DISCUSSION PAPER SERIES

IZA DP No. 5306

How to Think About Time-Use Data: What Inferences Can We Make About Long- and Short-Run Time Use from Time Diaries?

Harley Frazis
Jay Stewart

November 2010

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

How to Think About Time-Use Data: What Inferences Can We Make About Long- and Short-Run Time Use from Time Diaries?

Harley Frazis
U.S. Bureau of Labor Statistics

Jay Stewart
U.S. Bureau of Labor Statistics and IZA

Discussion Paper No. 5306
November 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 5306
November 2010

ABSTRACT

How to Think About Time-Use Data: What Inferences Can We Make About Long- and Short-Run Time Use from Time Diaries?*

Time-use researchers are typically interested in the time use of individuals, but time use data are samples of person-days. Given day-to-day variation in how people spend their time, this distinction is analytically important. We examine the conditions necessary to make inferences about the time use of individuals from a sample of person-days. We also discuss whether and how surveys with multiple household members or multiple days are an improvement over single-diary surveys.

JEL Classification: C81, D13, C83

Keywords: time use, estimation, survey methods

Corresponding author:

Jay Stewart
U.S. Bureau of Labor Statistics
2 Massachusetts Avenue, NE
Washington, D.C. 20212
USA
E-mail: Stewart.Jay@bls.gov

* Any views expressed in this article are ours and do not necessarily reflect those of the Bureau of Labor Statistics. We thank participants at the 2010 meetings of the International Association for Time Use Research for their comments.

Introduction

In recent decades, there has been an increased interest in using time diaries to analyze how people spend their time. The distinctive feature of time-diary data is the short reference period—usually one or two 24-hour periods—compared to reference periods of a week or more for most household surveys. The shorter recall period results in more accurate data, and because time diaries ask respondents to list all of their activities rather than asking about a few select activities, they are also relatively free of social desirability bias. These data would seem to be ideal for answering a number of questions about the time spent in many activities, including childcare, household work, and job search, that are of interest to policymakers. However, policymakers are typically interested in long-run time use—that is, the amount of time that individuals spend in an activity over the course of a month or a year. Thus a drawback to time diaries over retrospective questions is that the shorter reference period combined with the considerable day-to-day variation in the time spent in different activities means that an individual's time diary does not necessarily reflect his or her long-run time use. Put another way, there is a mismatch between the reference period of the data (one or two days) and the time period that is typically of interest to the researcher. This feature of time-diary data has important implications for its analysis.

To date, many studies that analyze time-diary data have failed to recognize the importance of the reference-period mismatch. As a result, many recent papers have analyzed time-diary data inappropriately. The published literature contains examples both of researchers estimating statistics that, if properly interpreted, have little analytic interest and of researchers failing to estimate easily-calculated statistics that are of interest.

The purpose of this paper is to provide guidance to time-use researchers on what can and cannot be estimated from time-diary data. For the most part, we focus on data collected for a single day from a single member of a household as is done in the American Time Use Survey (ATUS).[1] This framework brings out the statistical issues in sharpest relief. We begin by examining the most common type of analysis, where time spent in an activity is the dependent variable. We then turn to analyses that are becoming more common in which time use is an explanatory variable. Finally, we consider issues related to survey design, such as what researchers can learn from time-use surveys that collect data from multiple household members or for multiple days compared with surveys that do not collect this information.

Time-Use as a Dependent Variable

The time-use of a population of individuals can be thought of as a collection of person-days. For example, in 2008 the civilian non-institutional population of the U.S. ages 15 and older was 238 million, which translates into 87 billion person days. It is straightforward to use data collected for a single day per respondent to estimate summary measures over the population of person-days. Taking means and medians as examples of summary measures and using 2008 ATUS data, we estimate that the mean time spent doing household activities was 106 minutes per person-day and the median time was 60 minutes. However, researchers are usually interested in time use over a period of time that is long enough to reflect the long-run decision making of individuals and households—that is, researchers are typically interested in person-months or person-years, rather than person-days. Given this mismatch between the period of interest and

[1] To briefly describe the ATUS: One respondent is chosen from each sample household. The respondent is asked to take the interviewer through his or her day for one 24-hour period (beginning at 4 AM in the ATUS). The respondent describes each activity, which the interviewer records. For each episode the duration of the activity is also recorded. The activity descriptions are then coded.

the reference period of the data, it is important to know what summary measures can be estimated from person-day data that allow us to make inferences about time use over a person-year.

To illustrate, consider the two measures of central tendency mentioned above—the mean and the median. Estimates of long-term mean time use follow immediately from the short-term estimates. Estimated mean annual time spent by individuals in household activities in 2008 is clearly 366 x 106 minutes per year. In contrast, it is impossible to estimate median annual time use from daily data because there is no necessary relationship between the percentiles of the distribution of daily time use and the distribution over longer periods. To illustrate, suppose that every person in the population does a certain activity only once per week. The median time spent in the activity is zero for the population of person days even though every person in the population spends time doing the activity each week.

