Time Use and Labor Productivity: The Returns to Sleep

Matthew Gibson and Jeffrey Shrader*

August 27, 2017

Abstract

We investigate how the single largest use of time—sleep—affects labor productivity. Motivated by a theoretical model, we provide empirical evidence that sleep is complementary to work in the short run and complementary to home production for non-employed individuals in both the short and long run. Using time use diaries from the United States, we show that later sunset time reduces worker sleep and earnings. After investigating these relationships and ruling out alternative hypotheses, we implement an instrumental variables specification that provides the first causal estimates of the impact of sleep on earnings. A one-hour increase in location-average weekly sleep increases earnings by 1.1% in the short run and 5% in the long run. (JEL No. J22,J24,J31)

*Corresponding author: Jeffrey Shrader, School of Law, New York University, 139 MacDougal Street, New York, NY 10012; jeffrey.shrader@nyu.edu. Matthew Gibson, Department of Economics, Williams College, mg17@williams.edu. The authors thank Prashant Bharadwaj, Julie Cullen, Gordon Dahl, Joshua Graff Zivin, Daniel Hamermesh, Erik Hurst, Mark Jacobsen, Craig McIntosh, Sendhil Mullainathan, Dale Squires, Michelle White, and seminar participants at UC San Diego for helpful comments. Shrader gratefully acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144086. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Word count: 13,024.

1 Introduction

Questions of labor productivity are fundamental to economics, important for both individual decisions and public policy. While there are traditions of research on the labor productivity effects of human capital (Becker, 1962, 1964) and health (Leibenstein, 1957; Mushkin, 1962), less attention has been paid to time use. Many types of time use, from reading to vacationing, plausibly impact labor productivity in both market and non-market work. In this study we examine the effects of the most common human time use—sleep.

Medical research indicates that sleep may play an important part in determining labor productivity. Tired doctors make more mistakes (Ulmer et al., 2009). Tired students perform worse on tests (Taras and Potts-Datema, 2005). Poor sleep impairs health (Cappuccio et al., 2010). Moreover for the average individual, sleep takes up more time than any other activity. Despite the manifest importance of sleep, economists have largely treated it as a biological phenomenon outside their purview. We investigate two important questions that have been overlooked almost entirely. First, how do individuals trade off sleep against other time uses? Second, how does sleep affect earnings and wages?

To examine the first question, we extend the time use model of Gronau (1977) to include productivity-enhancing sleep as a fourth time use in addition to labor, leisure, and home production. For workers, the model shows that if sleep improves wages more than it complements home production, then sleep and market work will be complements while sleep and home production will be substitutes. In the short run, there is evidence that this complementarity pattern holds, suggesting that sleep is relatively more beneficial for market work than home work among employed individuals. The model also allows us to test the productive sleep hypothesis on the sample of non-workers. If an individual does not work, then the model predicts that sleep and home production will be complements while sleep and leisure time will be substitutes. We find support for this hypothesis in both the short and long run.

Our model implies that answering our second question is difficult, because the relationship between sleep and wages is causal in both directions. Sleep increases labor productivity and higher wages raise the opportunity cost of sleep time. There is empirical evidence of the latter relationship—a pioneering study by Biddle and Hamermesh (1990) finds that higher wages are associated with less sleep—but the former relationship has not been studied. In addition, sleep may be correlated with unobservable worker characteristics that also influence wages.

Motivated by medical research on circadian rhythm, we resolve the endogeneity by using sunset time as a source of exogenous variation. In general the first-stage relationship is straightforward: earlier sunset causes workers to begin sleeping earlier, and because work and school start times do not respond as strongly to solar cues (Hamermesh et al., 2008), this earlier bedtime translates into more sleep. In fact, sunset timing provides two types of variation, short-run and long-run. In the short run, within a location, earlier sunset in winter induces longer sleep duration. In the long run, comparing two locations in the same time zone, the location farther east will experience earlier average sunset than the location farther west. As a consequence, residents of the eastern location will sleep longer. These two types of sunset variation provide two instruments for sleep.

To implement our empirical strategy, we geocode observations from the American Time Use Survey (ATUS). ATUS provides rich labor market information about individuals, a wealth of control variables, and detailed time use data from daily diaries. Using the diary date and location, we assign each observation a diary-date sunset time and an annual-average sunset time. We then use these variables to estimate the short and long-run effects of sunset time on sleep and earnings, controlling—in the case of the short-run estimates—for fixed location characteristics, year effects, and individual characteristics, and—in the case of the long-run estimates—for geographic characteristics (distance to the coast, latitude) and location-level demographic characteristics.

Consistent with our hypothesized first stage relationship, we find that later sunset time significantly reduces sleep duration. The reduced-form effect of sunset time on earnings is also consistent with our sleep hypothesis. Intra-annually, a one-hour increase in sunset time decreases worker earnings by 0.4%, while a one-hour difference in long-run average sunset time decreases worker earnings by 4.5%. We find these earnings changes stem almost entirely from wage changes, rather than changes in hours worked. If labor markets are competitive and workers are paid their marginal revenue product, these wage changes derive from productivity changes. Using alternative econometric specifications, we rule out a number of other, non-sleep hypotheses and obtain similar results from multiple data sets in addition to ATUS. These results suggest that the exclusion restriction required for instrumental variables estimates—that the effect of sunset time on earnings operates only through sleep, conditional on our control variables—is reasonable. Under this assumption, we exploit sunset-induced sleep changes to identify both short- and long-run earnings effects.

Our results show that a short-run, one-hour increase in location-average weekly sleep increases worker earnings by 1.1%. A permanent one-hour increase in weekly average sleep increases location-average earnings by 5%. These are, to our knowledge, the first causal estimates of how sleep affects earnings. Because our identification relies on location-level variation, these estimates should not be interpreted as individual effects. Both short and long-run estimates potentially include productivity spillovers across workers. In addition, our long-run estimate may include general-equilibrium effects induced by exogenously higher worker productivity. We find no evidence of nonlinearity in the sleep-earnings relationship.

Our study demonstrates that sleep is not just an economic curiosity, but rather a vital determinant of productivity and an important part of an individual's time allocation decision. A one-hour increase in a location's weekly mean sleep raises earnings by roughly half as much as a one-year increase in education (Psacharopoulos and Patrinos, 2004).[1] These results point to the large impact that non-labor market activities can have on labor market performance. By examining the largest use of human time, our study contributes to the time-use literature following Becker (1965). It complements the important work on the evolution of leisure time by Aguiar and Hurst (2007). Our study also contributes to the growing literature on how environmental forces influence worker productivity (Graff Zivin and Neidell, 2012) and to the broader productivity literature on factors like information technology (Bloom et al., 2012) and workplace practices (Black and Lynch, 2001).

The rest of the paper proceeds as follows: Section 2 presents a time use model with sleep as a choice variable, illustrating identification challenges, and discusses related literature. Section 3 presents the estimating equations and discusses our identification strategy. Section 4 describes the data used in the study. Section 5 reports the main results. First we test model predictions about trade-offs across sleep and other time uses. We then report the reduced-form and first-stage effects of sunset time on earnings and sleep, followed by an extensive set of robustness checks aimed at instrument validity. Finally we present instrumental variables estimates of the effect of sleep on earnings. Section 6 concludes.

[1]While the effect of increasing earnings for all workers in a location might differ from the partial-equilibrium estimate of (Psacharopoulos and Patrinos, 2004), the latter nonetheless provides an instructive benchmark.

2 Identifying the effect of sleep on productivity, wages and earnings

2.1 Previous research

Existing studies of the relationship between sleep and wages in economics are few and largely concerned with addressing the question of whether sleep should be treated as a choice variable rather than simply a biological necessity. Biddle and Hamermesh (1990) is the first paper to provide empirical evidence on this issue and remains one of the only empirical investigations of labor market impacts of sleep. The authors lay out a model with agents optimizing over sleep, work, and leisure time in an otherwise standard setting. While their theoretical model allows sleep to affect productivity, Biddle and Hamermesh do not focus on this relationship in their empirical work. Instead they emphasize the causal mechanism operating in the opposite direction, modeling sleep as a function of instrumented wage (see, for instance, Biddle and Hamermesh (1990) Table 6). Brochu et al. (2012) and Szalontai (2006) also estimate the impact of wage changes on sleep using more recent data from Canada and South Africa. Finally, Bonke (2012) has examined the impact of two chronotypes—whether the individual is a "morning" or "evening" person—on income, finding that morning types earn more.

Daylight savings time (DST) has been used in a variety of settings in economics as a proxy for sleep changes. For example, Smith (2016) finds the spring DST transition results in more automobile accidents and attributes the change to sleepiness behind the wheel. However, the short-term nature of any sleep change induced by DST limits its use in studying slower-moving outcomes like wages. Moreover, while the spring transition into DST reduces sleep by 40 minutes on the day of the change, the transition out of DST does not induce a change in sleep (Barnes and Wagner, 2009).

Medical studies concerned with the effect of long-term differences in sleep on health or mortality[2] are closest to our study in terms of time horizon. In recent years, a series of papers have assessed the impact of short-term sleep loss on laboratory tasks. These studies provide suggestive insight into how sleep might impact labor productivity.

Van Dongen et al. (2003) conducted the longest laboratory-controlled study on the relationship between sleep levels and cognitive performance. The researchers kept subjects in the lab for two weeks, placing them into groups receiving 4, 6, and 8 hours of sleep. The subjects were given daily tests of attention, memory, and cognition. The research found that, relative to the

[2]See for instance Cappuccio et al. (2010) and Krueger and Friedman (2009).

Table 1: Causal medical studies of sleep and performance

Study	Sleep change (hr/day)	Study duration (days)	Outcome	Elasticities (abs. value)
Belenky et al. (2003)	-4, -2, -1, +1	7	PVT speed	.7, .5, .7, 0
Cohen et al. (2010)	-2.5	21	PVT reaction time	18
Dinges et al. (1997)	-2.4	7	PVT lapses	6
Landrigan et al. (2004)	+.82	21	Serious medical errors	4.5
Lockley et al. (2004)	+.82	21	Attention failures	4
Van Dongen et al. (2003)	-4, -2	14	Memory task	3.3, 2.2
Vgontzas et al. (2004)	-2	7	PVT lapses	2.9
Mean magnitude				3.9

Notes: Table includes all studies that experimentally manipulated sleep for at least 7 days, drawing on reviews by Van Dongen and Dinges (2005) and Banks and Dinges (2007). Studies of complete sleep deprivation (e.g. Mckenna et al. (2007)) were excluded. PVT stands for psycho-motor vigilance test, described in Section 2.1.

8-hour group, the groups subjected to 4 and 6 hours of sleep performed progressively worse on all three tests. Moreover, the performance decline was linear across the treatments.

Intriguingly, the subjects' subjective assessments followed a different pattern, declining for a few days and then leveling off. Observed cumulative effects quickly achieved large magnitudes: after one week, subjects in the 6-hour group performed as badly as subjects who were deprived of sleep entirely for one night. This indicates that sleep reductions well within the observed range, continued over long periods of time, can have very large effects.

We review the evidence from Van Dongen et al. (2003) and similar medium-term causal studies in Table 1. Each study manipulated sleep duration by one to four hours per night, over periods of one to three weeks. In almost every case they find very large effects. The typical elasticity of task performance with respect to sleep duration is approximately four. Although these studies provide evidence that a positive relationship between sleep and labor productivity is plausible, the question of external validity remains. Laboratory elasticities sometimes come from error rates or reaction times, and the relationship between such measures and real-world task performance is unknown. Moreover medical studies typically limit subjects' scope for adaptation, often prohibiting caffeine use and requiring subjects to undergo assessments at particular times.

2.2 A productive sleep model

A formal treatment of time use in the presence of productivity-enhancing sleep clarifies the identification challenge stemming from reverse causality and generates predictions about time-use trade-offs for workers—a group for whom we can observe wages or earnings as a proxy for productivity—and non-working individuals.[3] These predictions, along with similar predictions for workers, can be used to test whether sleep is productive. The model shows that productive sleep implies sleep and work (either in the market or at home) can be complements. In contrast, if sleep were not productive, then it would be a substitute with the other time uses. We empirically investigate trade-offs between time uses in Section 5.1 and find that sleep is indeed complementary to labor and home production.

To model productive sleep, we present an extension of the time use model of Gronau (1977). The individual's problem is to maximize a utility function $u(Z(x, T_L))$ where Z is a home production function in the style of Becker (1965) that takes goods, x, and leisure time, T_L, to produce consumables. Assuming as usual that u is increasing and quasi-concave, we can ignore the utility function and consider the individual's problem to be one of maximizing Z subject to the constraints below.

Goods can either be purchased in the market, at a price normalized to 1, or produced at home. Denote market goods as x_M and home produced goods as x_H so that total goods are given by $x = x_M + x_H$. Work time is denoted T_N, and the individual can gain market goods by working at wage $W(T_S) = \alpha + w(T_S(\theta))$, which is a function of an exogenous parameter, α, and time spent sleeping, $T_S(\theta)$. Sleep time is chosen by the individual but it is influenced by an exogenous price, denoted θ, which one can think of as sunset time. We assume that a higher price (or later sunset time) reduces sleep time, so $dT_S(\theta)/d\theta < 0$. Since we are modeling sleep as productivity-enhancing, we also assume that more sleep will, *ceteris paribus*, increase wage, so a change in wage due to an exogenous change in θ, $\frac{dW}{d\theta} = \frac{\partial W(T_S)}{\partial T_S}\frac{dT_S(\theta)}{d\theta} := w'(T_S)T_S'(\theta)$, will be negative. For convenience, T_S will be written with θ suppressed unless we explicitly want to model a change in sleep time due to an exogenous price change. The individual has non-labor income V so that $x_M = W(T_S)T_N + V$.

Home goods, x_H, are produced using the production function $x_H = f(T_H, T_S)$, where T_H is home production time. Given the assumption that sleep is work-productive, we naturally also assume that the change in home production with respect to sleep, f_2', is positive. Assume $f(0, .) = 0$ so that $T_H = 0$ is equivalent to $x_H = 0$.

Putting all time uses together, the total time constraint is $T = T_L + T_H + T_N + T_S$. Finally,

[3] In general earnings reflect both work hours and productivity. In what follows, we provide evidence that the earnings changes we estimate stem almost entirely from wage changes, rather than hours changes. This need not be true in other settings.

assume that sleep cannot be used as leisure to produce goods in Z.[4] Substituting the time budget into the goods budget, the combined budget constraint is

$$x_M + W(T_S)(T_H + T_L + T_S) = W(T_S)T + V$$

and the optimization problem is

$$\begin{aligned} \max_{T_L, T_H, T_S, x_M} \quad & Z\left(x_M + f\left(T_H, T_S\right), T_L\right) + \lambda_1 \left(W(T_S)T + V - x_M - W(T_S)\left(T_H + T_L + T_S\right)\right) \\ & + \lambda_2 x_M + \lambda_3 T_H + \lambda_4 T_S \end{aligned}$$

This is a concave problem since $W(T_S)T > W(T_S)(T_H + T_L + T_S)$.

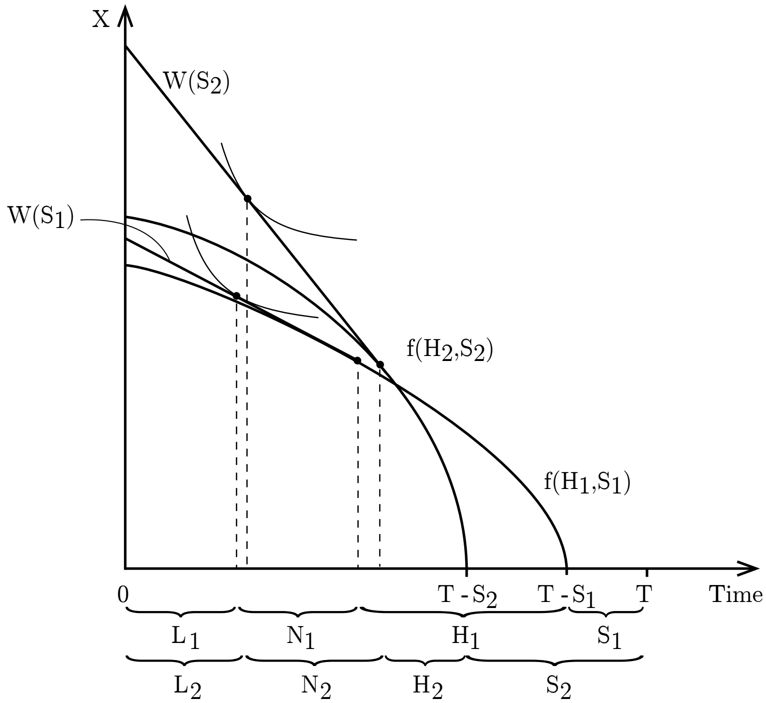
The main predictions of the model relative to the case where sleep is not productive are, first, that sleep can be complementary to either home production or labor for employed individuals, and second, that sleep can be complementary to home production for individuals who do not work.

We can get intuition for these results by utilizing Gronau's graphical device of representing all activities in just two dimensions: goods and net-of-sleep time. In Figure 1 the home production function, $f(T_H, T_S)$, should be viewed as a projection onto the $(x_H, -T_H)$ plane for a given value of T_S. The first order conditions show that the wage function will be tangent to the home production function. This point of tangency is where a worker will choose to divide time between leisure and work on the left and sleep and home production on the right. The tangency between utility and wages will determine the division between leisure and market work.

Changes in sleep influence the slope of the wage line, the steepness of the home production function, and the amount of time available for other time uses. The figure depicts two possible values for sleep: $T_{S,1}$, a low value, and $T_{S,2}$, a high value. By assumption, an increase in sleep raises the slope of the wage line. If this were the only effect, the wage, $W = \alpha + w(T_{S,2})$ would be tangent to the home production function $f(T_{H,1}, T_{S,1})$ at a point further to the right, causing the worker to expand total time devoted to work and leisure. Sleep also raises the productivity of home production, however, causing a reduction in time devoted to leisure plus work. For the changes shown in the figure, the wage increase from sleep is large enough to lead to an overall increase in the amount of market work and a decrease in the amount of home production. This

[4] We discuss relaxation of this assumption at the end of this section.

Figure 1: Productive sleep



Notes: A graphical depiction of the productive sleep model in the style of Gronau (1977, Figure 1). The figure depicts two possible choices for levels of sleep, showing that increasing sleep raises both work and home productivity, at the cost of taking time away from the total time budget. Time uses are denoted by their subscripts in the model for legibility.

provides intuition that in cases where sleep is especially beneficial for productivity in market work, sleep and labor will be complements. A similar condition holds for home production, as the following results make clear.

To formalize the time-use predictions, consider an individual who works in the market and at home. This case corresponds to $(x_M > 0, T_H > 0)$, so both λ_2 and λ_3 are 0. Assuming that sleep time is positive, $f_1' > 0, f_2' > 0, f_{11}'' < 0$, and $f_{22}'' < 0$, then the first order conditions can be written as

$$\frac{Z_2'(x, T_L)}{Z_1'(x, T_L)} = W(T_S) \tag{1}$$

$$f_1'(T_H, T_S) = W(T_S) \tag{2}$$

$$f_2'(T_H, T_S) + T_N w'(T_S) = W(T_S) \tag{3}$$

plus budget constraints. These equations implicitly define optimal time use quantities in terms of parameters, including sunset time, θ.

Taking a total derivative of Equation (2) with respect to θ and re-arranging yields,

$$\frac{dT_H}{d\theta} = \left(\frac{f''_{12}(T_H, T_S) - w'(T_S)}{f''_{11}(T_H, T_S)} \right) \frac{dT_S}{d\theta} \tag{4}$$

By assumption, $\frac{dT_S}{d\theta} < 0$, so the right hand side will be positive if $w' > f''_{12}$. In words, if the marginal increase in wages with respect to sleep is larger than the complementarity between sleep and home production, then home production will rise in response to a later sunset time. If the opposite condition holds, then later sunset time will decrease both sleep and home production time.

To see that sleep and work can also covary in the model, consider Equation (3). Again take the total derivative with respect to θ to find

$$\frac{dT_N}{d\theta} = (w'(T_S))^{-1} \left(\left(w'(T_S) - f''_{22}(T_H, T_S) - T_N w''(T_S) \right) \frac{dT_S}{d\theta} - f''_{12}(T_H, T_S) \frac{dT_H}{d\theta} \right)$$

Assuming that $W_{11} \leq 0$, if $\frac{dT_H}{d\theta} > 0$ and $f''_{12} > 0$, then the whole term will be negative.[5] Therefore, a later sunset time would lead to less sleep and less work. So, for a worker, depending on f''_{12} and w', productive sleep can be complementary to either home production or work time.

