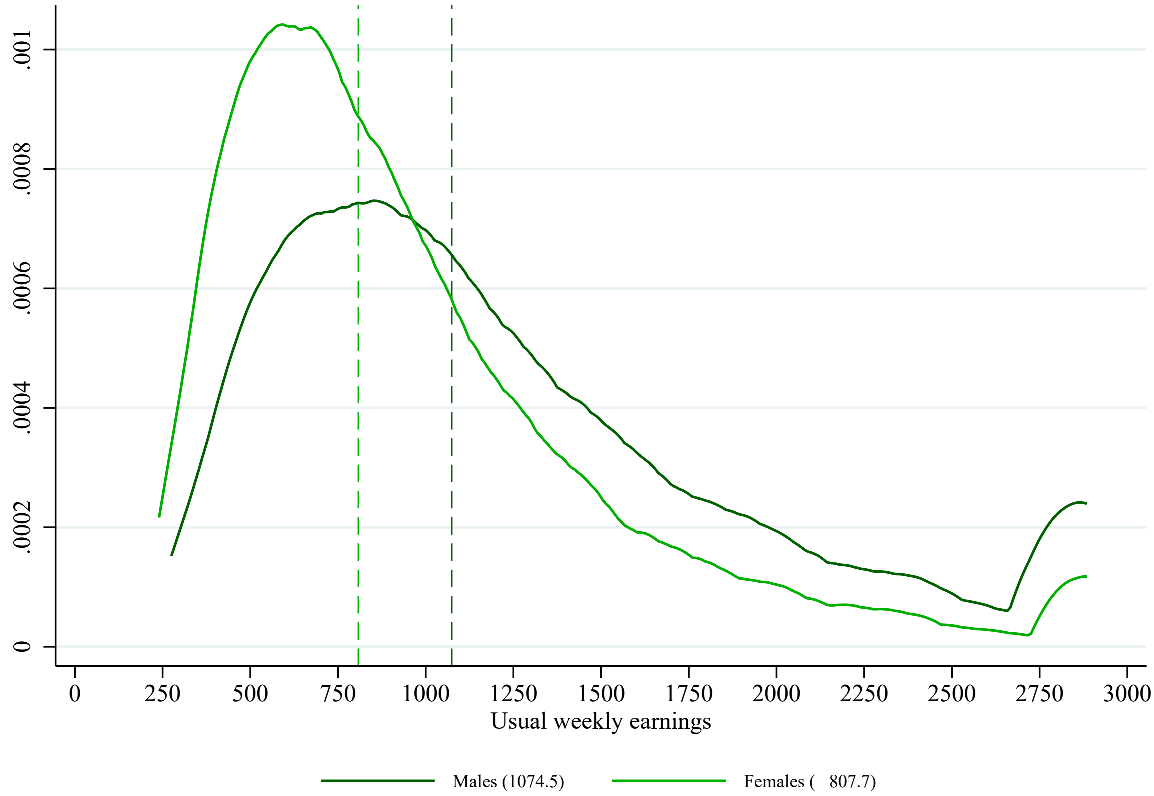


Figure B4: ln(earnings) distribution by gender

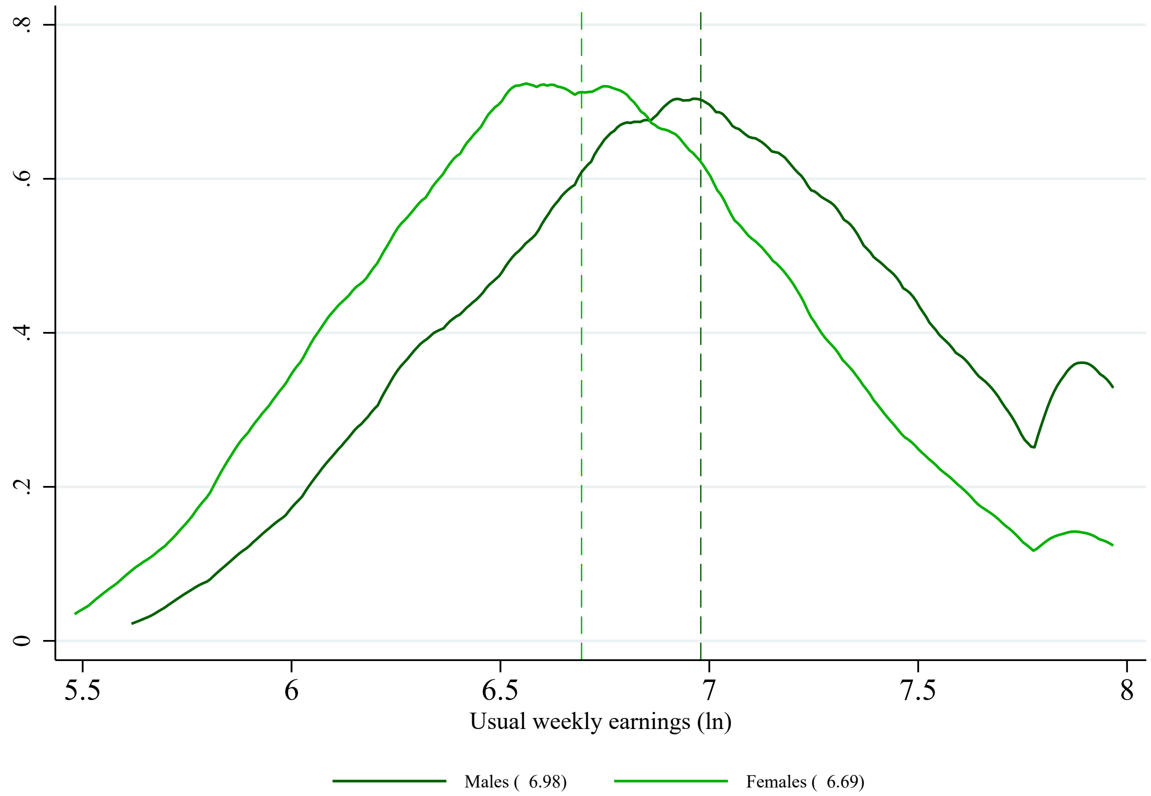


Figure B5: Fifths of the sleep distribution

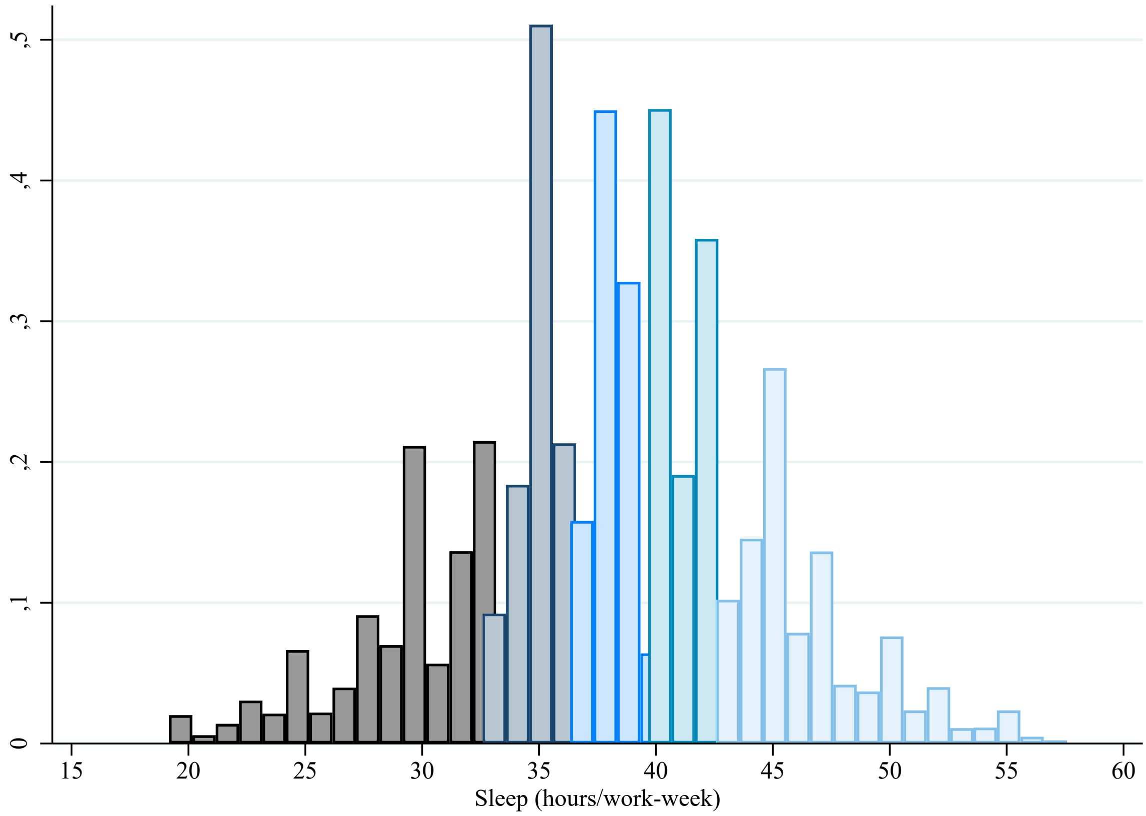
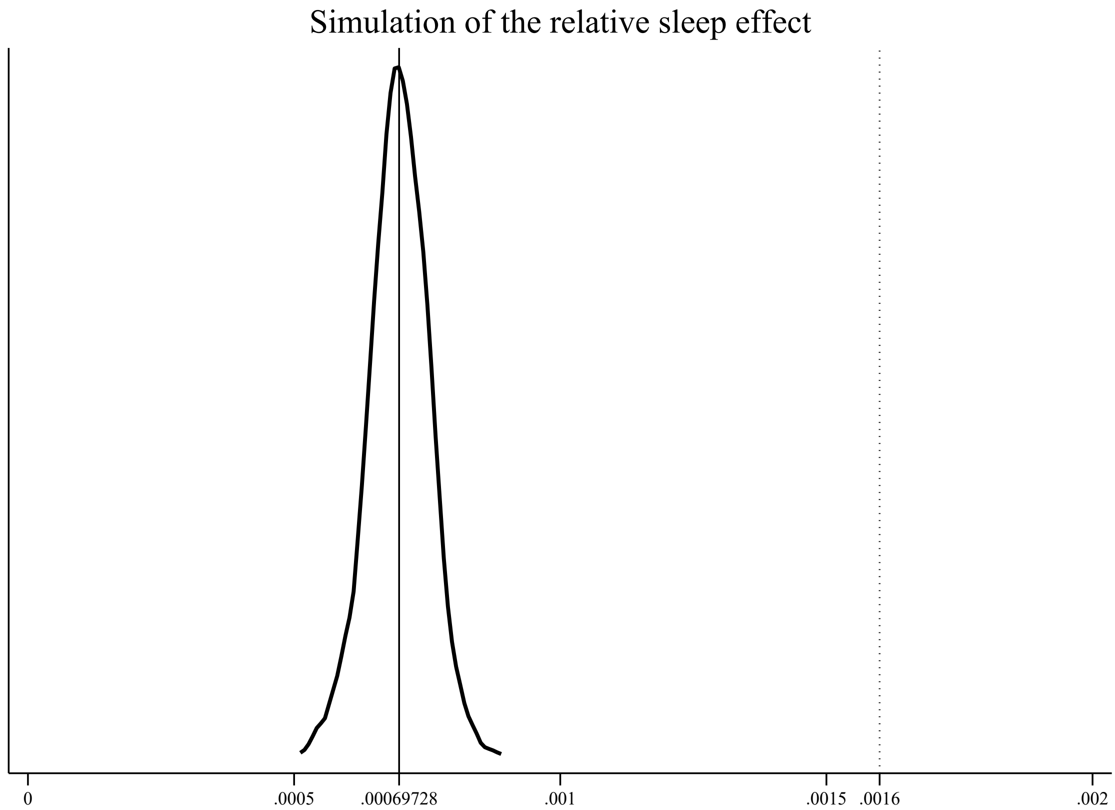
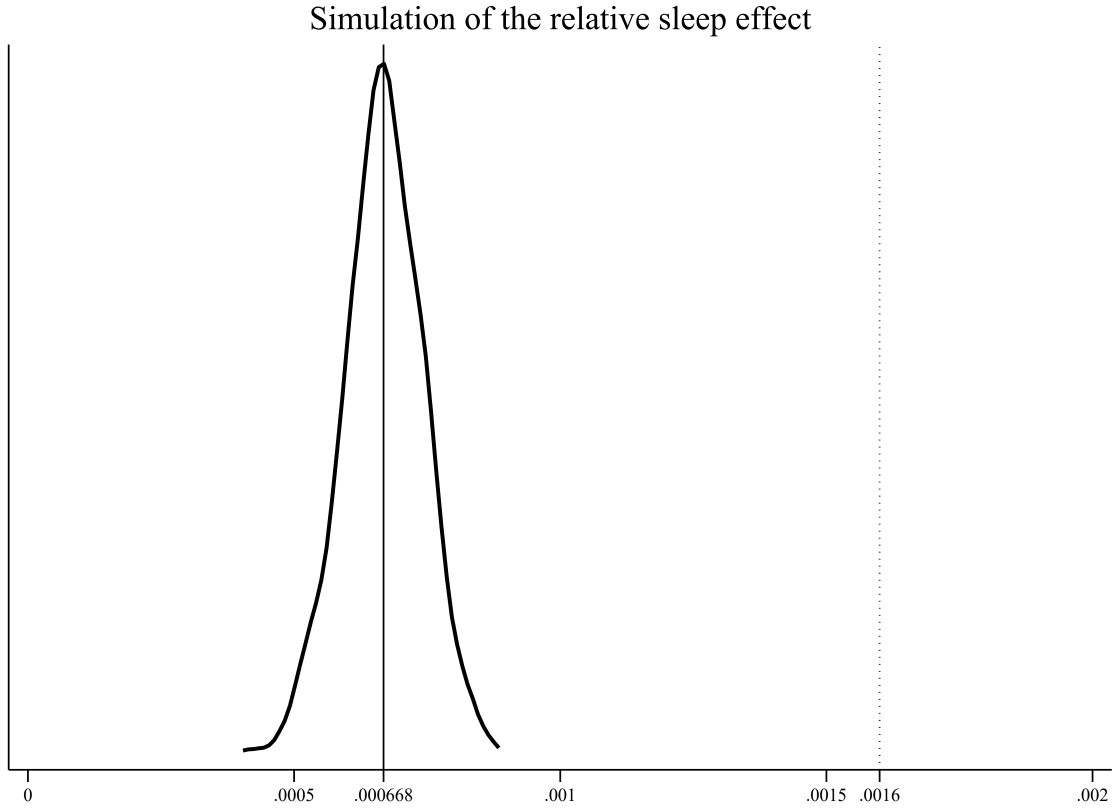


Figure B6: Simulation results to assess measurement error



Notes: The solid line represents the average estimated effect of one night sleep on the generated measure of productivity from 1000 draws. The dotted line represent the "true" effect of individual average sleep, known by construction. Parameters used to define the sleep distribution between and within individuals are taken from Jonasdottir et al. [2021], suggesting an average sleep duration of 7 hours per night, with a SD of around 1 hour, and an intra-individual variability in sleep during the work-week of about 1.1 hours.

Figure B7: Simulation results to assess measurement error



Notes: The solid line represents the average estimated effect of one night sleep on the generated measure of productivity from 1000 draws. The dotted line represent the "true" effect of individual average sleep, known by construction. Parameters used to define the sleep distribution between and within individuals are taken from [Willoughby et al., 2023], suggesting an average sleep duration of 412 minutes per night, with a SD of around 45, and an intra-individual variability in sleep during the work-week of about 53 minutes.