More generally, we seek to provide guidelines as to when it is possible to make inferences about the time use of individuals from a sample of person days. To generalize from the above examples, let t_{id}^{a} denote total time spent in an activity of interest a $(a \in A)$ by person i (i = 1,…,I) on day d (d = 1,…,D). Let t_{id}^{A} denote the vector whose elements are t_{id}^{a} for all $a \in A$, T_{ID}^{A} denote the set of t_{id}^{A} for all activities for all person days in the population, and let $T_i^A \subset T_{ID}^{A}$ denote the set of days for person i, where $\bigcup_{i=1}^{I} T_{id}^{A} = T_{ID}^{A}$. The population T_{ID}^{A} describes the sampling frame of the ATUS and many other time use surveys, which typically collect only one or two diaries per person. As our interest will be in one or two activities at a time, for the most part we confine our discussion to the population T_{ID}^{a}—the time spent engaged in activity a—and omit the superscript except where necessary for clarity. Finally, let $g(T)$ be a statistic of interest to the researcher that is defined over a set of days, T.

We would like to know the conditions under which the statistic $g(T_{ID})$ calculated over a sample of person-days can be used to make inferences about the long-run time use of individuals. Let $\overline{T}_{ID} = \{\overline{T}_1, \overline{T}_2, \cdots, \overline{T}_I\}$ denote the long-run time use by the I individuals in the population, where $\overline{T}_i = \{\bar{t}_i, \bar{t}_i, \cdots, \bar{t}_i\}$ and $\bar{t}_i = \frac{1}{D}\sum_{d=1}^{D} t_{id}$. Put another way, $\overline{T}_i$ is an artificial construct that represents the time spent in the activity by person i that would result if there were no day-to-day variation and he or she spent the same amount of time (his or her long-run average) in the activity every day. We can estimate average long-run time use from a sample of person days if:

(1) $$g(T_{ID}) = g(\overline{T}_{ID}).$$

This condition simply states that it is possible to make inferences about individuals' long-run time use from a sample of person days if the day-to-day variation does not affect the value of the statistic.

A couple of examples will illustrate the point. The mean clearly satisfies this condition:

$$g(T_{ID}) = \frac{1}{I \times D}\sum_{i=1}^{I}\sum_{d=1}^{D} t_{id} = \frac{1}{I}\sum_{i=1}^{I}\left(\frac{1}{D}\sum_{d=1}^{D} t_{id}\right) = \frac{1}{I}\sum_{i=1}^{I} \bar{t}_i = g(\overline{T}_{ID}).$$

The earlier example of the once-per-week activity makes it clear that the median does not satisfy this property. The variance is another example of a function that does not satisfy this property. To see this, decompose t_{id} as follows:

(2) $$t_{id} = m_i + e_{id}$$

where $m_i = E_d(t_{id})$ is the long-run average deviation of individual i's time spent in the activity of interest and e_{id} is the deviation from m_i for a given day. Note that m_i and e_{id} are orthogonal, that

(2) is true by construction, and that we have made no assumptions about the time-series structure of the disturbances. Letting g(T) = Var(T) and noting that e_{id} = 0 for each $\bar{T}_i$, then:

$$g(T_{ID}) = Var(m_i + e_{id}) = Var(m_i) + Var(e_{id}) \neq Var(m_i) = g(\bar{T}_{ID}),$$

which means that we cannot estimate the variance of the long-run time use of individuals from a population of person-days.

However, even though point estimates of the variance of long-term time use are not identified, it is possible to bound the variance. The addition of covariates X will help establish the lower bound, so we modify (2) to include covariates:

(2′) $$t_{id} = L(X_i) + m_i + e_{id},$$

where $L(X)$ is the long-run mean of t_{id} conditional on X, m_i is redefined to equal the long-run average deviation of individual i's time use from $L(X)$, and e_{id} is the deviation from ($L(X) + m_i$) for a given day. Note that $L(X)$, m_i, and e_{id} are all mutually orthogonal. It is clear from equation (2′) that the residual components m_i and e_{it} are not identified in cross-sectional data. Letting $Var(\bar{t}_i)$ denote the variance of long-term time use across individuals, consider two polar cases. If there is no day-to-day variation, then e_{id} = 0 for all members of the population, $Var(\bar{t}_i)$ = $Var(L(X_i)) + Var(m_i)$, and all variation in the data reflects permanent differences between individuals. If on the other hand m_i = 0 for all members of the population, then $Var(\bar{T}_{ID})$ = $Var(L(X_i))$ and all residual variation is day-to-day variation, which vanishes as averages are taken in the long run. This implies:

(3) $$Var(L(X_i)) \leq Var(\bar{t}_i) \leq Var(t_{id}).$$

The lower bound can be estimated by computing the variance of predicted time use from a regression of the t_{id} on the X_i. An illustration of this bounding, and an extension to multi-person households, can be found in Frazis and Stewart (2009) in the context of incorporating household production into income inequality measures.