For an individual who does not work outside the home, the predictions are starker than for a worker: so long as the labor productivity gain from more sleep is not so large that it induces the person to work, then home production should increase with an increase in sleep, while leisure time should fall. The formal statements for these relationships can be derived from the first order conditions, taking $\lambda_3 = 0$ and $\lambda_2 > 0$. In particular, in this model, an individual chooses not to work if the derivative of the home production function with respect to home production time is steeper than the market wage rate. At the margin, a change in sleep will not change this condition, so it must be the case that sleep and home production will be complements by the same logic used to sign Equation (4) above.

[5]Medical studies of the effect of sleep on health often find a nonlinear relationship (for example Cappuccio et al. (2010)) that suggests $W_{11} < 0$). Assuming $W_{11} = 0$ might be reasonable, at least for some range of sleep hours. Van Dongen et al. (2003), for instance, finds that performance on attention tasks declines linearly with sleep deprivation of up to 4 hours per night. To the limited extent that we can test for it in our empirical setting, we do not find evidence for nonlinearity in the sleep-wage relationship.

Contrast these predictions with a model of sleep that is not productivity-enhancing. Suppose that sleep is just a biological necessity—it grants neither utility nor productivity. In such a case, sleep only removes time that could have been devoted to other tasks. A later sunset time will decrease sleep, causing a pure income effect for the individual. Leisure, work, and home production time will rise, with the exact split between these time uses dependent on the production function, Z. Therefore, sleep will be a substitute for all other time uses.

If sleep time can be used for consuming goods but does not improve productivity, then it is indistinguishable from leisure. Therefore, sleep and leisure will be substitutes, and from Figure 1, one can see that inducing a worker to consume more leisure will decrease work time but will leave home production time unchanged. Therefore, such a model would predict that sleep and work would be substitutes while sleep and home production would be unrelated (at the margin). Neither of these models generate the prediction that sleep will be complementary to either work or home production.

Finally, we can use the model to motivate our empirical method by highlighting the potential reverse causality stemming from an increase in wages. Applying the implicit function theorem to the first order conditions, one can formally find the effect of an exogenous change in α on sleep. The expression for this change can be found in Appendix Section B.1. Intuitively, if sleep is relatively more important for work than for home production, then a dominant income effect coming from increased wages can cause the agent to reduce sleep in exchange for more home production. A naïve regression of wages on sleep would yield a negative coefficient in this case, even if sleep were productive.

3 Empirical strategy

3.1 Estimating equations

We would like to recover the relationship between sleep and wages, where $\partial W/\partial T_S > 0$ would provide evidence for productivity-enhancing sleep. Given the reverse causality between wages and sleep, however, we might erroneously find $\partial W/\partial T_S < 0$. To avoid this problem and to account for the wide variety of other omitted variables that might co-vary with sleep and wages, we predict sleep using two instruments based on local sunset time, then use the instrumented values of sleep to estimate earnings and wage impacts.

The first instrument uses daily variation in sunset within a given location. Because this instrument varies on a daily basis, it will identify short-run variation in sleep (Frazis and

Stewart, 2012). Thus, we can use it to estimate a short-run first stage

$$T_{S,ijt} = \alpha_1 \text{sunset}_{jt} + \gamma_{1,j} + \mathbf{x}'_{it}\delta_1 + \eta_{1,ijt} \tag{5}$$

and reduced form

$$\ln(W_{ijt}) = \alpha_2 \text{sunset}_{jt} + \gamma_{2,j} + \mathbf{x}'_{it}\delta_2 + \eta_{2,ijt} \tag{6}$$

where $T_{S,ijt}$ is nighttime sleep for individual i in location j on date t, sunset_{jt} is the sunset time on that date in that location, γ_j is a location fixed effect, $\mathbf{x}_{it}$ is a vector of individual level controls, W_{ijt} is a measure of wages or earnings observed at time t, $\eta_{1,ijt}$ is the error term for the first stage, and $\eta_{2,ijt}$ is the error term for the reduced form. Controls are four race indicators; age and age squared; a gender indicator; indicators for holidays, day of week, and year; and detailed occupation code indicators. More details on these control variables can be found in the data discussion in Section 4. Following the suggestions of Winship and Radbill (1994) and Solon et al. (2015), we do not weight observations, but we do control for weekends since they are over-sampled in our dataset.

If seasonal sunset time is a valid instrument for sleep, then this first stage and reduced form can be used to construct causal estimates of the effect of sleep on earnings or wages by taking the ratio of α_2 to α_1. In practice, we will calculate the instrumental variables estimator using two-stage least squares. We denote this estimate β_{SR}, or the effect of sleep in the short run.

The second instrument is annual average sunset. This instrument exploits spatial differences in sunset time within and across time zones. Because this is a fixed feature of a location, it will identify long-run differences in sleep (Frazis and Stewart, 2012). For estimation, we collapse the ATUS data to the location level. This serves to emphasize that variation in long-run sunset time is permanent and common to all workers in a location. We then estimate the following first stage

$$T_{S,j} = \varphi_1 \text{sunset}_j + \mathbf{x}'_j\zeta_1 + \varepsilon_{1,j} \tag{7}$$

and reduced form

$$\ln(W_j) = \varphi_2 \text{sunset}_j + \mathbf{x}'_j\zeta_2 + \varepsilon_{2,j} \tag{8}$$

where $T_{S,j}$ is average nighttime sleep in location j, sunset_j is the average sunset time in that location, $\mathbf{x}_j$ is a vector of controls, W_j is average earnings or wage in that location, and $\varepsilon_{k,j}$ is an error term for $k \in \{1, 2\}$. We control for both geographic characteristics (coastal distance and latitude) and demographics (gender, age, race, and occupation shares, plus population density).

Following the recommendation in Solon et al. (2015), we weight location-level observations using counts of the underlying individual ATUS observations to correct for heteroskedasticity. Appendix Section C.3 provides evidence of heteroskedasticity from a modified Breusch-Pagan test, presents unweighted results, and explores differences between weighted and unweighted results.

Again, if average sunset time is a valid instrument, the causal effect of long-run changes in sleep can be found by taking the ratio of φ_2 to φ_1. We denote this coefficient β_{LR} and estimate it by two-stage least squares. Although the control variables primarily serve to reduce residual variance, as discussed in Section 5, we do find evidence that average sunset is not unconditionally exogenous with respect to coastal distance and population density (see Appendix Tables 22 and 23). Therefore the identifying assumption underlying our long-run IV estimates is one of conditional exogeneity, as discussed in Sections 3.2.3 and 5.3.2.

3.2 Local sunset time instruments

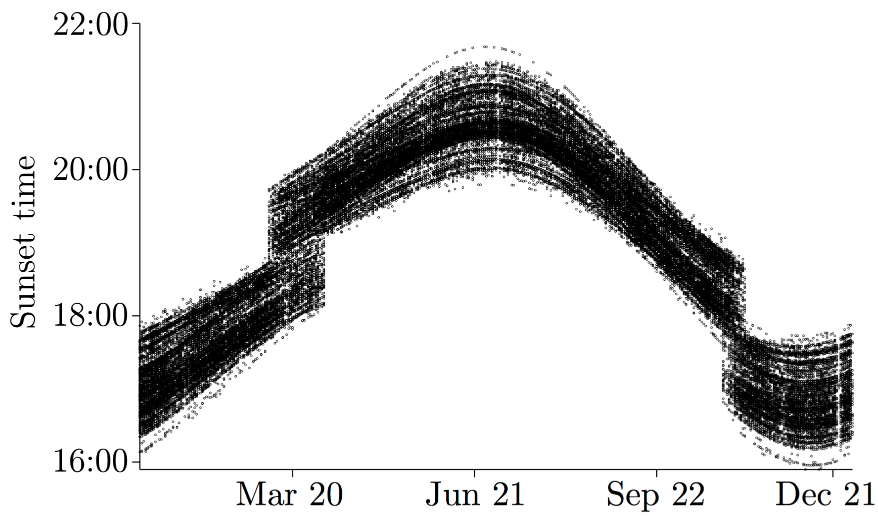
3.2.1 Relevance of sunset to sleep

The relevance of sunset time as an instrument for sleep stems from the biological relationship between sleep patterns and daylight. Human circadian rhythm is synchronized with the rising and setting of the sun through a process known as entrainment. This force is powerful, with Roenneberg et al. (2007) showing that "the human circadian clock is predominantly entrained by sun time rather than by social time." Using data from Germany, the authors demonstrate that living in a location with a later sunset induces individuals to begin sleep later. The detailed ATUS files enable us to reproduce this result: workers experiencing a later sunset go to bed later and this causal connection between sunset and bedtime persists even if the worker goes to bed well after dark. Intra-annual changes in sunlight also influence human sleep patterns through a similar process of entrainment (Hubert et al., 1998). In a vacuum, a later sunset time might cause workers to go to bed later and also rise later, leaving sleep duration unchanged. But workers face morning coordination constraints due to work and school scheduling (Hamermesh et al., 2008), so later sunset and later bedtime decrease sleep duration, both intra-annually and in the long run. Our estimates in Section 5.2 are consistent with these hypotheses. The evidence bearing on instrument validity is different for the short and long run. We discuss the two cases separately below.

3.2.2 Short-run validity

For our short-run estimates of the effect of sleep on wages and earnings, validity requires that other wage determinants not co-vary with daily sunset time within a location. The primary threat to this assumption is seasonally varying wage determinants, since sunset time exhibits a regular seasonal pattern, as shown in Figure 2 for all ATUS observations. One can see that sunset time is generally described by a cosine wave with a period of one year. This wave is phase shifted by roughly 10 days relative to the calendar year. The amplitude of the wave is determined by the latitude of the location, and vertical translations are due to within-time zone variation, which we use for our long-run estimate. The final important features of sunset time are the prominent jumps in the spring and fall caused by daylight savings time. A 2005 change in the timing of daylight savings observance makes these jumps non-sharp.[6]

Figure 2: Sunset time for locations sampled in ATUS



Notes: Vertical axis measured in 24-hour time (for example, 19:00 is equivalent to 7PM). Each point shows the sunset time for a location sampled by ATUS within the continental United States.

Knowing the exact pattern followed by sunset time allows us to characterize the degree of potential confounding from other seasonal variables. These calculations are given in Section 5.3.1. Because of daylight savings time and the phase shift relative to the calendar, our estimates

[6]For more discussion of solar mechanics, see Section A.

are robust to a wide range of seasonal confounders. Moreover, these features allow us to clearly disentangle short-run variation in sunset from calendar features like the December shopping season.

Our productivity measures are usual weekly earnings and usual hourly wage, rather than wages on the day of the interview. This raises an additional identification issue inherent in studying earnings and wages rather than productivity: timing mismatch between observations of sleep and wages combined with a potentially low-frequency relationship between wages and productivity. These issues mean that our short-run estimates of both the reduced form Equation 6 and the resulting two-stage least squares estimator will be attenuated. In appendix Section B.2, we derive an expression for asymptotic bias that depends only on the frequency of wage or earnings changes and a trigonometric function derived from solar mechanics. Using that expression we bound the degree of attenuation between 0.25 and 1, so we expect *a priori* that our short-run estimates should be between 0 and one-quarter of the true parameter value.

There is one short-run identification issue we cannot address: seasonal variation in sunset time is almost perfectly correlated with seasonal variation in sunrise time and daylight duration.[7] Therefore in purely statistical terms, all short-run results could be recast in terms of either of these other variables. Our interpretation of the short-run results could be incorrect if daylight affects both earnings and sleep through mood, the hedonic value of leisure, or some other channel.[8] Medical studies find positive effects of daylight duration on mood (Murase et al., 1995; Lambert et al., 2002), so if this is the dominant confounder our earnings and wage estimates will be biased downward in magnitude. We focus on sunset time rather than sunrise time because it is emphasized by existing medical literature and because it is the driver of long-run differences in sleep, as discussed in the next section.

3.2.3 Long-run validity

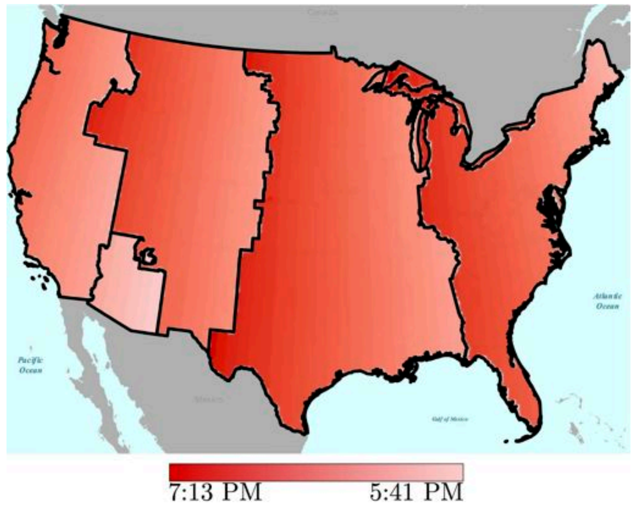
Figure 3 illustrates long-run variation in sunset time across locations. As the sun sets, eastern locations grow dark earlier than western locations, leading residents in more easterly locations go to bed earlier and sleep longer. By design, the maximum difference in sunset time within a US time zone is approximately one hour.

The difference in average sunset time between two locations over the year is plausibly

[7] Only changes in daylight savings time break this linkage.

[8] Daylight duration does not create an analogous problem for the long-run estimates, as all locations in the continental United States experience nearly the same average amount of daylight.

Figure 3: Average sunset in the United States



Notes: Map shows sunset time at the vernal equinox for the continental United States in 2012, which is a close approximation to average sunset time. Darker red indicates later sunset, lighter red indicates earlier. The time zone boundaries are given by bold black lines.

unrelated to other factors influencing the labor market. In particular, time zone boundaries break the link between average sunset time and longitude. Average sunset time is also, by construction, orthogonal to latitude. All locations in the continental US experience approximately the same average daylight duration over the year, so this is not an omitted variable in our long-run analysis, which is restricted to this geographic area. If average sunset time shifted the timing of work within the day, and if workers were more productive at particular times of day, that could violate the exclusion restriction. Hamermesh et al. (2008), however, find that the response of work schedules to sunrise and sunset times on the day of an ATUS diary is extremely small and not statistically significant. This suggests work schedules do not adjust to sunset time in either the short or long run.

Railroads implemented the first US time zones, called Standard Railroad Time (SRT), on November 18, 1883. They replaced a patchwork of railroad time standards and were quickly adopted by the US government and Western Union (Allen, 1883; Anonymous, 1883). While railroads were the first adopters, the primary impetus for standard time and the zone plan itself came from scientists concerned with problems like simultaneous observation of the aurora borealis at different points across the US (Bartky, 1989). The width of a zone, 15 degrees of longitude, was chosen to correspond with a one-hour difference in solar time (Library of Congress, 2010).

Endogenous modifications to time zone borders could have undermined this initial randomization. State and local governments may petition the Department of Transportation (DOT) to switch time zones. The DOT criterion for evaluating proposed time zone changes is "the convenience of commerce."[9] This process means that the precise location of the boundary is endogenous and a regression discontinuity design comparing nearby communities on opposite sides of the boundary could be biased. Moreover the integration of labor markets across time zone borders in some counties could pose identification problems. In Table 7 we show our results are robust to the exclusion of all counties on time zone borders. To avoid potential endogeneity, we drop locations that do not observe daylight saving time. Finally, while time zone borders often coincide with state borders, they frequently do not, and twelve of the lower 48 US states span multiple time zones (Hamermesh et al., 2008).

Current or past worker sorting on sunset time could threaten instrument validity if such sorting induced correlation between average sunset time and characteristics that influence worker productivity. We investigate this possibility in Appendix Section C.3.2 and find no evidence for it. Potentially more problematic in our finite sample is spurious correlation between the instrument and location-level determinants of productivity. We do find a statistically significant relationship between average sunset time and population density, which motivates our use of a flexible control for this variable. Within a time zone, long-run average sunset time is a linear function of longitude and thus also correlated with coastal distance. Table 7 suggests coastal distance is indeed an important potential confounder, and we control flexibly for it in all our specifications. Conditional on our flexible controls, we find no evidence that early and late-sunset workers differ on observables. The identifying assumption in our long-run models is exogeneity of average sunset time conditional on these controls. Having found two important variables correlated with our instrument, it is natural to ask whether there might be others. We cannot exclude this possibility. In Section 5.3.2 we investigate a wide range of possible confounders and find our results are robust.

[9]For details on the DOT process, see Valpando (2013) and Appendix Section C.3.3.

4 Data

The largest data set containing both sleep time and wage information is the American Time Use Survey, administered by the US Bureau of Labor Statistics (BLS) since 2003. ATUS includes a random sample drawn from households that have recently completed participation in the Current Population Survey (CPS). For example, if a household participated in the CPS from January through April 2013 and January through April 2014, it would be eligible for ATUS sampling in June, July, or August 2014. BLS employs a three-stage sampling procedure. First, the bureau allocates observations to states proportional to their populations. Second, it stratifies on "race/ethnicity of the householder, the presence and age of children, and the number of adults in adults-only households" (BLS, 2015). There is some oversampling of households where the householder identifies as Hispanic or non-Hispanic black, and of households with children. Third, within each household, ATUS respondents are chosen randomly from among individuals aged 15 or older. All such individuals within a household have equal probability of being selected. This sample is allocated uniformly across 12 months and then across 4 weeks within each month. Within each week, 10 percent of subjects are randomly selected for interviews on each weekday, and 25 percent on each weekend day (BLS, 2015). While this design does not enforce a uniform distribution of interviews over the year within location, it does imply that the distribution within a location is uniform in expectation. The use of telephone interviews rules out potential sampling biases from travel costs or the locations of government offices. Appendix Figure 9 confirms that interviews are, on average, uniformly distributed across months of the year within location, and that random deviations from uniformity are small in magnitude.

BLS collects the data using computer-assisted telephone interviews, which cover the period from 4AM on the previous day to 4AM on the interview day. By employing a very short recall period and forcing all time uses to sum to 24 hours, this method minimizes the possible biases associated with time diaries (Hamermesh et al., 2005). Interviewers use conversational techniques that have been shown to reduce recall bias in laboratory settings (Schober and Conrad, 1997). For each time use, the interviewer records either the duration or the end time. In most cases the interviewer records the respondent's description verbatim and it is coded later. Codes are very detailed, distinguishing for example between watching hockey and basketball on television. Because the sample is drawn from CPS respondents, CPS demographic and labor variables are available for almost all ATUS respondents. Each respondent participates in ATUS only once, however, so it is not possible to construct an individual time-use panel. For a more detailed description of ATUS, see Hamermesh et al. (2005) and BLS (2015). For the analysis of employed individuals, we use the sample of prime age individuals who report receiving positive weekly wages from a primary or secondary job and who work full time.[10] Summary statistics for the employed sample, broken down by sunset time, are given in Table 2.

Table 2: ATUS Summary Statistics

	Early sunset Mean/(SD)	Late sunset Mean/(SD)	Diff./(SE)	Total obs.
Weekly earnings ($/week)	965.4	954.2	11.2**	61,161
	(626.0)	(616.2)	(5.02)	
Hourly wage ($/week)	16.6	16.5	0.10	30,925
	(9.32)	(9.10)	(0.10)	
Sleep (hour/week)	58.4	57.6	0.82***	61,161
	(14.4)	(13.9)	(0.11)	
Sunset time (24 hr)	17.6	20.1	-2.45***	61,161
	(0.75)	(0.54)	(0.0053)	
Work (hour/week)	31.1	30.8	0.27	61,161
	(31.0)	(31.0)	(0.25)	
Female (0/1)	0.48	0.48	0.0043	61,161
	(0.50)	(0.50)	(0.0040)	
Age (years)	41.6	41.6	0.053	61,161
	(10.8)	(11.0)	(0.088)	
Race, white (0/1)	0.82	0.82	-0.0038	61,161
	(0.39)	(0.38)	(0.0031)	
Race, black (0/1)	0.12	0.12	0.000096	61,161
	(0.33)	(0.33)	(0.0027)	
Weekend (0/1)	0.51	0.51	0.0046	61,161
	(0.50)	(0.50)	(0.0040)	
HS or less (0/1)	0.31	0.32	-0.0075**	61,161
	(0.46)	(0.46)	(0.0037)	
College (0/1)	0.25	0.25	0.0019	61,161
	(0.43)	(0.43)	(0.0035)	
Number of children	0.98	0.97	0.0052	61,161
	(1.12)	(1.12)	(0.0090)	

Notes: Summary statistics for two sub-samples from ATUS are shown. Early sunset is defined as having a sunset time earlier than the median, and late sunset time is later than the median. Significance is determined from a t-test on the difference between means. Total observations are given in the far right column. The early and late sunset time groups each have half of the stated observations. Average work hours are approximately 31, despite the restriction of the sample to full-time workers, because half of ATUS diaries come from weekends.

Aside from giving basic information on the sample, Table 2 also provides initial evidence

[10]We exclude individuals in locations that do not observe daylight savings time, as local sunset time is potentially a choice variable in such locations.

in support of our main results.[11] Early sunset observations have significantly higher earnings and sleep duration than observations with later sunset times. In contrast, other individual characteristics are well balanced across the two groups. Only one difference is significant—the fraction of the population with a high school degree or less.[12]

To assign locations to individuals in ATUS, we first merge ATUS data with the corresponding CPS data. For a given individual, the CPS data often contain location at the county level. This variable is censored for individuals living in counties with fewer than 100,000 residents. When county is available, we assign the county centroid as an individual's location. We have county location for approximately 44% of ATUS observations. For an additional 28% of observations, we observe location at the level of Census CBSA, a small group of counties in the same metropolitan area. In total, we are able to geocode 72% of observations at the sub-state level.[13] For the remaining 28% of observations, ATUS contains location at the state level. We assign the 2010 population-weighted state centroid (computed by the Census) as the location for these individuals. Where we refer to Federal Information Processing Standards (FIPS) codes, we are referring to either the county (FIPS 6-4) or CBSA-level code, if available, or the state level code (FIPS 5-2) where more detailed location is unavailable. Using the interview date and respondent location, we are able to determine sunset time for each individual in the dataset using solar mechanics algorithms from Meeus (1991). We compute annual average sunset time by computing sunset for each day in an individual's location, then calculating the mean over days of the year.

Our sleep variable is nighttime duration from the ATUS diary, multiplied by seven to obtain a weekly measure. We employ a weekly rather than a daily sleep variable to match the frequency of our earnings variable. We remove any sleep that starts and ends during daylight hours on the date of diary entry. This will exclude naps, which might be an adaptation strategy for some short sleepers, however it also removes night-shift workers, for whom the sunset instrument should not be relevant. Empirically, our point estimates are practically unchanged

[11]Note that dividing sunset time in this way conflates short-run variation (within location over time) and long-run variation (across locations). In the regression models, we will explicitly break up these two sources of variation.

[12]This difference works in the direction of explaining the difference in wages in the two groups, but other (insignificant) differences work in the opposite direction. Results controlling for individual characteristics and broken down by short versus long-run sunset time are reported along with a non-parametric version of this binary division in Section 5.

[13]Even for county-geocoded observations, there will be measurement error in location and therefore sunset time. This could attenuate both first-stage and reduced-form estimates, but under classical assumptions will not attenuate IV estimates. County size is roughly constant within time zone, but greater in the Pacific and Mountain zones than in the Central and Eastern zones.

by the exclusion of daytime sleep, but precision of the first stage increases substantially.[14]

Our primary labor productivity measure is “usual weekly earnings” as reported in ATUS. While this measure could capture both changes in wages and changes in hours worked, our earnings results largely reflect wage changes, as discussed in Sections 5.1 and 5.3. Work time responds less than sleep time to sunset, and explicitly holding work hours fixed does not substantially change our estimates. The usual weekly earnings variable is defined for all respondents who have positive labor income and are not self-employed. It is top-coded above \$2,884.61. We also estimate a version of our model including only workers who receive an hourly wage, “hourly earnings at main job” as reported in ATUS. This variable is likewise top-coded at the level such that hourly earnings multiplied by usual weekly hours equals \$2,884.61. In the primary estimation sample, 52% of individuals report an hourly wage. Some control variables (e.g. occupation codes) appear in both ATUS and CPS files. Where possible we use ATUS variables, which are more recent. Our preferred regression specifications include a set of 22 occupation dummies or shares based on the ATUS “trdtocc1” variable, which categorizes the respondent’s main job. Examples include “education, training, and library occupations” and “food preparation and serving related occupations.” As discussed in Section 3.1, in our long-run analysis we collapse ATUS to a cross section in locations.

To investigate the robustness of our long-run results, we also use the Quarterly Census of Employment and Wages (QCEW), which covers all US counties. The QCEW is collected by the BLS as part of the Covered Employment and Wages Program (ES-202), under which state employment security agencies report data to BLS. Covering all workers eligible for unemployment insurance, the data provide “a virtual census” of nonagricultural employees and cover 47 percent of agricultural employees (BLS, 1997). BLS estimates that the QCEW covers 96% of civilian jobs, accounting for 92.5% of the wage and salary component of national income. Self-employed workers and members of the US military are excluded. Earnings measures include bonuses, stock options, and gratuities, but exclude employer contributions to unemployment insurance, worker’s compensation, and pensions (BLS, 1997). We collapse a county-level panel (1990-2013) to a cross section in order to investigate the reduced-form effects of our long-run instrument. Appendix Table 27 presents summary statistics.

[14]ATUS gathers data on all sleep during the course of a single 24 hour period for each individual, so there are potentially other ways to calculate naps, and our results are robust to alternative definitions.

5 Empirical results

5.1 Time use complementarity

We begin by testing theoretical predications about time-use trade-offs. Our model predicts that, if sleep is relatively more work-productivity increasing than home-productivity increasing, an increase in sleep should increase work time and decrease home production time. Moreover, leisure should respond negatively to an increase in sleep. For non-workers, sleep should be complementary to home production.

Table 3 estimates these time complementarities for full time workers and non-workers in both the short-run and long-run ATUS samples. To estimate complementarity, we use a new method by Allen and Rehbeck (2016). With a simplex budget, as in the case of time use, complementarity can be assessed using a ratio estimator similar in spirit to a Wald estimator. In our setting, this means that we can determine whether two time uses are complements or substitutes by regressing each on sunset time, then taking the ratio of the two resulting regression coefficients. A positive ratio indicates that the time uses are complements while a negative ratio indicates they are substitutes. This measure is theoretically motivated but also intuitively appealing: because the time budget must bind, any two large uses of time are likely to be negatively correlated. Therefore, complementarity must be assessed by comparing changes in the two time uses with respect to a time shifter.

To bring this method to the data, we categorize time uses from the ATUS diaries as either sleep, work, leisure, or home production.[15] Estimating the effect of sunset time on these four time use categories, one can see that they provide evidence in favor of the productive sleep model. In the short-run sample, complementarity between work and sleep for workers and complementarity between home production and sleep for non-workers are consistent with productive sleep and are inconsistent with a model where sleep is not productive. Similar evidence is provided by the long-run relationship between sleep and home production in both the worker and non-worker sample, although the evidence is less precise for workers.

In the context of the model, the short-run worker results show that sleep is relatively

[15]The sleep definition is the same as that used throughout the rest of the paper. Work time is all time that the individual reports spending at any job (ATUS two-digit time category 05). Home production time is the sum of time spent in personal care (category 01); housework (category 02); caring for a household or non-household member (categories 03 and 04); in education (category 06); shopping (category 07); or using professional, household, or government services (categories 08, 09, and 10). Leisure time is all remaining time categories (11 through 18) including eating, socializing, sports and recreation, attending religious activities, volunteering, and traveling.

Table 3: Time use responses to sunset

	Sleep	Work	Leisure	Home production
Panel A: Short run				
Full-time workers				
Daily sunset time	-0.38***	-0.11*	0.21***	0.085
	(0.041)	(0.065)	(0.060)	(0.055)
Mean of dep. var.	57.9	31.2	49.5	26.6
Observations	61,161	61,161	61,161	61,161
Non-workers				
Daily sunset time	-0.38***		0.26***	-0.19*
	(0.065)		(0.097)	(0.10)
Mean of dep. var.	62.5		61.5	38.4
Observations	28,126		28,126	28,126
Panel B: Long run				
Full-time workers				
Avg. sunset time	-0.93***	0.66	0.73	-0.28
	(0.28)	(0.62)	(0.57)	(0.47)
Mean of dep. var.	57.9	31.4	49.4	26.5
Observations	529	529	529	529
Non-workers				
Avg. sunset time	-0.11		1.49**	-1.30*
	(0.48)		(0.71)	(0.79)
Mean of dep. var.	62.5		61.5	38.1
Observations	527		527	527

Notes: Data are from ATUS. The table shows results from 14 separate regressions estimating versions of the first stage Equation (5) where the dependent variable is time use in one of four categories, indicated at the top of each column. Controls, number of observations, and standard error clustering are the same as in Tables 4 and 5 for the short-run (Panel A) and long run (Panel B) samples, respectively. The full-time workers sample is the same as that used throughout the paper. The non-workers are prime age individuals who do not report having a full or part-time job. The weights in the long-run sample are based on the population of all workers in each location. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

more work-productivity enhancing than home-productivity enhancing because instrumented sleep and work positively covary while instrumented home production and sleep negatively covary. For the long-run, full-time worker sample, there is not a precise relationship between instrumented sleep and any of the other three time categories. The point estimates indicate that sleep is relatively more home-productivity enhancing for both the worker and non-worker samples.[16]

The other results do not help distinguish between productive and non-productive sleep, but they do provide context for the rest of the results in the paper. In all samples, leisure and sleep are substitutes, and in all but one case, this is the strongest relationship with sleep. Effects on work time are smaller in magnitude than those on sleep time in both the short and long run, but such changes nonetheless could influence our results using usual weekly earnings as a dependent variable. In Sections 5.2 and 5.4 we obtain similar results when we control for a quadratic in usual hours worked or use only the sample of workers who report an hourly wage, indicating that wage changes are quantitatively more important than changes in work time.

5.2 Effect of sunset on sleep and earnings

Table 4 shows results from estimating Equations (5) and (6) on daily ATUS data. The first column shows that the sun setting one hour later within a location reduces nighttime sleep by roughly 20 minutes per week, which is statistically significant at the 1% level. We observe about 5 hours of variation in daily sunset time across our sample, meaning that we identify 1.9 hours per week of intra-annual sleep variation. Practically, this represents a substantial change in time use—roughly equal in magnitude to the 2.1 weekly work hours lost by the average individual during the most recent recession (Aguiar et al., 2013).[17] The time use results in Table 3 show that this short-run effect is uniform across employed and non-employed individuals, and unreported heterogeneity analyses show that worker sleep responds fairly equally to solar cues regardless of gender, age, race, or other demographic characteristics. In Appendix Section C.3.1 we compare sleep distributions in early and late-sunset diaries. Early seasonal sunsets shift the whole distribution rightward, increasing sleep for most quantiles. The second column of Table 4 shows that daily sunset time also affects earnings in a location. A sunset time one hour later reduces earnings by a statistically significant 0.44%.

[16]In Section B.1, we theoretically investigate reverse causality flowing from wages to sleep, showing that if sleep is relatively more work productive than home productive, an exogenous increase in wages can lead to a decrease in sleep. Consistent with the model, Table 3 shows that short-run sleep is relatively more work productive, and an OLS regression of wages on sleep for this sample yields a negative coefficient. In contrast, sleep appears to be relatively more home productive in the long run, and an OLS regression of wages on average sleep in that sample yields a positive coefficient.

[17]Our results suggest that the sunset time influences sleep across the sample. Recession-induced work time changes are concentrated among those who lose their jobs, so individual responses to these two shocks will likely differ even though the averages are the same.

For workers who report an hourly wage, the first stage estimate is nearly identical to the full sample estimate. The reduced form effect of sunset time on wages is less than half the size of the effect on earnings, consistent with the rigidity-attenuation model discussed in Section 3.2.2. In Appendix Section B.2, we derive exact attenuation predictions based on wage and earnings flexibility estimates from Barattieri et al. (2014). Based on the analysis, we expect the non-hourly worker estimate to be 1.4 times the size of the hourly worker estimate. Estimates comparing earnings effects for the non-overlapping samples of works who report an hourly wage and workers who do not, reported in Appendix Table 12, show that the non-hourly worker reduced form estimate is 1.5 times larger. We find additional corroborating evidence for the attenuation model when we examine the reduced form estimates for union members and government employees. Both of these workers have higher wage rigidity than the typical worker, so our measurement error model predicts that the reduced form estimate of sunset on earnings will be attenuated for these workers. Indeed, estimates in Table 13 show that this is the case.

Table 4: Short-run effects of sunset on sleep, earnings and wages from ATUS

	(1) First stage Sleep	(2) Reduced form ln(earnings)	(3) First stage Sleep	(4) Reduced form ln(wage)
Daily sunset	-0.38***	-0.0044***	-0.42***	-0.0020
	(0.041)	(0.0017)	(0.061)	(0.0016)
Individual controls	Yes	Yes	Yes	Yes
Time controls	Yes	Yes	Yes	Yes
Location FEs	Yes	Yes	Yes	Yes
Mean of dep. var.	57.9	6.67	58.5	2.68
Sample	Earnings	Earnings	Hourly wage	Hourly wage
Observations	61,161	61,161	32,040	32,040
Adjusted R^2	0.12	0.30	0.10	0.35

Notes: The table shows results from estimating Equation (5) (first column) and Equation (6) (second column) on ATUS data. The dependent variable is indicated at the top of each column. Earnings refers to "usual weekly earnings." Sleep is measured in hours per week and sunset time in hours. Controls are discussed in Section 3.1 and are location fixed effects; race indicators; age indicators; a gender indicator; indicators for holiday, day of week, and year; and occupation indicators. Standard errors, clustered at the FIPS code (location) level, are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Both estimates include individual race indicators (white, black, Asian, and other), age, age squared, a gender indicator, and detailed occupation code indicators (23 categories). Time controls are an indicator for holidays, separate indicators for each day of the week, and year fixed effects. The location fixed effects are at the most disaggregated FIPS code level available for each observation (county, CBSA, or state). The fixed effects absorb any spatial differences in sunset time, leaving only the seasonal component with which to identify the coefficient of interest.

We cluster standard errors at the FIPS code level, as the exogenous variation is at the group rather than the individual level. As shown in the robustness checks below, clustering at higher levels does not change the inference. Other robustness checks related to concerns about seasonal confounders are in Section 5.3.1.

Table 5 presents estimates of the long-run effects of sunset on sleep (Equation 7), earnings and wages (Equation 8) using a cross-section of continental US locations from ATUS. In the US, time zones create just over an hour of variation in long-run sunset time across locations. Thus average sunset induces about one hour per week of sleep variation—roughly half the variation induced by the short-run instrument. Column 1 shows that average weekly sleep falls by just under one hour in a location where the sun sets one hour later. This is similar to the result obtained by Giuntella et al. (2015) using Chinese data. In Appendix Table 29, we obtain a roughly similar first-stage estimate using US data collected by Jawbone, a manufacturer of wearable health and sleep trackers. We also show that the sleep difference arises because later sunset time delays sleep onset by more than it delays awakening. These Jawbone estimates indicate that our ATUS results do not arise from reporting biases in the time diaries. In Appendix Section C.3.1 we compare sleep distributions in early and late-sunset locations. Early long-run average sunsets move mass out of the lower tail of sleep and into the center of the distribution. The second column of Table 5 shows that for a location where average sunset is one hour later, earnings are more than 4% lower on average. To evaluate the relative importance of wage and hours effects, we also estimate the models for workers who report an hourly wage. Column 3 reports the first stage for wage workers and column 4 reports the reduced form. Precision is reduced, in part because there are fewer individual observations underlying the location-level averages. Point estimates are similar to those from earnings models, and we cannot reject null hypotheses of equal coefficients across wage and earnings models. While the sample of hourly wage workers is different than the full sample, this pattern of results provides some evidence that wage changes are driving most of the estimated effects on earnings.

We report heteroskedasticity-robust standard errors (White, 1980). We do not cluster

Table 5: Long-run effects of sunset on sleep, earnings and wages from ATUS

	(1) First stage Sleep	(2) Reduced form ln(earnings)	(3) First stage Sleep	(4) Reduced form ln(wage)
Avg. sunset time	-0.93***	-0.045***	-0.73*	-0.061***
	(0.28)	(0.017)	(0.40)	(0.020)
Geographic controls	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes
Mean of dep. var.	57.9	6.67	58.5	2.68
Sample	Earnings	Earnings	Hourly wage	Hourly wage
Observations	529	529	529	529
Adjusted R^2	0.125	0.811	0.097	0.565

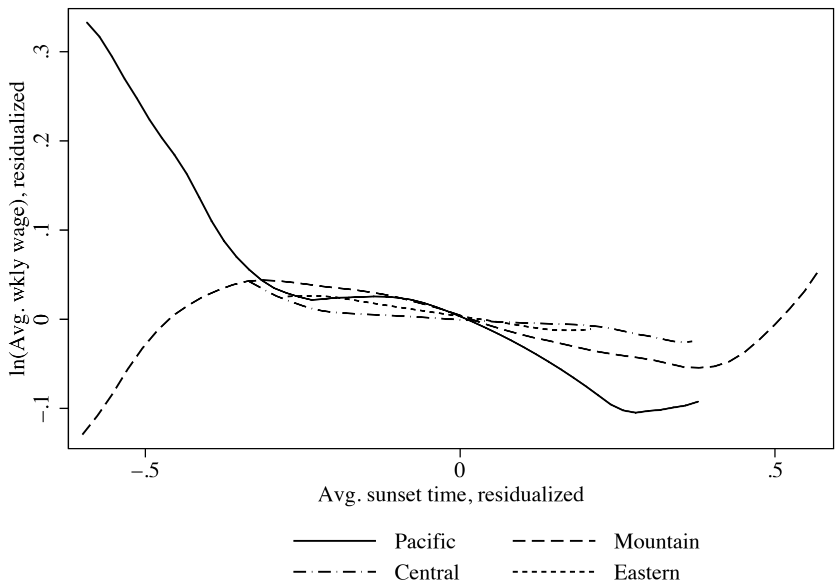
Notes: The table shows results from estimating Equation (7) (columns 1 and 3) and Equation (8) (columns 2 and 4) on ATUS data. The dependent variable is indicated at the top of each column. Earnings refers to "usual weekly earnings" and wage to hourly wage for those workers who reports one. Sleep is measured in hours per week and sunset time in hours. Controls are discussed in Section 3.1 and are: an indicator for coastal county, coastal distance, and their interaction; a ten-piece linear spline in latitude; mean age and mean squared age; percent female; race and occupation shares; and a five-piece linear spline in population density. White heteroskedasticity-robust standard errors are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

because annual average sunset varies exogenously across locations, however we show in Section 5.3.2 that our results are robust to clustering standard errors at the state level. Section 5.3.2 also includes a more focused discussion of robustness to potential spatial confounders. While weighted OLS is plausibly more efficient in this setting, unweighted results are presented and discussed in Appendix Section C.3.

Together Tables 4 and 5 demonstrate that sunset time has practically important effects on time allocation and labor markets. These results are novel and do not require an assumed exclusion restriction. Comparing the two reduced form estimates, one sees the short-run estimate is about 10% the size of the long-run estimate. While these estimates may differ for many reasons, including different complier populations, this is within the 0 to 25% range suggested by our the measurement error analysis, as discussed in Section 3.2.2.

The cross-sectional results pool locations from all four continental US time zones. If our hypothesized causal relationships between sunset, sleep, and wages hold, however, we would expect to find similar results within each time zone. The incomplete geographic coverage of ATUS limits our ability to explore within time zones, so we turn instead to the QCEW, which includes all US counties. The added spatial richness of QCEW also allows for semi-parametric estimates of the reduced form relationship between average sunset time and wages, shown in Figure 4. The figure shows a locally-weighted kernel regression of residual log earnings on residual average sunset time within each time zone.

Figure 4: Long-run effects of sunset on wages in QCEW



Notes: Underlying wage data are a cross-section in locations from QCEW, 1990-2013. The figure shows the relationship between residualized log average wage and residualized sunset time for counties in the Eastern, Central, Mountain, and Pacific time zones. Time zone borders, defined as longitudes at which multiple time zones overlap, are excluded for the reasons discussed in Section 3.2.3. Residualization is with respect to coastal distance, latitude, median age, percent female, percent black, and percent white. Demographic controls are from 2010 Census data.

To arrive at the figure, we first residualize log wage and average sunset time using a control set similar to our preferred ATUS cross-sectional specification. Controls are time zone indicators, interactions of those indicators with linear splines in coastal distance for the Pacific and Eastern time zones, a linear spline in latitude, median age, percent female, percent black, and percent white. We plot a separate kernel regression of the relationship between this residualized wage and residualized average sunset time for each time zone. There are some striking non-linearities at the edges of the plot, where data are sparse. In the central region, however, all four time zones exhibit similar negative, roughly linear relationships.[18] In

[18]This suggests that differential attenuation from the larger counties in the Pacific and Mountain zones is not

Section 5.3.2 we present the pooled parametric analog of these QCEW analyses, together with robustness checks.

5.3 Robustness

The following sections explore the robustness of our reduced-form specifications. In general we cannot reject a null hypothesis of equality with our preferred estimates. The differences in coefficients are sometimes large in proportional terms, but uniformly small in absolute terms. In general we can reject a null hypothesis that the coefficient on sunset time is zero, and we comment on the exceptional cases.

5.3.1 Short-run robustness

The primary identification concern with the short-run estimates is that another seasonal variable might be omitted from the sunset time-earnings regressions. We can address this concern in four ways. First, we introduce another seasonal variable—daily temperature—that is plausibly correlated with earnings and which covaries with sunset time. We merge temperature data from the NCEP/NCAR reanalysis produced by Kalnay et al. (1996) at the day-location level. The data is available on a two-by-two degree latitude-longitude grid, and we use the daily average temperature for the grid point corresponding to our geocoded centroid for estimation.

The reduced form coefficients including weather controls are the first two entries in Table 6. All of the baseline controls are also included in these regressions. Including temperature changes the coefficient estimate by less than one-third of a standard error. While the temperature control reduces precision, the estimate remains statistically significant at the five percent level. Temperature is highly correlated with sunset time, and it has a small, nominally significant effect on earnings. Adding rainfall reduces significance slightly but does not change the point estimate, since rain is not strongly correlated with either sunset time or earnings. We have also investigated alternative weather controls including diurnal temperatures and wind speed. Daily average temperature and minimum or maximum temperatures are highly correlated, so the effect on inference is similar. Wind speed does not change the result.

Next, we can include seasonal fixed effects or other calendar controls. The reduced form estimate including quarter fixed effects is shown in the third entry in the table. The estimate is larger, although precision again declines. With month fixed effects, the first stage becomes zero to the second decimal place, and the reduced form also becomes zero to the fourth decimal

practically important. The relationship between within-time zone longitude and wage is shown in Appendix Figure 16.

place. Within a month, there is not enough variation to identify either of these regression coefficients. The fourth entry in Table 6 removes the holiday season (Thanksgiving through January 15th), again finding little change. In appendix Table 16, we show that dropping the entire first and fourth quarters does not substantially change inference—the results show that the reduced form effect is strongest in quarters 2 and 3.[19]

We can also parsimoniously assess seasonal confounding by including control variables that have the same seasonal pattern as sunset time but different phase or frequency. Figure 2 shows that sunset time follows a sinosoidal pattern within a year, peaking around June 21st and reaching a trough around December 21st. Appendix Section C.2.1 contains results from including seasonal control variables with this same pattern but with peaks and troughs on other days of the year. The results show that the sign of the baseline estimate is robust to including any phase-shifted seasonal control variable of this type.

We find no evidence of confounding seasonal changes in ATUS sample composition. Appendix Figure 8a shows that occupation shares in our sample are constant over the months of the year, and Appendix Figure 8b shows that the share of ATUS respondents reporting a positive wage (i.e. the fraction engaged in market work) is likewise constant over the year. Together, these figures reassure us that such selection bias is not driving our results.

Table 6: Robustness of ATUS short-run reduced form estimate

	ln(earnings)			ln(earnings)	
Temperature control			*No controls*		
Daily sunset	-0.0049**	(0.0025)	Daily sunset	-0.0058***	(0.0022)
Temperature and rain			*Work hours quadratic*		
Daily sunset	-0.0048*	(0.0025)	Daily sunset	-0.0042**	(0.0017)
Quarter FEs			*State clustering*		
Daily sunset	-0.0086**	(0.0034)	Daily sunset	-0.0044**	(0.0018)
No holiday season			*Education controls*		
Daily sunset	-0.0049**	(0.0020)	Daily sunset time	-0.0033**	(0.0016)
Observations	52,856				

Notes: The table shows results from estimating Equation (6) estimated on ATUS data. Dependent variable is indicated at the top of each column. Unless otherwise noted, controls, number of observations, and standard errors are the same as in Table 4. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

[19]In unreported results, the estimate is robust to dropping any two quarters from the sample. Dropping three quarters removes too much data for precise inference, but the point estimates are similar.

Additional robustness checks are reported in Table 6 and Appendix Section C.2. The first entry in column 2 re-estimates the reduced-form relationship with only time control variables and fixed effects. The estimate is similar in magnitude to the baseline estimate, but with lower precision. The next specification controls for a quadratic in usual hours worked and returns estimates very similar to our primary short-run results.[20] Per Borjas (1980), a control of this type is preferable to dividing on the left-hand side of the equation because the latter course may introduce division bias (Kronmal, 1993). We prefer not to include such a control in our main specification because omitting a work hours control allows workers to take additional sleep time out of either work time or other (non-work, non-sleep) time. Controlling for work time forces all changes in sleep to come out of other time. That might bias our estimates if, as theory predicts, sunset time responses co-vary with unobserved earnings determinants. In the penultimate entry we cluster standard errors at the state level for all observations and find that inference does not change. The final robustness check adds indicator variables for highest educational attainment, since Table 2 revealed potential imbalance on these variables. Controlling for educational attainment categories reduces the estimated reduced form effect slightly. A specification including many of these controls simultaneously is in Table 17.

5.3.2 Long-run robustness

We now test the sensitivity of our long-run reduced form estimate. Again, the first stage estimates are stable across specifications and are omitted. Table 7 shows estimates of the reduced form equation (8) with variations in controls, sample, and clustering. First, we report estimates using only geographic controls, without any other covariates. Consistent with our conditional exogeneity assumption, the coefficient is similar in magnitude to our preferred estimate. It is however much less precise (p=.13 and the standard error is almost three times larger), indicating the demographic controls meaningfully reduce residual variance. Next, we control for a quadratic in usual hours worked to test whether our primary estimate is driven by wage changes. This estimate is slightly larger than the estimate from our preferred specification, which may be because in the long run a later sunset time may induce a small increase in work hours (the point estimate is positive but imprecise; see Table 3). If workers' production functions are concave, this increase in hours could also tend to decrease our long-run estimate by lowering marginal productivity. Thus our primary long-run result may be interpreted as a

[20] By holding hours fixed, we isolate changes in usual weekly earnings from wages, rather than hours.

lower bound on the wage effect. A linear control for longitude does not change the estimate substantially. Clustering standard errors at the state level for all observations does not change the results of our hypothesis tests.

Table 7: Robustness of ATUS long-run estimates

	ln(earnings)			ln(earnings)	
Only geographic controls			*Time zone indicators*		
Avg. sunset time	-0.075	(0.049)	Avg. sunset time	-0.043*	(0.023)
Work hours quadratic			*No Eastern time zone*		
Avg. sunset time	-0.052***	(0.017)	Avg. sunset time	-0.049*	(0.029)
Longitude control			Observations	244	
Avg. sunset time	-0.042**	(0.018)	*No high-wage cities*		
State clustering			Avg. sunset time	-0.033*	(0.019)
Avg. sunset time	-0.045***	(0.015)	Observations	474	
No time zone border counties			*Albouy QOL control*		
Avg. sunset time	-0.037	(0.024)	Avg. sunset time	-0.032*	(0.019)
Observations	450		*Educational attainment shares*		
100 most populous locations			Avg. sunset time	-0.035**	(0.015)
Avg. sunset time	-0.045	(0.066)			

Notes: The table shows results from estimating Equation (8) estimated on ATUS data. Dependent variable is the log of average earnings. Unless otherwise noted, controls, number of observations, and standard errors are the same as in Table 5. Results reported under "No high-wage cities" exclude workers in San Francisco, Los Angeles, Chicago, Boston, and New York. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

The next robustness check mimics the specification estimated using QCEW data and reported in Figure 4. We drop counties at border longitudes where time zones overlap, as such counties might have selected into a time zone based on economic considerations. The standard error is 41 percent larger, so the test is inconclusive. The point estimate is not statistically distinguishable from zero, or from our preferred estimate. Restricting the sample to the 100 most populous locations similarly decreases precision. Adding time zone indicators yields an estimate that is statistically significant at the ten percent level and similar to our main result.

Next, we directly test for spatial confounding by excluding the wealthier, denser Eastern time zone, then excluding selected high-wage cities. Both of these models return estimates (statistically significant at the ten percent level) similar to our baseline model, reinforcing the result from the QCEW sample that this effect is not restricted to urban areas or a single time zone. To test whether long-run average sunset time co-varies with local amenities, we include a control for the state-level quality-of-life index from Albouy (2008). Again the change in the coefficient is less than one standard error and it remains significant at the ten percent level. Finally we control for shares in the educational attainment bins from Table 2. The estimate is slightly smaller in magnitude, but remains significant at the five percent level.

We next use QCEW data to test the external validity of the long-run ATUS reduced form estimate. Table 8 presents the linear counterpart to the semi-parametric analysis in Figure 4, pooling across time zones. The base specification seeks to match our long-run ATUS specification as closely as possible, though directly analogous variables are not available in all cases. The point estimate resulting from these controls, -.056, is similar to the long-run ATUS estimate, -.045. The pattern of robustness checks is similar to those in Table 7 above; we will comment only on the differences. When standard errors are clustered at the state level in QCEW data, the estimate is no longer statistically significant. Leaving out the Eastern time zone decreases magnitude to -.022 and again the estimate is not statistically significant. In general the QCEW estimates are similar to those from ATUS, however, and in none of the robustness checks can we reject a null hypothesis of equality with the base estimate.

Table 8: Robustness of QCEW long-run estimates

	ln(earnings)	
Base specification		
Avg. sunset time	-0.056**	(0.025)
Only geographic controls		
Avg. sunset time	-0.097***	(0.027)
Longitude control		
Avg. sunset time	-0.11***	(0.025)
State clustering		
Avg. sunset time	-0.056	(0.048)
100 most populous locations		
Avg. sunset time	-0.088	(0.10)

	ln(earnings)	
No Eastern time zone		
Avg. sunset time	-0.022	(0.027)
No high-wage cities		
Avg. sunset time	-0.046*	(0.026)
Albouy QOL control		
Avg. sunset time	-0.060**	(0.026)
Educational attainment shares		
Avg. sunset time	-0.077***	(0.025)

Notes: The table shows the linear regression analogue of Figure 4. Heteroskedasticity-robust standard errors are reported in parentheses. Data are from the BLS Quarterly Census of Employment and Wages 1990-2013. Sunset time is the long-run county average. Results reported under "No high-wage cities" exclude workers in San Francisco, Los Angeles, Chicago, Boston, and New York. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Both ATUS and QCEW estimates, then, show long-run location-level effects of sunset time on wages. A sorting model like Roback (1982) predicts that under perfect worker and firm mobility, the gains from a productive location-specific amenity accrue to owners of land, the fixed factor. In this setting, the prediction is that locations with earlier average sunset times will have higher rents and house prices. Using county-level Census data from 2010, Table 9 provides evidence consistent with that prediction. We regress log median county home value on average sunset time and a set of controls similar to our long-run ATUS specification. A county experiencing sunset one hour earlier will have, on average, a median home value approximately 6% higher. In levels, the estimated effect on median home value is approximately \$7,900 to \$8,800. Based on the discussion following Table 10, a worker's annual income gain from moving to a location where sunset is an hour earlier is approximately \$1,570. Thus the capitalization of sunset time into home prices reflects the present discounted value of wage gains from approximately 8 to 9 years of work (assuming a five percent discount rate). This result is roughly consistent with the prediction of the Roback model. Given that an optimizing worker responds to real, not nominal, income, this finding suggests workers do not have a financial incentive to move to an early-sunset location. In Appendix Table 28 we show this hedonic estimate is robust to additional controls.

Table 9: Effects on log median home value

	(1) Log value	(2) Log value
Sunset time	-0.0640***	-0.0574***
	(0.0232)	(0.0191)
Geographic controls	Yes	Yes
Demographic controls	No	Yes
Observations	2824	2824
Adjusted R^2	0.399	0.617

Notes: White's heteroskedasticity-robust standard errors are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1. Data are 2010 5-year ACS estimates. Sunset time is the average for a given county. Geographic controls include coastal distance and a ten-piece linear spline in latitude. Demographic controls include percent female, percent in four race categories, occupation shares, and a five-piece linear spline in population density.

5.4 IV estimates of the effect of sleep on earnings and wages

In the preceding sections, we have discussed potential threats to instrument validity (Section 3.2) and conducted empirical tests for violations of the exclusion restriction (Section 5.3). Table 10 presents our instrumental variables estimates of the effect of sleep on earnings. A one-hour increase in short-run mean weekly sleep in a location increases earnings by 1.1%. A one-hour increase in long-run mean weekly sleep in a location increases earnings by 4.9%. Together these results demonstrate that exogenous location-level sleep changes have important earnings effects.[21]

Table 10: IV estimates of the effect of sleep on earnings

	(1) Short run ln(earnings)	(2) Long run ln(earnings)
Sleep	0.011**	0.049**
	(0.0048)	(0.024)
Short-run controls	Yes	No
Long-run controls	No	Yes
Observations	61,161	529
Adjusted R^2	0.22	0.70
F-stat on IV	95.16	10.66
Elasticity	0.72	2.58

Notes: The table shows instrumental variables estimates using ATUS data and based on the first stage and reduced form Equations (5) and (6) (column 1) and Equations (7) and (8) (column 2). The dependent variable is the natural log of weekly earnings. Controls for column 1 are the same as in Table 4 and for column 2 are the same as in Table 5. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

In the short-run earnings model, the first-stage F statistics of 95.2 exceeds the relevant Stock-Yogo critical value of 16.38, so we reject the null hypothesis of weak instruments, where "weak" is defined as true size greater than 10% for a nominal 5% test (Stock and Yogo, 2005). In the long-run earnings model, the first-stage F statistic of 10.66 falls short of the Stock-Yogo critical value for true size of 10%, but exceeds the critical value of 8.96 for true size of 15%. The result of our long-run t-test should thus be viewed with more skepticism, but we can rule out gross failures of size control.

While our earnings estimates largely reflect wage changes rather than changes in hours

[21] Appendix Section C.4 investigates possible nonlinearity in the sleep-earnings relationship and finds no evidence for it.

worked, Table 11 directly examines effects on hourly wage for workers who report one. A one-hour increase in short-run mean weekly sleep in a location increases wages by 0.5%. A one-hour increase in long-run mean weekly sleep in a location increases wages by 8.3%. Precision is reduced relative to the earnings models. The short-run estimate is not statistically significant and the long-run estimate is statistically significant only at the 10% level. In the long-run wage model, the first-stage F statistic is quite small at 3.27 and we strongly caution against placing too much weight on this particular point estimate. Consistent with the analysis in Section B.2 and the earnings flexibility estimates of Barattieri et al. (2014), the short-run estimates are smaller relative to the long-run estimates when using wages rather than earnings.

Table 11: IV estimates of the effect of sleep on wages

	(1) Short run ln(wage)	(2) Long run ln(wage)
Sleep	0.0047	0.083*
	(0.0039)	(0.050)
Short-run controls	Yes	No
Long-run controls	No	Yes
Observations	32,040	529
Adjusted R^2	0.32	0.14
F-stat on IV	51.83	3.27
Elasticity	0.29	2.54

Notes: The table shows instrumental variables estimates using ATUS data and based on the first stage and reduced form Equations (5) and (6) (column 1) and Equations (7) and (8) (column 2). The dependent variable is the natural log of wages. Controls for column 1 are the same as in Table 4 and for column 2 are the same as in Table 5. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Our estimates reflect location-level local average treatment effects, so several nuances of interpretation warrant discussion. First, all workers in a location experience the same sunset time. If sunset-induced sleep differences generate productivity spillovers across workers, our estimated β will capture not the effect of increasing individual sleep, but rather the effect of increasing mean sleep in a location. The empirical literature on peer effects provides evidence that such spillovers may be large. For example, Chetty et al. (2011) estimates an elasticity of age-27 earnings with respect to kindergarten peer quality of roughly one. Similarly, Carrell et al. (2009) estimate an elasticity of academic achievement with respect to peer quality of roughly 0.9. Mas and Moretti (2009) find evidence of large positive peer effects in supermarket cashiers. Moretti (2004) finds that human capital spillovers can operate at the municipal level, estimating that a one percent increase in the share of college-educated workers in a city increases output by roughly half of one percent. This is similar in spatial scale to our analysis. It suggests that if earlier sunset time increases sleep and productivity for all workers in a location, spillover effects may be broadly felt. Because spillovers and peer effects are externalities, it would not necessarily be individually rational for a worker to sleep more. This body of evidence implies that an individual worker would not see a 4.9% wage increase from an additional hour of sleep, but instead something potentially much smaller. The second interpretative nuance is that differences in long-run sunset time across locations are nearly permanent. For most locations, long-run sunset time has changed only when the US has revised its time zones, which it has not done since the first world war. If labor and capital are on average complements, it is possible that sleep-driven labor productivity differences have influenced the long-run growth of the capital stock. Thus the effects we estimate may have emerged over many years, and short run effects could be different. Third, a positive marginal effect does not imply optimization failure. If the marginal willingness to pay for leisure exceeds the marginal wage increase from additional sleep, an optimizing worker will not increase sleep.

Our analysis demonstrates that workers experiencing an earlier sunset get more sleep. As discussed in Section 5.1, in the short run the additional sleep largely comes at the expense of leisure, while in the long run it comes at the expense of both work and leisure. Insofar as these changes in other time uses impact worker productivity, our instrumental variables estimates of the effect of sleep on wages will also contain those effects. In other words, we estimate the effect of a quasi-random change in sleep, allowing other time uses to adjust endogenously. While this might seem undesirable at first glance, it is unavoidable. An agent's time constraint always binds with perfect equality. Even in a laboratory setting, it is not possible to change the time use of interest without also changing at least one other time use.

Expressed as an elasticity, our short-run earnings estimate is 0.84 and our long-run earnings estimate is 2.6. Medical researchers have typically found elasticities of task performance with respect to sleep duration of approximately four (Table 1). If wages are equal to a worker's marginal physical product multiplied by output price, we expect such performance effects to produce equally large wage effects. Our smaller estimated elasticities may reflect differences between laboratory tasks and actual work tasks or the broader scope for adaptation (for instance the use of stimulants like coffee) outside the lab. The difference between short- and long-run elasticities is consistent with the attenuation bounds calculated in appendix Section B.2. Unaided intuition might suggest smaller effects of sleep on performance, but intuition provides a poor sense of this relationship: Van Dongen et al. (2003) showed that subjects' self-reported fatigue quickly stabilized after a few days of sleep reduction, even as their performance continued to decline.

Taking average values for earnings and assuming 50 work weeks per year, one can calculate the annual income effects implied by our long-run estimates. If mean weekly sleep in a location increased by one hour and work time remained unchanged, mean annual income would rise by about \$2,350. In reality, extra sleep comes out of both work and non-work time. If workers took roughly 70% of the extra sleep hour out of work time then a one-hour increase in weekly mean sleep in a location would increase mean annual income by about \$1,570. If extra sleep came solely at the expense of work time, the income increase would be \$1,250. As we showed analytically in Section 2.2, an optimizing agent might in fact choose to increase work time in response to an increase in sleep.

6 Conclusion

Although time use is entangled in a causal web with labor market outcomes, economists have paid little attention to these relationships. In particular, the profession has scarcely examined sleep. In this paper, we formalize the links between productivity-improving sleep, wages, and other time uses in an optimizing model. We generate theoretical predictions on complementarity between sleep and other time uses for both workers and non-workers. Empirical tests yield results largely consistent with those predictions. Our earnings and wage regressions demonstrate that sleep has a powerful impact on labor market outcomes and should be considered an integral part of a worker's utility maximization problem. Using individual time-use diaries matched with labor market variables from ATUS, we show that increasing short-run weekly average sleep in a location by one hour increases worker earnings by roughly 1%. Increasing long-run weekly average sleep in a location by one hour increases earnings by roughly 5%. Our use of instrumental variables techniques addresses the reverse-causality and omitted variable problems that would bias naïve estimates. We buttress this finding with a battery of short and long-run robustness checks, and a hedonic model of home prices showing that long-run wage increases are partially capitalized into housing.

These results suggest that sleep is a crucial determinant of wages, rivaling ability and education in importance. Figures of the magnitude shown here naturally lead one to ask why workers do not work less and sleep more. One possible explanation lies in the spillovers and general-equilibrium effects our estimate incorporates. It is also conceivable that observed sleep reflects optimization failure by workers, as hypothesized by Mullainathan (2014). Such failure could occur even under classical assumptions. For example, the inaccurate self-perceptions of fatigue found by Van Dongen et al. (2003) could lead to sleep below the utility-optimizing level even if workers are behaving optimally, conditional on their information set. Such sub-optimal sleep could, in turn, contribute to the type of time-use poverty trap analyzed by Banerjee and Mullainathan (2008).[22] On the other hand, sub-optimal sleep could arise from behavioral considerations like time inconsistency or constraints on cognitive resources (see for example Mani et al. (2013)). While optimization failure and its possible mechanisms are beyond this scope of the current paper, we are exploring them in ongoing experimental work.

Further attention should be paid to industries characterized by chronic sleep shortages. In addition to wages, optimal sleep plausibly depends on other factors like leisure complementarities, direct sleep utility, and health optimization. Each of these trade-offs suggests an interesting research question. More broadly, our results demonstrate that non-labor time uses can have first-order effects on labor outcomes and these effects warrant further investigation.

References

Aguiar, M. and E. Hurst (2007). Measuring trends in leisure: the allocation of time over five decades. *The Quarterly Journal of Economics* (August), 969–1006.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013). Time use during the great recession. *The American Economic Review 103*(5), 1664–1696.

Albouy, D. (2008). Are big cities bad places to live? estimating quality of life across metropolitan areas. *NBER Working Paper No. 14472*.

Allen, R. and J. Rehbeck (2016). Complementarity in perturbed models. pp. 1–41.

Allen, W. F. (1883). Open letters: Standard railway time. *The Century*, 796.

Anonymous (1883, December). The new standard time and its advantages. *The Manufacturer and Builder 15*(12).

Banerjee, A. V. and S. Mullainathan (2008). Limited attention and income distribution. *The American Economic Review*, 489–493.

[22] While the authors interpret their work as a model of inattention, they note, ”In fact, our model is formally identical to a rational time allocation model, if we think of comfort goods as time saving devices.”

Banks, S. and D. F. Dinges (2007, August). Behavioral and physiological consequences of sleep restriction. *Journal of Clinical Sleep Medicine 3*(5), 519–28.

Barattieri, A., S. Basu, and P. Gottschalk (2014). Some evidence on the importance of sticky wages. *American Economic Journal: Macroeconomics 6*(1), 70–101.

Barnes, C. M. and D. T. Wagner (2009). Changing to daylight saving time cuts into sleep and increases workplace injuries. *Journal of Applied Psychology 94*(5), 1305.

Bartky, I. (1989). The Adoption of Standard Time. *Technology and Culture 30*(1), 25–56.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy 70*(5), 9–49.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal 75*(299), 493–517.

Belenky, G., N. J. Wesensten, D. R. Thorne, M. L. Thomas, H. C. Sing, D. P. Redmond, M. B. Russo, and T. J. Balkin (2003, March). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research 12*(1), 1–12.

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

Black, S. E. and L. M. Lynch (2001). How to compete: The impact of workplace practices and information technology on productivity. *The Review of Economics and Statistics 83*(August), 434–445.

Bloom, N., R. Sadun, and J. Van Reenen (2012). Americans do it better: US multinationals and the productivity miracle. *The American Economic Review 102*(1), 167–201.

BLS (1997). Chapter 5. Employment and Wages Covered by Unemployment Insurance. In *BLS Handbook of Methods*, Number April.

BLS (2015). American Time Use Survey User's Guide.

Bonke, J. (2012). Do morning-type people earn more than evening-type people? How chronotypes influence income. *Annales d'Economie et de Statistique* (105), 55–76.

Borjas, G. J. (1980). The relationship between wages and weekly hours of work: The role of division bias. *Journal of Human Resources*, 409–423.

Brochu, P., C. D. Armstrong, and L.-P. Morin (2012, September). The 'trendiness' of sleep: An empirical investigation into the cyclical nature of sleep time. *Empirical Economics 43*(2), 891–913.

Cappuccio, F., L. D'Elia, P. Strazzullo, and M. Miller (2010). Sleep Duration and All-Cause

Mortality: A Systematic Review and Meta-Analysis of Prospective Studies. *Sleep*.

Carrell, S., R. Fullerton, and J. West (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics 27*(3), 439–464.

CFR (2010). Relocation of Standard Time Zone Boundary in the State of North Dakota: Mercer County.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics 126*(4), 1593–1660.

Cohen, D. A., W. Wang, J. K. Wyatt, R. E. Kronauer, D.-J. Dijk, C. A. Czeisler, and E. B. Klerman (2010, January). Uncovering residual effects of chronic sleep loss on human performance. *Science Translational Medicine 2*(14), 14ra3.

Dinges, D. F., F. Pack, K. Williams, K. A. Gillen, J. W. Powell, G. E. Ott, C. Aptowicz, and A. L. Pack (1997, April). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep 20*(4), 267–77.

Frazis, H. and J. Stewart (2012, January). How to Think about Time-Use Data: What Inferences Can We Make about Long- and Short-Run Time Use from Time Diaries? *Annals of Economics and Statistics* (105/106), 231–245.

Giuntella, O., W. Han, and F. Mazzonna (2015). Circadian Rhythms, Sleep and Cognitive Skills. Evidence From an Unsleeping Giant.

Graff Zivin, J. and M. Neidell (2012, December). The Impact of Pollution on Worker Productivity. *American Economic Review 102*(7), 3652–3673.

Gronau, R. (1977). Leisure, Home Production, and Work-the Theory of the Allocation of Time Revisited. *The Journal of Political Economy 85*(6), 1099–1123.

Gutknecht, D. (2013). Testing monotonicity under endogeneity. *Oxford Department of Economics Discussion Paper*.

Haines, M. R. and Inter-university Consortium for Political and Social Research (2010). Historical, Demographic, Economic, and Social Data: The United States, 1790-2002.

Hall, R. E. (1975). The Rigidity of Wages and the Persistence of Unemployment. *Brookings Papers on Economic Activity 2*, 301–349.

Hamermesh, D. S., H. Frazis, and J. Stewart (2005). The American Time-use Survey. *Journal of Economic Perspectives 19*(1), 221–232.

Hamermesh, D. S., C. K. Myers, and M. L. Pocock (2008). Cues for Timing and Coordination: Latitude, Letterman, and Longitude. *Journal of Labor Economics 26*(2), 223–246.

Hubert, M., M. Dumont, and J. Paquet (1998). Seasonal and diurnal patterns of human illumination under natural conditions. *Chronobiology International 15*(1), 59–70.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, et al. (1996). The ncep/ncar 40-year reanalysis project. *Bulletin of the American Meteorological Society 77*(3), 437–471.

Kim, K. and A. Petrin (2011). A new control function approach for non-parametric regressions with endogenous variables. *NBER Working Paper No. 16679*.

Kronmal, R. A. (1993). Spurious Correlation and the Fallacy of the Ratio Standard Revisited. *Journal of the Royal Statistical Society 156*(3), 379–392.

Krueger, P. M. and E. M. Friedman (2009, May). Sleep duration in the United States: a cross-sectional population-based study. *American Journal of Epidemiology 169*(9), 1052–63.

Lambert, G. W., C. Reid, D. M. Kaye, G. L. Jennings, and M. D. Esler (2002). Effect of sunlight and season on serotonin turnover in the brain. *The Lancet 360*, 1840–1842.

Landrigan, C. P., J. M. Rothschild, J. W. Cronin, R. Kaushal, E. Burdick, J. T. Katz, C. M. Lilly, P. H. Stone, S. W. Lockley, D. W. Bates, and C. A. Czeisler (2004, October). Effect of reducing interns' work hours on serious medical errors in intensive care units. *The New England Journal of Medicine 351*(18), 1838–48.

Leibenstein, H. (1957). *Economic Backwardness and Economic Growth*. Wiley.

Leng, Y., F. P. Cappuccio, N. W. J. Wainwright, P. G. Surtees, R. Luben, C. Brayne, and K.-T. Khaw (2015, February). Sleep duration and risk of fatal and nonfatal stroke A prospective study and meta-analysis. *Neurology*.

Library of Congress (2010, November). Time! Accessed Feb. 4, 2014.

Lockley, S. W., J. W. Cornin, E. E. Evans, B. E. Cade, C. J. Lee, C. P. Landrigan, J. M. Rothschild, J. T. Katz, C. M. Lilly, P. H. Stone, D. Aeschbach, and C. A. Czeisler (2004). Effects of Reducing Interns' Weekly Work Hours on Sleep and Attentional Failures. *The New England Journal of Medicine*, 1829–1837.

Mani, A., S. Mullainathan, E. Shafir, and J. Zhao (2013). Poverty impedes cognitive function. *science 341*(6149), 976–980.

Mas, A. and E. Moretti (2009). Peers at Work. *American Economic Review 99*(1), 112–145.

Mckenna, B. S., D. L. Dickinson, H. J. Orff, and S. Drummond (2007). The effects of one night of sleep deprivation on known-risk and ambiguous-risk decisions. *Journal of Sleep Research 16*(3), 245–252.

Meeus, J. H. (1991). *Astronomical Algorithms*. Willmann-Bell, Incorporated.

Moretti, E. (2004). Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions. *The American Economic Review 94*(3), 656–690.

Mullainathan, S. (2014, February). Get some sleep, and wake up the G.D.P. *The New York Times*.

Murase, S., M. Kitabatake, T. Yamauchi, and A. Mathe (1995). Seasonal mood variation among japanese residents of stockholm. *Acta Psychiatrica Scandinavica 92*(1), 51–55.

Mushkin, S. J. (1962). Health as an investment. *The Journal of Political Economy*, 129–157.

Newey, W., J. Powell, and F. Vella (1999). Nonparametric Estimation of Triangular Simultaneous Equations Models. *Econometrica 67*(3).

Nolan, T. (2014, October). Dance to the (circadian) rhythm.

Psacharopoulos, G. and H. A. Patrinos (2004, August). Returns to investment in education: A further update. *Education Economics 12*(2), 111–134.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy 90*(6), 1257–1278.

Roenneberg, T., C. J. Kumar, and M. Merrow (2007). The human circadian clock entrains to sun time. *Current Biology 17*(2), R44–R45.

Schober, M. F. and F. G. Conrad (1997). Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly 61*, 576–602.

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

Solon, G., S. J. Haider, and J. M. Wooldridge (2015). What are we weighting for? *Journal of Human resources 50*(2), 301–316.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear iv regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80.

Szalontai, G. (2006). The demand for sleep: A South African study. *Economic Modelling 23*(5), 854–874.

Taheri, S., L. Lin, D. Austin, T. Young, and E. Mignot (2004). Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. *PLoS medicine 1*(3), e62.

Taras, H. and W. Potts-Datema (2005). Sleep and student performance at school. *Journal of School Health 75*(7), 248–254.

Ulmer, C., D. M. Wolman, M. M. Johns, et al. (2009). *Resident duty hours: Enhancing sleep, supervision, and safety*. National Academies Press.

USNO (2014). U.S. Time Zones. Accessed Feb. 4, 2014.

Valpando, A. (2013). DOT Procedure for Moving an Area from One Time Zone to Another.

Van Dongen, H. P. A. and D. F. Dinges (2005, April). Sleep, circadian rhythms, and psychomotor vigilance. *Clinics in Sports Medicine 24*(2), 237–49, vii–viii.

Van Dongen, H. P. A., G. Maislin, J. M. Mullington, and D. F. Dinges (2003, March). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep 26*(2), 117–26.

Vgontzas, A. N., E. Zoumakis, E. O. Bixler, H.-M. Lin, H. Follett, A. Kales, and G. P. Chrousos (2004). Adverse effects of modest sleep restriction on sleepiness, performance, and inflammatory cytokines. *The Journal of Clinical Endocrinology and Metabolism 89*(5), 2119–26.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817–838.

Winship, C. and L. Radbill (1994). Sampling weights and regression analysis. *Sociological Methods & Research 23*(2), 230–257.

Appendix for online publication only

A Solar mechanics

Here, we provide a brief summary of how sunset time is calculated and a glossary of terms. For a detailed glossary, see NOAA's ESRL website. We calculate sunset, sunrise, solar declination, and sunlight duration each day using the algorithm of Meeus (1991) as implemented by NOAA's Earth Systems Research Laboratory (ESRL). The calculator takes as inputs the date, time zone offset, latitude, and longitude. It outputs the solar declination, right ascension, sunrise time, sunset time, and sunlight duration. The Stata code that we used for calculation is available on Github.

Solar declination is the angle of a line segment from the sun to the earth relative to a plane projected from the equator of the Earth. The solar declination is a function only of the day of year and time zone offset (to compute fractional days for high-resolution local time sunset), and changes in solar declination correspond to the average seasonal movement of the sun. The highest solar declination, 23.44° occurs on the summer solstice, and the lowest solar declination, -23.44°, occurs on the winter solstice. On the equinox, solar declination is 0°. A rough calculation of solar declination, which we use as the basis of our seasonal control variable in Section 5.3.1, can be calculated as

$$-23.44 \cos\left(\frac{360}{365}(d+10)\right)$$

where d is the day of the year.

Sunset and sunrise time are both calculated assuming 0.833° of atmospheric refraction, or the bending of the path of light as it passes through the Earth's atmosphere. A more precise refraction correction would need to incorporate information on air pressure and humidity. Also, we calculate sunset assuming an observer with a zero-elevation change view of the horizon (not to be confused with the assumption of being at 0 elevation). These two variables take latitude as an additional input, reflecting the location-specific amplitude of sunlight changes.

Sunlight duration is simply calculated as the difference between sunrise and sunset time for a location on a given day. We do not use right ascension in this study.

B Additional theory results and calculation of timing mismatch attenuation

B.1 Theoretical effect of wages on sleep

Applying Cramer's rule to the first order conditions in Section 2.2, the numerator of the derivative of sleep with respect to α is

$$B\left(f_1'w' + f_{11}''(W + w'N) - f_{12}''W\right) - A\left(f_{11}'' - f_{12}'' + w'\right) - f_{11}''w'$$

where we define

$$A = \frac{Z_1'Z_{22}'' - Z_2'Z_{12}''}{\left(Z_1'\right)^2}$$

and

$$B = \frac{Z_1'Z_{12}'' - Z_2'Z_{11}''}{\left(Z_1'\right)^2}$$

The denominator is the determinant of the Hessian. To demonstrate the possibility of a negative effect of α on sleep, consider the case where $Z_{12}'' > 0$ so that $A < 0$ and $B > 0$ and where $f_{12}'' > 0$. For simplicity, also assume that $f_{11}'' = 0$. The numerator will be positive (and therefore the derivative of sleep with respect to α will be negative) if $f_{12}'' < w'$ and $f_{12}''W < f_1'w'$. The first expression is the familiar condition that sleep is relatively more work productive than it is an enhancement of home production time. The second condition will also be met if sleep is highly work productive.

B.2 Timing mismatch and short-run estimate bias appendix

Our preferred productivity measure is the answer to a question about "usual weekly earnings" rather than wages on the day of the interview. Thus, even in the case where we correctly isolate exogenous changes in short-run sleep using daily sunset time, there is an additional identification issue inherent in studying earnings (or wages) rather than productivity: timing mismatch between observations of sleep and wages combined with a potentially low-frequency relationship between wages and productivity. These issues mean that our short-run estimates of both the reduced form Equation 6 and the resulting two-stage least squares estimator will be biased.

This bias is present in all of our short-run estimates, but can be relatively benign. Consider, for instance, a piece-rate worker paid each day. Our observation of this worker's earnings could be based on yesterday's earnings and therefore sleep the previous night. We observe tonight's sleep, however, so the timing of our sleep observation is off by one night. Daily sunset time is highly autocorrelated, so the error in our estimate will be slight; we are using almost the correct observation. In general, however, earnings change more slowly. In Section B.2.1 below we show that using the known seasonal pattern in sunset time and assuming earnings are a linear function of average productivity, the estimated seasonal coefficient has an asymptotic bias that depends only on the frequency of earnings changes and a trigonometric function derived from solar mechanics.

Barattieri et al. (2014, Figures 12 and 13) provide estimates of the density function for the frequency of earnings changes, allowing us to evaluate this expression. The authors provide two sets of estimates: one based on raw, reported earnings (what we use in this study) and another based on earnings that have been cleaned to remove measurement error.[23] For a given individual, reported earnings can vary over time due to contractual wage changes, changes in real take-home pay unrelated to wage (like overtime or commission), and measurement error. Using this measure we expect our estimate to be 0.25 times the true coefficient. Ideally, we would calculate the distribution of earnings changes using only changes in take-home income caused by productivity, but due to the presence of measurement error, we view 0.25 as a lower bound on the attenuation of our estimate. There is one important caveat: since Barattieri et al. (2014) provide estimates of earnings changes only at 4 month intervals, this bound could overstate attenuation because it under-weights changes that occur in less than 4 months. Underweighting these high-frequency changes could substantially overstate bias because our estimation strategy is most accurate for individuals with frequent changes in earnings.

Using the cleaned series from Barattieri et al. (2014), the bias expression, Equation (16), evaluates to -0.006, indicating that our estimate would be fully attenuated. The cleaned series removes measurement error but also likely removes real take-home pay changes, which would raise the frequency of earnings changes. Thus, we view this as an upper bound on the degree of

[23]The cleaning procedure involves applying time-series tests for structural breaks at all possible dates in a series. The authors explain, "The adjusted series have wage changes at all dates where we can reject the no-break hypothesis, and are constant otherwise. This results in the exclusion of many instances of transitory wage changes that look very much like measurement error." For more detail, see Barattieri et al. (2014).

attenuation. In conclusion, we expect, *a priori* that our short-run estimates should be between 0 and one-quarter of the true parameter value. The long-run estimator captures permanent shifts in sunset time and sleep, and it is thus unaffected by this source of bias. If the same structural model relates long and short-run sunset to wages, then we also expect that the short-run estimates will be less than one-quarter the size of the long-run estimates.

B.2.1 Derivation of short-run bias

Here we derive expressions for the expected bias in our short-run estimates. Although we state the results in terms of wages and sleep, the model applies equally well to estimates of the reduced form equation (6). We also present the model in the form of a univariate equation, which can either be viewed as our reduced form or 2SLS models residualized on the short-run control set or as a statement about the unconditional behavior of the estimator.[24]

Assume that for a given individual i surveyed on day t, wages are equal to the average of D past sleep observations plus random noise. Therefore the true model relating sleep to wages is

$$\begin{aligned} w_{it,\tau} &= \beta \left(D_i^{-1} \sum_{k=\tau-T}^{\tau-1} T_{S,ik} \right) + \varepsilon_{i\tau} \qquad (9) \\ &= \beta T^*_{S,i\tau} + \varepsilon_{i\tau} \qquad (10) \end{aligned}$$

Thus, we are assuming that earnings change for this individual every D days, and sleep only matters during the earnings determination period. The subscript τ indexes the day that these earnings start to be *observed* in the data. Because of the fixed earnings change frequency for a given individual, these earnings will be observed for days τ through $D + \tau - 1$. To be concrete, consider the case of $D = 2$. Then we, the researchers, can only sample the individual on either the day after they received an earnings change or 2 days after, so τ will be equal to t or $t - 1$.

We further assume that τ is uniformly distributed across the year (a person has an equal probability of receiving an earnings change on any given day). This is a strong assumption, but the best available evidence from Barattieri et al. (2014) suggests that it is not broadly incorrect. Of course, for a given year, there will be weekend or holiday effects, but asymptotically, these become less relevant. Moreover, we do not have any information on when a given individual in our sample last experienced an earnings change, so this uniform assumption is a non-dogmatic

[24]This model could also apply to other covariates, so we implicitly treat a worker's observable characteristics on date t as the correct observables at the time of wage setting. Since many such characteristics are fixed or vary extremely slowly (for example race, occupation, and gender), we believe this assumption is benign.

baseline.

Finally, we assume that the researcher has isolated exogenous variation in sleep so that $\mathbb{E}T_{S,t}\varepsilon_\tau = 0$ for all t and τ.

If we observed past sleep and knew the earnings change frequency, we could estimate Equation (9) and return the correct estimate. Instead, we observe wages and sleep on date $t \geq \tau$, with which we estimate

$$w_{it,\tau} = \beta_1 T_{S,it} + \varepsilon_{i\tau}$$

We wish to know the relationship between β_1 and β.

We will exploit the wage setting structure given above and the functional form for the time series of sunset time from Section A to calculate this relationship. First, given the results from Frazis and Stewart (2012) we can use sunset time to both isolate daily, exogenous variation in sleep and to predict daily sleep for any day of the year, even though we only observe sleep on one day. This individual time series of sleep will have a similar functional form to the instrument, namely

$$T_{S,it} = A\cos(\theta t)$$

where A is the population coefficient on the unconditional version of the first stage Equation (5) and where we drop an ignorable, uncorrelated error term. The value $\theta = 360/365$ scales the wavelength to one year, so we make an additional simplification by assuming that a year is 360 days long so that this term can be ignored. Alternatively, one could, as we do when we analytically calculate the bias, rescale t to incorporate the term. Thus

$$T_{S,it} = A\cos(t) = A\cos(\tau + j) \tag{11}$$

where $j = t - \tau$ is the number of days since the latest earnings change for this observation.

We apply Lagrange's identity to rewrite earnings-relevant sleep.

$$T^*_{S,i\tau} = D_i^{-1}\sum_{k=1}^{D_i} A\cos(\tau) = \frac{A}{D_i 2\sin(1/2)}(\cos(\tau - D_i + (\pi - 1)/2) - \cos(\tau + (\pi - 1)/2))$$

The two cosine functions are simply phase shifts of each other, so we apply phasor addition to reduce this to

$$T^*_{S,i\tau} = AB_1^2 \cos(\tau + \omega) \tag{12}$$

where

$$B_1^2 = \frac{(\cos((\pi-1)/2 - D_i) + \cos((\pi-1)/2))^2 + (\sin((\pi-1)/2 - D_i) + \sin((\pi-1)/2))^2}{2D_i \sin(1/2)}$$
$$\omega = \arctan(\cot((D_i + 1)/2))$$

This form is convenient because observed sleep can now be written as a phase shift of earnings-relevant sleep and thus suggests that a version of the error-in-variables formula will apply in this setting since we can linearly relate observed sleep to earnings-relevant sleep plus a correlated error term.

Now note that the only individual heterogeneity is in terms of the frequency of earnings changes, so without loss of generality, we can replace the D_i index with just D. Then, applying the usual variance-covariance formula for the OLS estimator of a single coefficient, we have that our estimator relative to the true coefficient is given by[25]

$$\operatorname{plim} \hat{\beta}_{1,D} = \frac{\operatorname{Cov}(w_D, T_{S,D})}{\operatorname{Var}(T_{S,D})} \tag{13}$$
$$= \beta \frac{\operatorname{Cov}(T^*_{S,D}, T_{S,D})}{\operatorname{Var}(T_{S,D})} \tag{14}$$

Where we have dropped the time subscripts based on the calculations below.

To derive closed-form expressions for Equation (14) observe that since seasonal sleep is mean zero, and, for all D, we are equally likely to observe sleep on any day of the year, then by the double angle formula and Lagrange's identity, the denominator is

$$\begin{aligned}\operatorname{Var}(T_{S,D}) &= \lim_{T\to\infty} T^{-1} \sum_{t=0}^{T} A^2 \cos^2(t) \\ &= \frac{A^2}{2} + \lim_{T\to\infty} \frac{A^2 \csc(1)\sin(2T+1) + 3}{T} = \frac{A^2}{2}\end{aligned}$$

Where T (not sleep time, T_S) is the total number of time observations and the last equality

[25] We loosely call this attenuation even though in practice the estimate can be negative even when the true coefficient is positive.

follows from the boundedness of sine.

The numerator is, by application of the product-to-sum and Lagrange identities

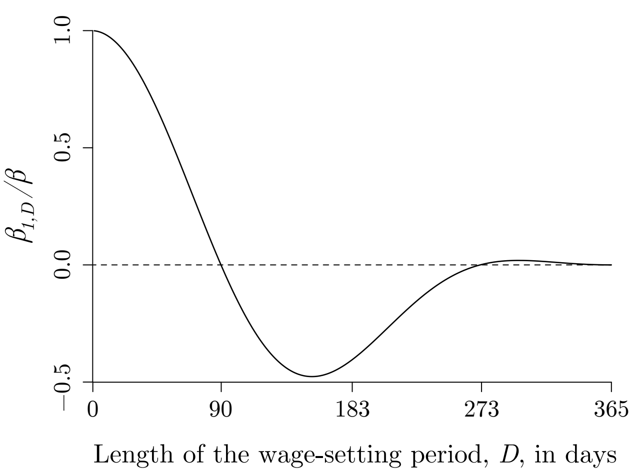
$$\begin{aligned}
\mathrm{Cov}(T^*_{S,D}, T_{S,D}) &= D^{-2} \sum_{k=1}^{D} \sum_{j=1}^{D} \lim_{T\to\infty} T^{-1} \sum_{\tau=0}^{T} A^2 \cos(\tau - k)\cos(\tau + j) \\
&= D^{-2} \sum_{k=1}^{D} \sum_{j=0}^{D} \lim_{T\to\infty} T^{-1} \sum_{\tau=0}^{T} A^2 \cos(\tau - k)\cos(\tau + j) \\
&= \frac{A^2}{2D^2} \sum_{k=1}^{D} \sum_{j=0}^{D} \cos(k + j)
\end{aligned}$$

Taking the ratio of these two values gives the relative bias of the short-run estimate with respect to the true estimate for a given D.

$$\hat{\beta}_{1,D} \xrightarrow{p} \beta \left(D^{-2} \sum_{k=1}^{D} \sum_{j=0}^{D-1} \cos(k + j) \right) \tag{15}$$

Figure 5 shows this value for all frequencies of earnings changes less than a year. This shape

Figure 5: Short-Run estimate versus true estimate



Notes: The figure shows the ratio of the probability limit of the short-run estimate to the true estimate on the y-axis for a range of possible frequencies of earnings changes on the x-axis. A value of 1 on the y-axis indicates no bias, while a negative value indicates that the estimated coefficient has the wrong sign.

is the result of two factors. First, for any two of the same sinosoidal functions that are phase shifted from each other by less than a quarter or more than three-quarters of a wavelength, the product will be positive because the two functions are "in phase enough". For a phase shift greater than one-quarter but less than three-quarters of a wavelength, the product will be negative. The attenuation of the estimate is an average of these products, so for frequent earnings changes (small D), we are largely averaging sleep that is less than a quarter wavelength off from the truth. For intermediate values of D, we are averaging in sleep that is phase shifted enough to flip the sign on the estimate. For D near a year, however, we have "crossed the hump" again and are averaging in sleep values that are phase shifted so much that they are back to the beginning of the cosine wave. Beyond $D = 365$, the estimate remains nearly fully attenuated, with slight oscillations around 0.

The expected attenuation for the full population will depend, therefore, only on the distribution of the frequency of earnings changes. Assuming that the frequency of earnings changes, D, is a discrete random variable, this expectation can be calculated by a sum over D, weighted by the probability of observing that frequency.

$$\hat{\beta}_1 \xrightarrow{p} \beta \sum_D D^{-2} \sum_{k=1}^{D} \sum_{j=0}^{D-1} \cos(k + j) \Pr(D) \tag{16}$$

Barattieri et al. (2014) provide estimates of this density function (derived from Figures 12 and 13), with which we can calculate Equation (16).

Finally, a note on alternative assumptions about the earnings determination process (Equation (9)): if earnings are based more on recent productivity rather than historical productivity (for instance if the manager is myopic when writing wage contracts), then our estimate will be closer to the true coefficient because we will be more likely to average together observed sleep that is less than a quarter-wavelength phase-shifted from the truth. If earnings are based on longer-term sleep or productivity patterns (for instance, the manager is very slow to update the wage contract and needs two earnings change cycles to fully incorporate current productivity changes), then our estimate will either be more biased or will be more likely to be attenuated all the way to zero.

B.2.2 Estimates of short-run bias

The timing mismatch model discussed above provides testable predictions about the relative magnitudes of short-run reduced form estimates from different populations. In particular, estimates based on workers with more flexible earnings should be closer to the long-run estimates presented in Table 5.[26] Recent research by Barattieri et al. (2014) provides information on wage and earnings flexibility for three groups in our data that can be used to test predictions from the model.

In particular, Barattieri et al. (2014) show that hourly wages exhibit less flexibility than weekly earnings by salaried workers before accounting for measurement error. This suggests that our reduced form estimate should be smaller for hourly wages than for weekly earnings reported by workers who do not receive an hourly wage. Columns 1 and 2 of Table 12 show that this is indeed the case.

Table 12: Relative attenuation of the short-run reduced form

	(1) Hourly workers ln(wage)	(2) Non-hourly workers ln(earnings)	(3) Hourly workers ln(earnings)
Daily sunset	-0.0020	-0.0042*	-0.0064***
	(0.0016)	(0.0025)	(0.0019)
Observations	32,040	30,232	30,921

Notes: The table shows results from estimating Equation (6) on daily ATUS data for three sub-samples. The dependent variable is indicated at the top of each column. Earnings refers to "usual weekly earnings" while wage refers to hourly wage. Controls are discussed in Section 3.1 and are location fixed effects; race indicators; age indicators; a gender indicator; indicators for holiday, day of week, and year; and occupation indicators. Standard errors, clustered at the FIPS code (location) level, are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

After accounting for measurement error, Barattieri et al. (2014) show that the corrected hourly wages are more flexible than corrected weekly earnings. This implies that when using the same measure of earnings for both hourly workers and non-hourly workers, the hourly worker sample should have a higher reduced form coefficient than the non-hourly worker sample. Comparing Columns 2 and 3 from the table shows that this prediction is also borne out.

When we evaluate the attenuation Equation (16) using the uncorrected series from Barattieri *et al*, we expect the non-hourly worker estimate to be 1.4 times the size of the hourly worker estimate. Indeed, when we compare estimates from Columns 1 and 2, the non-hourly worker reduced form estimate is 1.5 times larger.[27] Such close accordance with the Barattieri *et*

[26] The relative magnitudes should, strictly speaking, be compared to the long-run estimates from the same sub-samples, but the reduced form estimates for each of these groups is so similar that the ordering of coefficients does not change.

[27] This result has further implications for the amount of measurement error in CPS wage variables that is beyond the

al.-derived prediction provides evidence that the timing-based attenuation model is accurately capturing important aspects of the sunset-earnings relationship.

Table 13: Union and government workers have smaller reduced form estimates

	(1) ln(earnings)
β_1: Union member	0.16***
	(0.0093)
β_2: Government employee	0.053***
	(0.0071)
β_3: Sunset time	-0.0063***
	(0.0019)
β_4: Union × sunset	0.0087**
	(0.0044)
β_5: Government × sunset	0.0049
	(0.0043)
Observations	61,161
$\beta_3 + \beta_4$	.0024
P-value for test $\beta_1 + \beta_4 = 0$	0.60
$\beta_3 + \beta_5$	-0.0014
P-value for test $\beta_1 + \beta_5 = 0$	0.73

Notes: The table shows results from estimating Equation (6) on daily ATUS data for four subsamples. The dependent variable, earnings, refers to "usual weekly earnings." Controls are discussed in Section 3.1 and are location fixed effects; race indicators; age indicators; a gender indicator; indicators for holiday, day of week, and year; and occupation indicators. Standard errors, clustered at the FIPS code (location) level, are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

We can further test the model by comparing reduced form estimates for groups known to have a high degree of wage or earnings rigidity. Two groups that have long been known to exhibit especially high nominal wage rigidity are unionized workers and government employees (Hall, 1975). If wage rigidity translates into earnings rigidity, then in the context of the model, this rigidity implies that the reduced form estimates for union members and government employees should be closer to zero than the estimates for other workers.

To test this prediction, we add four new variables to Equation (6): an indicator for whether a worker is in a union, an indicator for whether the worker is a government employee, the interaction between sunset time and union status, and the interaction between government

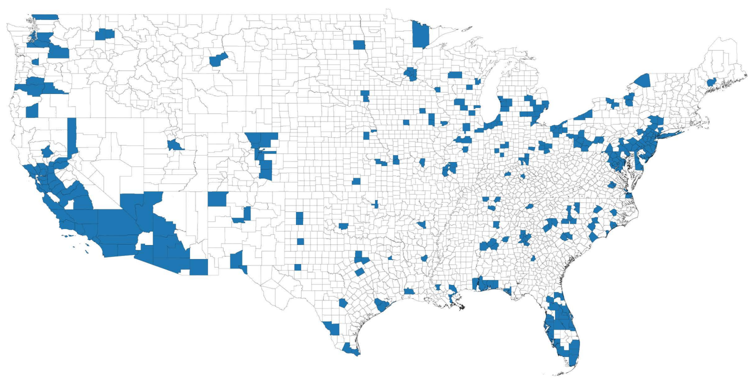
scope of this paper.

employment and sunset time.[28] Given the overall negative effect of sunset time on earnings, the model predicts that the two interaction terms will be positive and that the sum of the sunset time effect and the interaction terms will be close to zero. We show the results of this test in Table 13. As the model predicts, the effect of sunset time on earnings for both union members and government employees is substantially attenuated.

C Additional ATUS results

Figure 6 highlights counties in which we observe a respondent's location at county, rather than CBSA or state level.

Figure 6: ATUS county-level geocoding



Notes: The map shows, in blue, locations in the continental United States where we are able to geocode ATUS records at the county level.

Sunset time effects on bedtime and wake time are in Table 14. Naps may also be affected, and results including naps are below in Sections C.2 and C.3. In our sample of full-time workers, 10% report sleeping sessions that begin and end during daylight hours on the diary date. This sleep could be naps, night-shift workers, or individuals who sleep early or late but experience disrupted sleep. Median total daytime sleep is approximately 1.5 hours per week and mean total daytime sleep is 2 hours per week.

[28] Union membership and government employment are recorded during the CPS interview, not the ATUS interview, so it could be the case that individuals change status between the two interviews. In that case, we will have measurement error on our classifications. 8% of the estimation sample report a different detailed industry job during the ATUS and CPS interviews. Excluding this group does not substantively change the results.

Table 14: Bedtime and wake time

	(1) Bedtime	(2) Wake time
Panel A: Short-run		
Daily sunset time	0.12***	-0.034***
	(0.0050)	(0.0042)
Panel B: Long-run		
Avg. sunset time	0.46***	0.31***
	(0.039)	(0.035)

Notes: Dependent variable is given at the top of the column. Controls, number of observations, and standard errors are the same as in Table 4 (Panel A) and Table 5 (Panel B). Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

C.1 Results for part-time workers

Table 15 tests whether sunset time has an effect on whether a worker works part or full time. The table reports two regressions that are variants of the reduced form specification where the dependent variable is an indicator equal to one if the worker works part time. The results show that in the short run, there is a small decrease in the probability of a worker being part time if daily sunset time is later. In the long run, the reverse effect holds, but the estimate is less precise. Overall, there is not a strong relationship between employment status and sunset time.

Table 15: Sunset effect on part-time employment

	(1) Part-time Short-run	(2) Part-time Long-run
Daily sunset	-0.0016**	
	(0.00081)	
Average sunset		0.0011
		(0.0064)
Obs	85,243	529

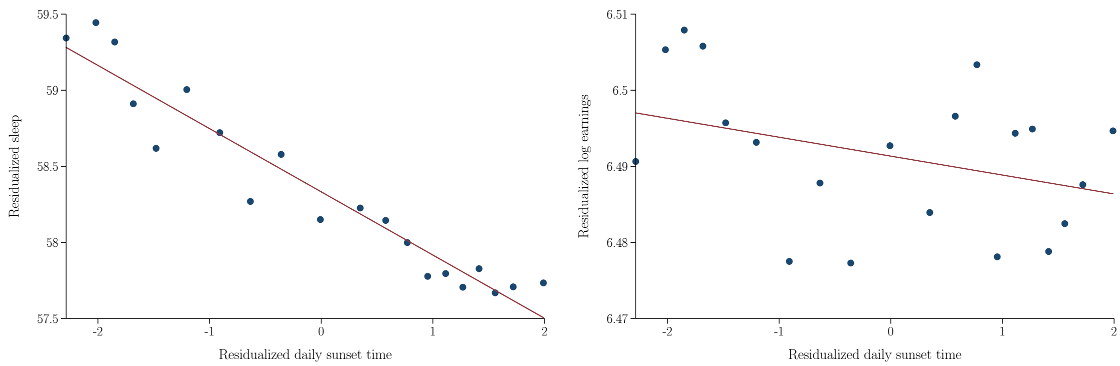
Notes: Dependent variable is given at the top of the column. Controls, number of observations, and standard errors are the same as in Table 4 (Column 1) and Table 5 (Column 2). Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

C.2 Additional short-run ATUS results

The following additional robustness checks are discussed in Section 5.3.1.

Figure 7: Short-run estimates of the conditional expectation function

(a) First stage (b) Reduced form



Notes: Created using binscatter. Sunset time is divided into 20 bins by quantile and means for the y-axis variables are computed within each. Sample and controls are the same as in Table 5, so the fitted line (dashed) which is estimated using OLS, corresponds to the estimates presented in that table.

Table 16: Additional short-run robustness checks

	ln(earnings)			ln(earnings)	
Only time controls and location FEs			*No occupation codes*		
Daily sunset	-0.0048**	(0.0021)	Daily sunset	-0.0037*	(0.0019)
No weekends			*No 4th or 1st quarter*		
Daily sunset	-0.0038	(0.0024)	Daily sunset	-0.016**	(0.0062)
Observations	30,041		Observations	29,895	

Notes: The table shows results from estimating Equation (6). Dependent variable is indicated at the top of each column. Unless otherwise noted, controls, number of observations, and standard errors are the same as in Table 4. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

As discussed in Section 4, ATUS documentation implies that the within-location distribution of interviews over the year is uniform in expectation and we test this implication in Figure 9. Monthly interview shares were computed within each location. For example, the January

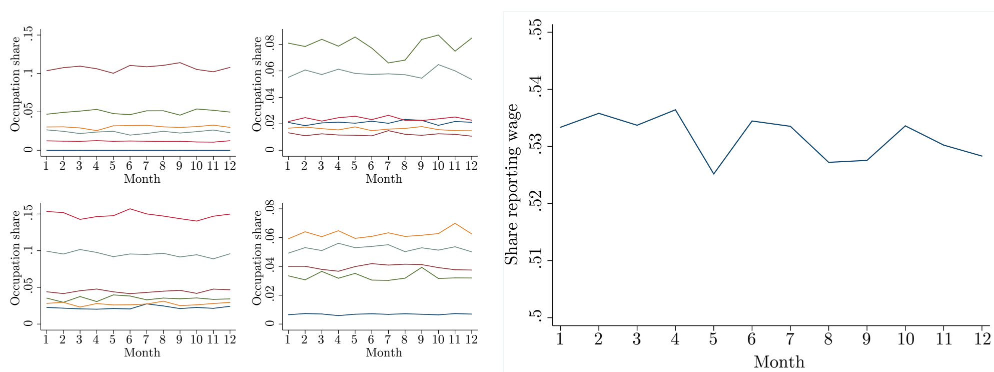
Table 17: Short-run ATUS reduced form effects of sunset on earnings

	(1) ln(earnings)	(2) ln(earnings)	(3) ln(earnings)	(4) ln(earnings)	(5) ln(earnings)
Daily sunset time	-0.0044***	-0.0033**	-0.0034**	-0.0045*	-0.0053
	(0.0017)	(0.0016)	(0.0016)	(0.0023)	(0.0038)
Education	No	Yes	Yes	Yes	Yes
Married, children	No	No	Yes	Yes	Yes
Temperature	No	No	No	Yes	Yes
Quarter FE	No	No	No	No	Yes

Notes: Dependent variable is usual weekly earnings. Unless otherwise noted, controls, number of observations, and standard errors are the same as in paper Table 4. "Education" adds 4 educational attainment indicators; "Married, children" adds controls for marital status, number of children, and detailed race categories; "Temperature" adds interview day average temperature; "Quarter FE" adds indicators for each quarter of the year. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Figure 8: ATUS occupations do not exhibit seasonality

(a) Occupation shares by month (b) Employment rate by month

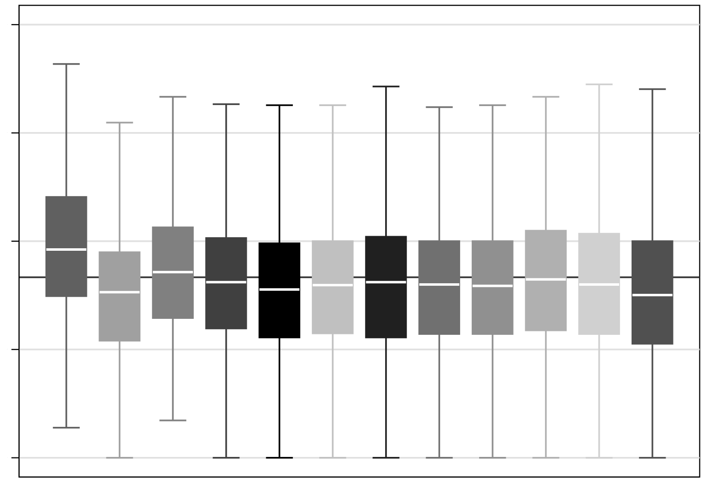


Notes: Panel (a) Lines show average values of our 23 occupation dummies by month, pooled over the period 2003-2013 for our estimation sample. The occupation exhibiting a modest summer dip in the upper-right panel is "Arts, design, entertainment, sports, and media occupations." Excluding this occupation does not change our results. Panel (b) Line shows average value of dummy that equals 1 if the respondent reports a non-zero weekly or hourly wage, by month, pooled over the period 2003-2013 for all ATUS respondents.

share in Los Angeles County might be .1, while the January share in King County might be .06. This box plot illustrates the distributions of these monthly shares over locations. Each distribution is centered close to the uniform level of .083, and the interquartile range is always less than .05. There is no evidence that the ATUS sampling procedure introduces seasonal biases in our estimates.

Table 18 shows first stage, reduced form, and IV estimates of where the sleep measure is

Figure 9: ATUS interview shares by month within location



Notes: Monthly interview shares were computed within each location. For example, the January share in Los Angeles County might be .1, while the January share in King County might be .06. This box plot illustrates the distributions of these location-level monthly shares. White horizontal lines represent medians. Boxes represent extend to the 25th and 75th percentiles. Whiskers are constructed by multiplying the interquartile range by 1.5, then adding it to the 75th percentile for the upper whisker and subtracting it from the 25th percentile for the lower. A uniform distribution of interviews over months within location would imply a share of 1/12=.083 for each month and the horizontal black line behind the boxes marks this value.

total daily sleep, including daytime sleeping. As expected, the first stage is smaller using this measure, leading to a larger IV estimate.

Table 18: Short-run effects from ATUS, sleep including naps

	(1) First stage	(2) Reduced form	(3) IV
Daily sunset	-0.20*** (0.044)	-0.0044*** (0.0017)	
Sleep and naps			0.022** (0.010)
Observations	61,161	61,161	61,161

Notes: The table shows results from estimating Equation (5) (columns 1 and 3) and Equation (6) (columns 2) on ATUS data. The dependent variable is indicated at the top of each column. Earnings refers to "usual weekly earnings" and wage to hourly wage for those workers who reports one. Sleep is measured in hours per week, including naps, and sunset time in hours. Controls are discussed in Section 3.1 and are location fixed effects; race indicators; age indicators; a gender indicator; indicators for holiday, day of week, and year; and occupation indicators. Standard errors, clustered at the FIPS code (location) level, are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

C.2.1 Assessing seasonal confounding using the known functional form of sunset time

One can use the approximate seasonal pattern of sunset time to parsimoniously assess seasonal confounding. This method is similar to including seasonal fixed effects, as in the above robustness checks, but it requires only a single additional parameter and it can be used to determine robustness of the sunset-wage relationship to seasonal patterns that peak at different times of year.

An omitted variable with a seasonal pattern similar the one exhibited by sunset could bias our estimate.[29] A seasonal variable with a very different phase or frequency poses no omitted variable problem, because the correlation with sunset time will be low.[30] Between these two extremes, a non-zero amount of confounding can occur. We can characterize this potential bias by including seasonal control variables that mirror the sinusoidal pattern of daily sunset time: $x_{\theta,\varphi} = -\cos[(360/365)(d + 10 + \theta)\varphi]$, where $\theta \in \mathbb{R}$ is the phase shift relative to sunset time and $\varphi > 0$ is the frequency relative to sunset time. We focus first on the set of variables where the relative frequency is the same ($\varphi = 1$) and the phase shift, θ, ranges over the set of natural numbers less than 183.[31] We then, one at a time, include these variables as additional covariates when estimating Equation (6). The black curve in Figure 10a shows the resulting coefficients on sunset time from these regressions along with the 95% confidence interval in dashed lines. The red horizontal line shows the reduced form coefficient reported in the baseline results in Table 4.

One can see that over a wide range of seasonal controls, the reduced form estimate is essentially unchanged from the baseline. For controls that are phase shifted less than about 50 days, the confidence interval widens to include zero, but in all cases, the sign of the point estimate is the same as in the baseline estimate. Overall, our estimate is robust to sign error for any seasonal confounder that has the same frequency but differs by even a day from sunset time.

Figure 10b shows the same result, but now for a set of regressions that include covariates

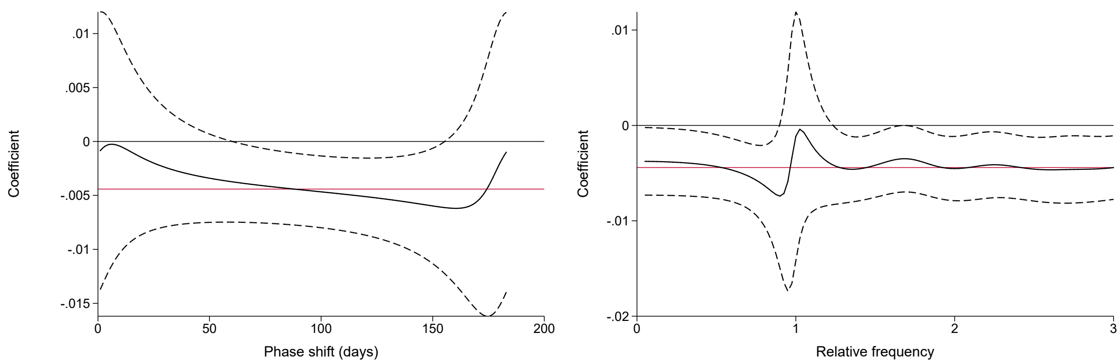
[29]As discussed in Section 3.2.2, daylight savings time breaks the smooth sinusoidal pattern of sunset time, potentially allowing for identification of the short-run effect even in the presence of seasonal confounders, but daylight savings time does not, by itself, change sleep over a long enough period to credibly estimate wage effects.

[30]For example, a variable with the same frequency but phase shifted by one-quarter wavelength will have zero correlation with daily sunset time.

[31]Since sunset time has a period of one year, this set of phase shifted variables will cover the full range of possible correlations with sunset time. In theory, we can let θ range over $\mathbb{R}^+$, but our highest data frequency is daily.

Figure 10: Robustness of short-run ATUS reduced form to seasonal controls

(a) Phase shifted control (b) Frequency control



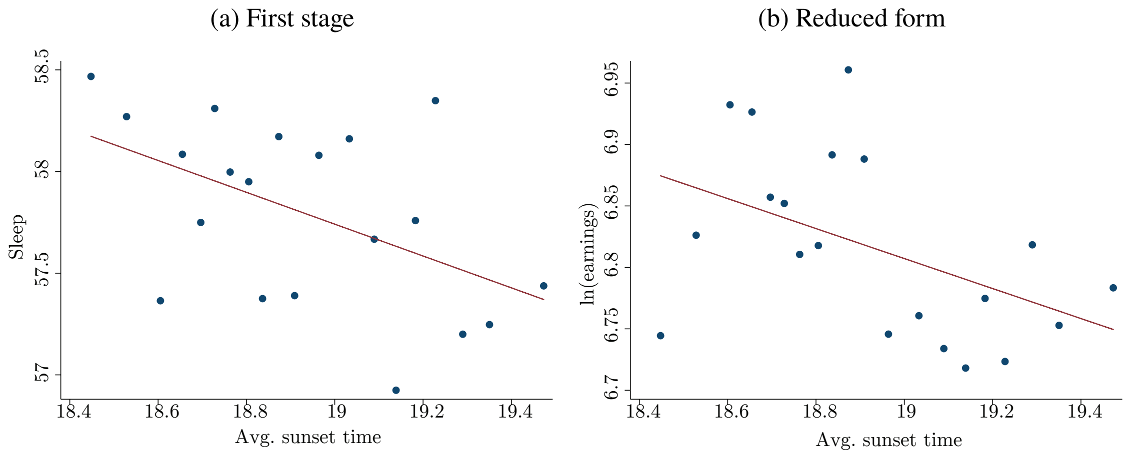
Notes: Panel 10a shows estimates of the coefficient on sunset time in the reduced form Equation (6) when a phase-shifted seasonal control variable is included. The degree of phase shifting is indicated on the x-axis. Panel 10b shows estimates of the same coefficient when seasonal control variables with differing relative frequency are included. 95% confidence intervals are shown by dashed lines and the baseline reduced form estimate is given in red.

with different relative frequencies ($\theta = 0$ and $\varphi \in (0, 3)$). Here, one can see that the reduced form estimate is highly robust to seasonal patterns with alternative frequencies. Outside of a narrow band of relative frequencies between 0.96 and 1.09, the estimate has the expected sign.

In principle we cannot rule out the possibility that the short-run estimates are the result of a spurious correlation between an unrelated seasonal pattern in wages that closely matches the frequency and phase of sunset time. Practically, however, Figure 10 suggests that such a variable would need to exhibit a pattern that is strongly similar to sunset time to introduce large bias. The corroborating results we recover using the orthogonal identification provided by long-run sunset further mitigate confounding concerns.

C.3 Additional long-run ATUS results

Figure 11: Long-run estimates of the conditional expectation function



Notes: Created using binscatter. Sunset time is divided into 20 bins by quantile and means for the y-axis variables are computed within each. Sample and controls are the same as in Table 4, so the fitted line (dashed) which is estimated using OLS, corresponds to the estimates presented in that table.

Per the recommendation in Solon et al. (2015), we conduct a modified Breusch-Pagan test for heteroskedasticity of the residuals from the unweighted 2SLS model. The results in Table 19 show that location-level observations with smaller underlying counts of ATUS observations exhibit higher variance, as expected, and the relationship is statistically significant at the one percent level. While the constant term is statistically significant, it is an order of magnitude smaller. This suggests that the common error component within location is minimal, so weighting will likely result in an efficiency improvement, and indeed that is what we see in Table 20.

Table 20 reproduces our long-run results from Table 5 above their unweighted counterparts. Weighting does indeed improve efficiency, reducing the standard errors in both the first stage and reduced form models. It does not appreciably alter the first-stage coefficient on sleep, but it does increase the magnitude of the reduced-form estimate. Note that we cannot reject a null hypothesis of equality of the weighted and unweighted estimates. The weighted and unweighted estimators have the same probability limit under the assumption of homogeneous treatment effects, however, so we investigate further below.

There is no guarantee weighted and unweighted estimators will produce similar estimates

Table 19: Modified Breusch-Pagan heteroskedasticity test

	Residuals^2
1/Observations	0.10***
	(0.020)
Constant	0.011***
	(0.0014)
Observations	529
Adjusted R^2	0.045

Notes: The dependent variable is the squared residual from estimating the unweighted version of (8). The variable "1/Observations" is the reciprocal of the number of ATUS interviews underlying a given location-level observation. Because the modified Breusch-Pagan test relies on the assumption of homokurtosis, we compute unmodified OLS standard errors. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

in a finite sample, and indeed sampling variance appears to drive the difference in the estimates. Table 21 reports unweighted reduced-form results from two subsamples: locations with fewer or more than the median 58 individual observations. In the sample of many-observation locations the reduced form estimate is -.047, unsurprisingly very close to the weighted estimate from the full sample, and statistically significant at the 10% level. In the sample of few-observation locations the estimate is positive (.041), but the standard error is large (also .041). While the many-observation locations have higher population, on average, than few-observation locations, there is substantial overlap in the distributions. It appears that the difference between weighted and unweighted results arises from sampling variance, rather than heterogeneity in the treatment effect under study.

We have also performed a variety of additional robustness checks with little or no change in estimates. We list them here without full tables, but all results are available upon request. 0.2% of the sample has topcoded wages. A tobit accounting for this does not change the results. Likewise, accounting for the truncation of sleep does not change inference. We have also estimated the models on only the sub-sample that is geocoded at the county or CBSA level. All of these robustness checks do not change inference.

In Table 22 we report estimates of county level characteristics as functions of average sunset time. We find a large and statistically significant relationship with population density, which motivates our use of a flexible control for this variable in estimating long-run effects. We also see a statistically significant relationship with unemployment, consistent with our estimated

Table 20: Long-run effects, weighted and unweighted

Panel A: Weighted			
	First stage	Reduced form	2SLS
	Sleep	ln(earnings)	ln(earnings)
Avg. sunset time	-0.93***	-0.045***	
	(0.28)	(0.017)	
Sleep			0.048**
			(0.023)
Observations	529	529	529
Adjusted R^2	0.125	0.811	0.699
F-stat on IV			11.22

Panel B: Unweighted			
	First stage	Reduced form	2SLS
	Sleep	ln(earnings)	ln(earnings)
Avg. sunset time	-0.86**	-0.0013	
	(0.44)	(0.022)	
Sleep			0.0015
			(0.025)
Observations	529	529	529
Adjusted R^2	0.119	0.648	0.650
F-stat on IV	3.86		

Notes: The table shows results from estimating Equation (8). In Panel A location-level observations are weighted by the count of underlying ATUS respondents, while in Panel B they are unweighted. The dependent variable is indicated at the top of each column. Earnings refers to "usual weekly earnings". Controls are as reported below Table 5. White's robust standard errors reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

long-run wage effect. Similarly, Table 23 regresses our long-run sunset time instrument on the coastal distance and population density controls employed in our preferred specification. As discussed in the body of the paper, long-run sunset time is endogenous to these variables and indeed they explain roughly one fourth of the variation in the long-run instrument.

We report additional long-run ATUS robustness checks in Table 24. First, we remove all weekend diary observations. ATUS oversamples weekends so that roughly half of the total observations are from weekend dates (see Table 2). We test the sensitivity of our results to this by dropping the weekend diary entries entirely. The estimate is similar to baseline, albeit less precise. While the number of location-level observations is the same, this specification drops roughly half of the underlying ATUS sample. Second, we add education controls, and

Table 21: Unweighted reduced form, subsamples of locations below and above median observation count

	(1)	(2)
	<58 obs	>=58 obs
Avg. sunset time	0.041	-0.047*
	(0.041)	(0.025)
Observations	264	265

Notes: The dependent variable is log weekly earnings. Earnings refers to "usual weekly earnings." Sunset time in hours. Controls are discussed in Section 3 of the paper and are splines in population density and coastal distance; race indicators; age indicators; a gender indicator; indicators for holiday, day of week, and year-month; and occupation indicators. Standard errors, clustered at the FIPS code (location) level, are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Table 22: Robustness: County characteristics

	Log pop. density	Pop. change frac.	Net migration frac.
Sunset time	-0.642***	-0.000931	-0.000980*
	(0.110)	(0.000814)	(0.000533)
Observations	3104	3104	3104
Adjusted R^2	0.012	0.000	0.001

	Log poverty rate	Labor force change	Unemployment rate
Sunset time	0.0157	0.00184	-1.412***
	(0.0221)	(0.00342)	(0.169)
Observations	3103	3103	3103
Adjusted R^2	-0.000	-0.000	0.023

Notes: Dependent variable is indicated at the top of each column. All data are from the Census and observations are at the county level. Population, net migration, and unemployment rate are all 2012 values. Poverty is from 2011. Labor force change is from 2000 to 2010. White heteroskedasticity-robust standard errors are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

finally we add industry controls. Limiting the sample to hourly workers, controlling for region, or limiting the sample to county-geocoded observations does not produce large changes in the estimate.

For controls that operate on the full estimation sample, we can also add multiple controls to see their combined effect on the estimate. Table 25 presents such a series of robustness checks for the long-run specification. One can see that the inclusion of many controls leaves

Table 23: Average sunset time, coastal distance, and population

	Avg. sunset time
Coastal distance	0.013**
	(0.0059)
Coastal county×Coastal distance	-0.19***
	(0.044)
Coastal county	-0.0099
	(0.067)
Pop. density spline - piece 1	0.000014
	(0.0000091)
Pop. density spline - piece 2	-0.0000060
	(0.000016)
Pop. density spline - piece 3	-0.000014
	(0.000018)
Pop. density spline - piece 4	0.000011
	(0.0000088)
Observations	529
Adjusted R^2	0.259

Notes: Estimates are from a regression of long-run average sunset time on the population and coastal distance controls used in our preferred specification. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

the estimate magnitude largely unchanged. The significance of the estimate falls in some specifications due to the inclusion of variables that are weakly correlated with both sunset and earnings but which do introduce noise.

Finally, Table 26 presents our preferred earnings models using all sleep recorded in the ATUS diary, including naps. Results are essentially unchanged.

Table 24: Additional robustness of ATUS long-run estimates

	Reduced form ln(earnings)	
No weekend diaries		
Avg. sunset time	-0.072***	(0.017)
Observations	3479	
Education controls		
Avg. sunset time	-0.035**	(0.015)
Industry controls		
Avg. sunset time	-0.050***	(0.018)
Hourly workers only		
Avg. sunset time	-0.061***	(0.020)
Region indicators		
Avg. sunset time	-0.049**	(0.019)
County geocoded only		
Avg. sunset time	-0.050**	(0.020)
Observations	452	

Notes: The table shows results from estimating Equation (8). Dependent variable is the log of average earnings. Unless otherwise noted, controls, number of observations, and standard errors are the same as in Table 5. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Table 25: Long-run ATUS reduced form effects of sunset on earnings

	(1) ln(earnings)	(2) ln(earnings)	(3) ln(earnings)	(4) ln(earnings)	(5) ln(earnings)
Avg. sunset time	-0.045***	-0.052***	-0.053**	-0.045*	-0.031
	(0.017)	(0.017)	(0.023)	(0.026)	(0.025)
Work hrs quadratic	No	Yes	Yes	Yes	Yes
Time zone indicators	No	No	Yes	Yes	Yes
Albouy QOL	No	No	No	Yes	Yes
Educ. attainment shares	No	No	No	No	Yes

Notes: Dependent variable is the log of average earnings. Unless otherwise noted, controls, number of observations, and standard errors are the same as in paper Table 5. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Table 26: Long-run effects from ATUS, sleep including naps

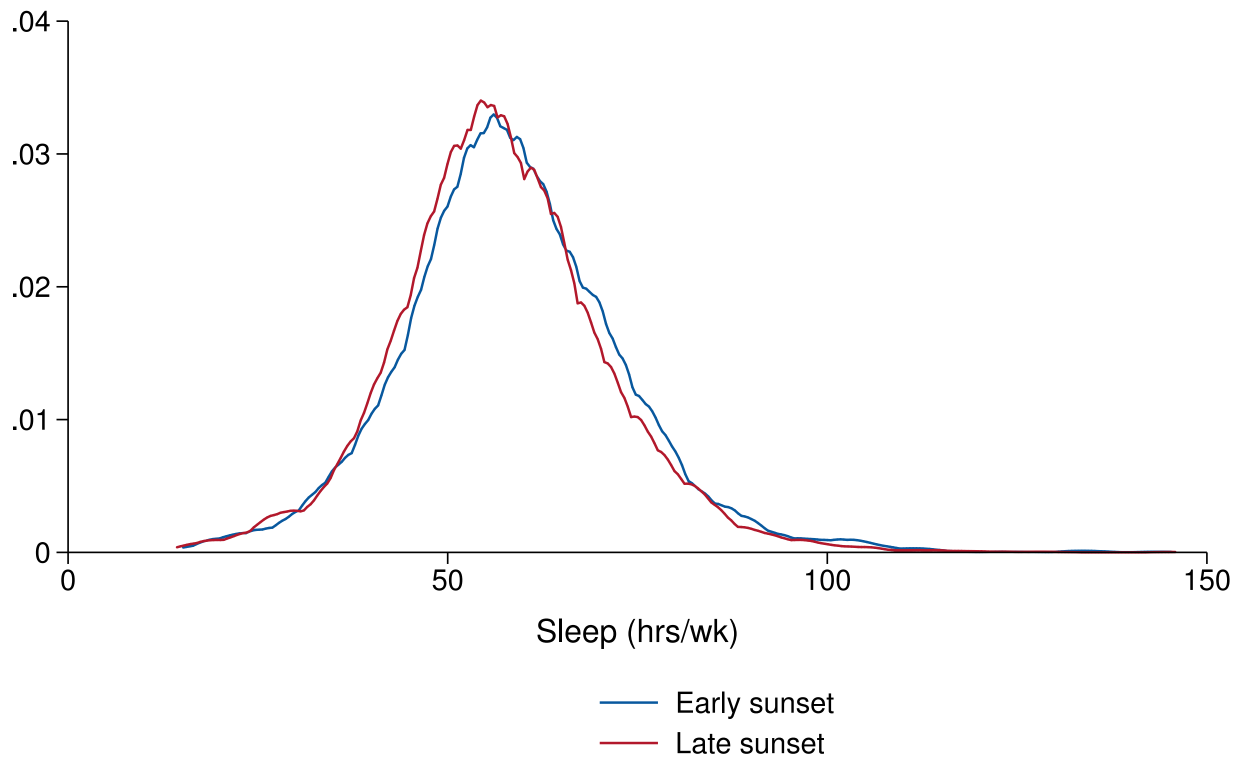
	(1) First stage Sleep and naps	(2) Reduced form ln(earnings)	(3) IV ln(earnings)
Avg. sunset time	-0.97***	-0.045***	
	(0.31)	(0.017)	
Sleep and naps			0.047**
			(0.023)
Observations	529	529	529
Adjusted R^2	0.151	0.811	0.685
F-stat on IV			9.54

Notes: The table shows results from estimating Equation (7) (columns 1 and 3) and Equation (8) (columns 2 and 4) on ATUS data. The dependent variable is indicated at the top of each column. Earnings refers to "usual weekly earnings" and wage to hourly wage for those workers who reports one. Sleep is measured in hours per week, including naps, and sunset time in hours. Controls are discussed in Section 3.1 and are: an indicator for coastal county, coastal distance, and their interaction; a ten-piece linear spline in latitude; mean age and mean squared age; percent female; race and occupation shares; and a five-piece linear spline in population density. White heteroskedasticity-robust standard errors are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1.

C.3.1 Sleep distributions by sunset time

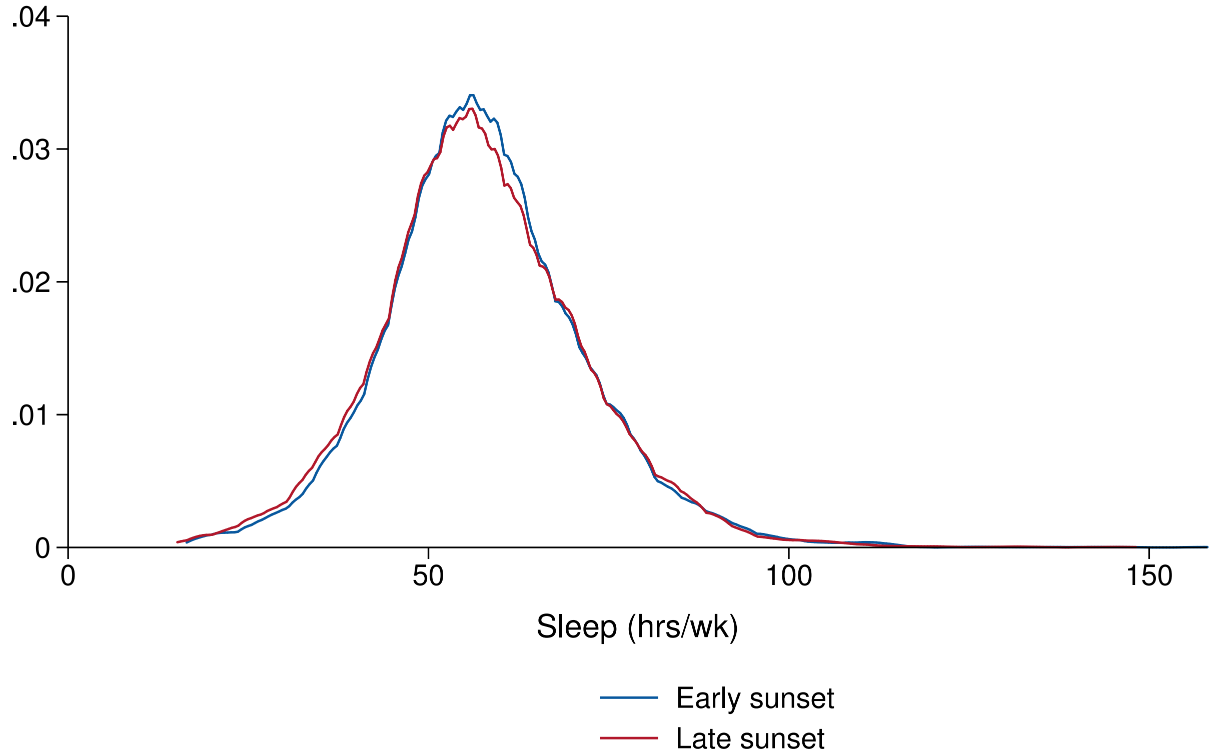
As indicated in Section 5.2, we have not found informative first-stage heterogeneity on standard observables like age and gender. We have, however, investigated the distribution of sleep among individuals with early and late sunset times. We cannot draw strong conclusions about who is affected from these graphs without assuming rank preservation, but the distributions are informative. Early seasonal sunsets shift the whole distribution rightward, increasing sleep for most quantiles. Early long-run average sunsets, by contrast, move mass out of the lower tail of sleep and into the center of the distribution. The right tail is unaffected. It is possible that long-run effects come from increased sleep among those who would otherwise be low sleepers, but we emphasize that this is speculation.

Figure 12: Sleep distributions, by early and late seasonal sunset time



Notes: "Early" denotes seasonal sunsets at or below the 10th percentile, while "late" denotes seasonal sunsets at or above the 90th percentile. Kernel and bandwidth are Stata defaults.

Figure 13: Sleep distributions, by early and late average sunset time

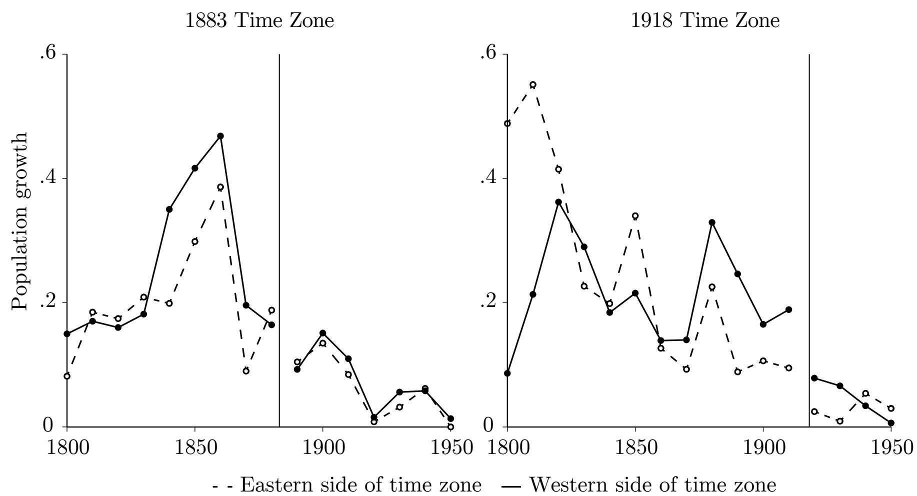


Notes: Underlying observations are individuals, not locations as in our primary analysis. "Early" denotes long-run average sunsets at or below the 10th percentile, while "late" denotes long-run average sunsets at or above the 90th percentile. Kernel and bandwidth are Stata defaults.

C.3.2 Historical sorting

As discussed in Section 3.2.3, sorting would be problematic only if it induced correlation between a determinant of productivity and long-run average sunset time. In our view this is relatively unlikely. Nonetheless we investigate the possibility of sorting. Figure 14 shows the county-level growth patterns around the dates of the 1883 and 1918 time zone implementations. For both figures, the 10% of counties that are closest to the eastern or western time zone boundary are considered to be on the eastern or western side, respectively. The dashed lines show median population growth rates (inter-census) for eastern side counties, and the solid lines show the same for western side counties. The composition of these groups differs between the two panels due to changes in the location of the 1883 versus 1918 time zones.

Figure 14: Historical time zone sorting



Notes: The figure shows median growth rates between censuses in counties on the eastern and western edges of the 1883 (left panel) and 1918 (right panel) time zones. Eastern counties are represented by the dashed line and western counties are the solid line. All data are from Haines and Inter-university Consortium for Political and Social Research (2010).

If gross sorting were occurring, one would expect eastern side counties to grow faster than western side counties after time zone implementation. Indeed, one might even expect the incentive to sort with respect to the 1883 time zones to be stronger than in the present day due to the lack of electrification. Instead, one can see that there is no evidence of gross sorting in response to the 1883 time zone. After implementation, the two regions of the time zones grow at almost identical rates. Growth rates around the 1918 law are more volatile but tell a similar story. Western side counties experience a slightly larger drop in growth rates after 1918 compared to eastern side counties, but the difference in changes between the two groups is not significant.

C.3.3 County time zone changes

State and local governments may petition the Department of Transportation (DOT) to switch time zones. The DOT criterion for evaluating proposed time zone changes is "the convenience of commerce." While most proposals succeed, some have been rejected under this standard (USNO, 2014). Counties have switched in both westerly and easterly directions, with the latter more common. Note that eastward switches are the opposite of what we expect if counties are choosing their time zone based on sleep-driven productivity considerations. Switching from being on the eastern side of a time zone to the western side moves the county from getting the "best" average sunset treatment to getting the "worst" in terms of sleep duration. Moreover, our design does not depend on the exact location of the boundary, but on the relative longitudes of cities within a time zone; the distance between the easternmost city in our data and the border is common to all observations in the time zone and does not contribute to our coefficients of interest. Proposals to switch typically cite coordination with workers or firms in an adjacent time zone. For example, the most recent such change moved Mercer County, North Dakota from the Mountain to the Central time zone in 2010. The final rule issued by DOT cited the facts that most Mercer County firms buy inputs from firms in the Central time zone and that broadcast media serving Mercer County operate on Central time (CFR, 2010).

C.4 Nonlinear effect of sleep on wages

Although our setting is not ideally suited to study nonlinear effects of sleep on earnings, it is a question of natural interest due to the routine reporting of such relationships in medical research.[32] Nonlinearity in the sleep-wage relationship is intuitively appealing—at the extremes, a worker cannot work if she sleeps all day, and this logic might well extend to shorter sleep durations that still impinge on work hours. Over more moderate sleep durations, however, the question of whether the marginal effect of sleep on wages is non-monotonic has important implications, and the answer is not obvious. For instance, in contrast to the above cited studies, Van Dongen et al. (2003) shows that the marginal effect of sleep on attention is linear over a wide range of sleep durations. Moreover, the reverse causality discussion that prefaced the

[32] See, for instance Cappuccio et al. (2010) and Leng et al. (2015) on mortality or Taheri et al. (2004) on BMI.

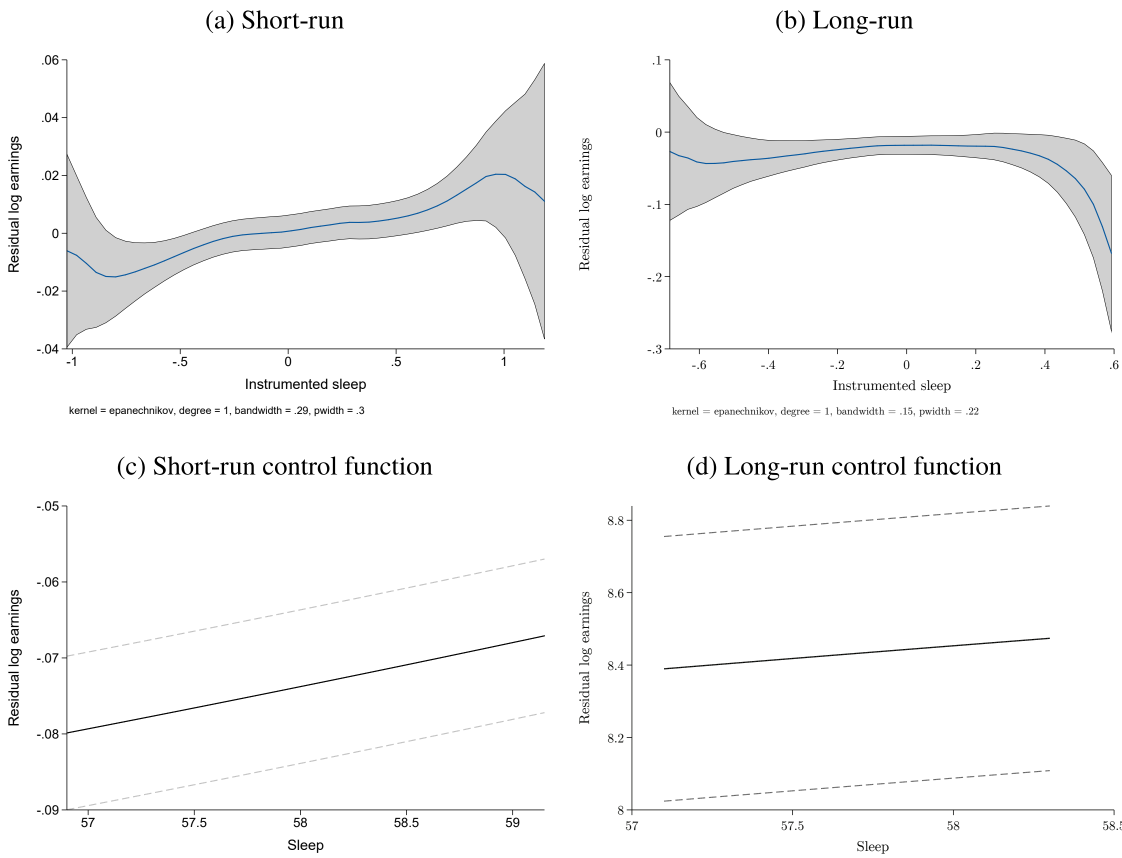
analysis of this paper suggests that workers have an incentive to sleep moderate amounts even if sleep is productivity-improving. Thus, selection into working as well as other forces might lead one to erroneously conclude that long, in addition to short sleep, is bad for productivity, health, and other outcomes.

We can empirically examine this question to some degree in our setting. We are limited by the small changes in sleep induced by sunset time relative to many laboratory studies, but we benefit from much larger sample sizes and the fact that the nonlinearity—if the medical research is correct—should be most apparent near mean sleep levels where our identifying variation exists.

We use a variety of methods to investigate nonlinearity. For the top two panels of Figure 15, in a first stage, we regress log earnings and sleep on all controls from Equation (6). We then predict the sleep residuals using daily sunset time. Finally, we fit a kernel regression through the resulting values. Under similar logic to the calculation of linear regression by fitting a line through conditional expectation functions, this figure provides a simple, semi-parametric method for showing the causal relationship between wages and sleep without imposing linearity. Note that a causal interpretation of this plot requires strong exogeneity (zero expectation of the error of log earnings conditional on sunset time) to rule out nonlinear dependence, as is typical of any nonlinear causal inference method. The bottom two panels use the control function approach of Kim and Petrin (2011), which extends the nonparametric IV methods initially proposed by Newey et al. (1999). We allow sleep to enter as a quadratic, controlling for a quadratic in the first-stage residuals and the interaction of that quadratic with our instruments. The quadratic is chosen by an information criterion.

Visually, for the short-run sample, this relationship is linear over the identified range. The Kim and Petrin (2011) estimates allow us to conduct a hypothesis test on higher order polynomial terms, and if anything, these terms suggest that the relationship is convex—more sleep has a slightly increasing marginal effect on earnings. In the long-run case, the kernel regression suggests slight concavity of the sleep-earnings relationship, but hypothesis testing of the Kim and Petrin (2011) estimates fails to reject linearity. A formal test of monotonicity based on Gutknecht (2013) also fails to reject monotonicity of effect in both the short and long-run estimates, but this test is conservative. Even though we do not find evidence for a nonlinear sleep-wage relationship, the time-intensiveness of sleep means that there is an inherent nonlinearity in the relationship between sleep and income.

Figure 15: Non-parametric causal relationship between weekly earnings and sleep



Notes: The top two panels shows two kernel regressions of residual log wage on sleep instrumented by daily sunset time. Both regressions use an Epanechnikov kernel. The left panel uses the short-run ATUS sample and a bandwidth of 0.29, and the right panel uses the long-run ATUS sample and a bandwidth of 0.14. The bottom two panels show control function estimates based on Kim and Petrin (2011). The left panel again uses the short-run ATUS sample, and the right panel uses the long-run ATUS sample.

D Additional hedonic and QCEW results

Using county-level Census data from 2010, Table 9 provides evidence that the reduced-form wage effects we identify are partially capitalized into housing prices. We regress log median county home value on average sunset time and a set of controls similar to our long-run ATUS specification.

$$\ln(\text{median home value})_j = \beta\overline{\text{sunset}}_j + \mathbf{x}_i'\gamma + \varepsilon_j$$

Table 27: QCEW summary statistics

Variable	Mean	Std. Dev.
Weekly wage	492.37	171.88
Weekly wage - goods	609.35	240.53
Weekly wage - services	431.84	161.19
Sunset time	18.38	.94
Observations	285,680	

Notes: All data are from the Quarterly Census of Employment and Wages at the county level from 1990-2013.

Appendix Table 28 below investigates the robustness of this hedonic result.

Table 28: Hedonic robustness

	Log value	Log value	Log value	Log value
Sunset time	-0.0574***	-0.0599***	-0.0432***	-0.0536***
	(0.0191)	(0.0154)	(0.0162)	(0.0188)
Base controls	Yes	Yes	Yes	Yes
Industry shares	No	Yes	No	No
Educational attainment	No	No	Yes	No
Longitude	No	No	No	Yes
Observations	2824	2824	2824	2824
Adjusted R^2	0.617	0.769	0.740	0.621

Notes: The table shows robustness checks for the hedonic results. White heteroskedasticity-robust standard errors are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1. Data are 2010 5-year ACS estimates. Sunset time is the average for a given county. Column 1 reproduces the final column of our preferred results from Table 9 and "Base controls" denotes the controls from that model. Additional controls employed in this table are a 3-piece linear spline in coastal distance, 13 industry shares, and longitude.

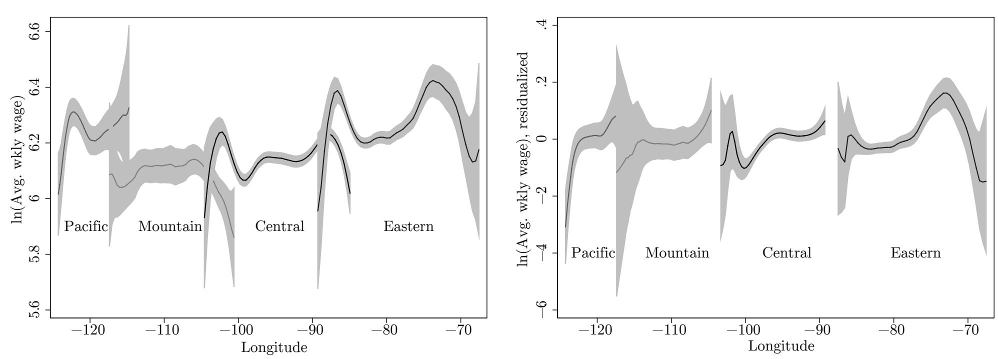
Within a time zone, average sunset time is a linear function of longitude.[33] To illustrate what the controls are doing in the QCEW estimates and to provide intuition for the sunset time instrument, we investigate the unconditional relationship between log wage and longitude within each time zone in Figure 16. Approaching the data in this way allows for an intuitive, map-like presentation, with the Pacific time zone at the left of the figure and the Eastern time zone at the right. For each time zone, we estimate a separate linear kernel regression

[33]Ignoring legal changes to the timing of daylight savings time, which did occur once during our sample period.

between longitude (and, implicitly, average sunset time) and log wage. Gray regions represent 95 percent confidence intervals. This figure strongly suggests that average sunset time is not unconditionally exogenous. Wages spike near the Pacific and Atlantic coasts. At inland time zone boundaries, where sorting is a concern (as discussed in Section 3.2) and labor markets may be integrated, the fitted relationships tend to intersect. To address these potential sources of endogeneity, we residualize log wages using time zone dummies, and interactions of the time zone dummy with coastal distance in the Pacific and Eastern time zones. This controls for the well-known gradients in population and home prices near the coast. We also exclude longitudes at which time zones overlap. Figure 16 plots a local linear relationship between residualized wage and longitude for each time zone. The pattern of results is striking, with similar linear relationships and upward slopes for all four time zones. Conditional on this parsimonious control set, western locations earn less on average than eastern locations within the same time zone.

Figure 16: Long-run effects of longitude on wages in QCEW

(a) Raw data (b) Residual wage



Notes: The figures show kernel regressions of the log earnings on longitude within time zone. The shaded area represents a 95% confidence interval around the local linear fit. While the shaded regions for the Pacific and Mountain time zones intersect, the fitted lines do not have any common support. The gap between Pacific and Mountain fitted lines is smaller than for the Mountain-Central and Central-Eastern borders because the Pacific and Mountain zones share fewer common longitudes.

E Jawbone results

To probe the validity of our first-stage ATUS results, we evaluate the relationship between average sunset time and sleep using data collected by Jawbone, a manufacturer of wearable health and fitness trackers. Jawbone's UP wristbands track user sleep onset and duration using an accelerometer and, for some models, sensors for heart rate, respiration rate, body temperature, and galvanic skin response. In a blog post, Jawbone disclosed county-average bedtime, wake time, and sleep duration for all counties in the United States (Nolan, 2014). These data derive from more than 1 million underlying users, who chose to make their data available to the firm. Thus there are two margins of selection, first into Jawbone ownership and second into data sharing. Jawbone performed some interpolation for less populous counties, but does not provide a detailed description of the interpolation procedure, so we cannot evaluate its effects on the results below. Bearing these caveats in mind, the Jawbone data nonetheless provide a useful independent check on our ATUS results.

Table 29: Jawbone first stage estimates

	Bedtime	Wake time	Sleep
Avg. sunset time	0.31***	0.27***	-0.30***
	(0.016)	(0.016)	(0.090)
Observations	2410	2410	2410
Adjusted R^2	0.142	0.108	0.005

Notes: Heteroskedasticity-robust standard errors are reported in parentheses. Significance indicated by: *** p<0.01, ** p<0.05, * p<0.1. Data are from Jawbone (Nolan, 2014). Sunset time is the long-run county average.

We estimate effects of average sunset time on bedtime, wake time, and sleep duration in simple bivariate models. As in ATUS, a later sunset time delays bedtime: an average sunset one hour later postpones bedtime by .31 hours. This effect is not fully offset by a later wake time (postponed by .27 hours), resulting in lower sleep duration (.30 hours per week). Because the Jawbone data are not a random sample, these magnitudes are not directly comparable to those from ATUS, however the similar signs and relative magnitudes suggest our primary results are not driven primarily by reporting biases in the ATUS time diaries. Because Nolan (2014) does not discuss the within-location distribution of observations over the year, it is possible these estimates capture an unknown average of short- and long-run effects.