Regression Functional Forms and Reporting

The ability to estimate means of long-run time use from a sample of daily observations implies that researchers can estimate the effect of covariates on time use using regression analysis. That said, the distinctive features of time-use data have implications for the functional form of regression functions. For some activities it makes sense to assume that $L(X)$ in (2′) is strictly greater than zero. This would be the case for activities such as childcare done by parents and time spent working by the employed. Now let $S_d(X)$ denote the corresponding expectation for a given day *d*. While it is obvious that the functional form of the $S_d(\cdot)$ implies a functional form for $L(\cdot)$, the reverse is also true. If, as is commonly assumed, $L(X)$ is linear, then the $S_d(X)$ function will also almost always be linear.[2]

This implies that the long-term expectation, $L(X)$, can be consistently estimated using OLS on a sample of person-day data, t_{id}.[3] This is true despite the possibility that a large proportion of t_{id} observations have a zero value. In contrast, and somewhat ironically, estimation techniques for limited dependent variables that assume a non-linear functional form, such as the Tobit, will be inconsistent. For a comparison of OLS and Tobit, see Stewart (2009).

Another specification issue is whether to estimate means and regression coefficients separately for weekends and weekdays. This is sometimes done on the grounds that these days are very different from each other. Although separate estimation can be justified, it raises the issue of how to report results. It is not uncommon for researchers to report means and

[2] One can construct cases where the non-linearities of different S_d functions exactly cancel each other out. However, this will only happen coincidentally.
[3] OLS, however, will be inefficient due to the non-normality of the residuals.

regression coefficients separately for weekends and weekdays with no mention of overall means or the effect of covariates on overall time use (for example, Kalenkoski, Ribar, and Stratton, 2007; Kimmel and Connelly, 2007; Connelly and Kimmel, 2009; and Vernon, 2010). Yet there is no reason not to report overall effects. Assuming (for convenience) time use is homogenous within weekdays and weekend days, it is a straightforward consequence of the linearity of means that weekly mean time use can be estimated as $w(t) \equiv 5S_D(t_{iD}) + 2S_E(t_{iE})$, where *S* denotes the daily mean, *D* denotes weekday and *E* denotes weekend. More importantly, it is not clear how one should interpret the separate coefficients. The effect of a covariate on, say, weekend time use can be thought of as the combined effect of the effects of the covariate on overall time use and on the weekday-weekend distribution of time spent in the activity. Overall time-use is relevant for most policy-related questions, and substitution effects can shed light on how people rearrange their schedules in response to outside influences. But it is not clear what questions the combined effects address.

Short-run Time Use

Focusing on short-run time-use is appropriate when the researcher is interested in the shifting of activities between days. For example, Connolly (2008) uses ATUS data to examine how workers shift work and leisure activities in response to changes in the weather.[4] She finds that people shift work activities from sunny days to rainy days, while leisure activities are shifted in the opposite direction. To illustrate the use of single-day data in analyzing substitution, consider the following regression equation:

---

[4] Other researchers (Hamermesh, et al, 2008 and Stewart, 2010) have examined the timing of activities within a day. Hamermesh, et al, examined how light cues, television schedules, and the synchronization of work activities across time zones affect the timing of work and leisure activities. Stewart examines how work schedules affect when during the day that mothers spend time with their children and how mothers adjust their schedules to spend time with their children at more-preferred times of day.

(4) $t_{id} = X_i\beta + (X_{id} - X_i)\gamma + m_i + e_{id}$

where X_i denotes the mean value of covariate X across time and X_{id} denotes the value on day d. The person fixed-effect m_i may be correlated with X_i –in the Connolly (2008) example, persons who spend above-average time in leisure may migrate to low-precipitation areas—but the parameter of interest γ will be consistently estimated if $(X_{id} - X_i)$ is uncorrelated with e_{id}.

Because Connolly is interested in the short-run responsiveness of time-use to shocks in a covariate, there is no mismatch between the reference period and the period of interest. Accordingly, summary measures that do not satisfy equation (1), such as coefficients from median regressions or the probability that $t_{id} > 0$, may be of analytic interest.

Time Use as an Independent Variable

Several recent studies have included time use as an explanatory variable (Pinkston and Stewart 2009, Christian 2009, Hersch 2009, Hamermesh 2010). This use of the data opens up new avenues of research, but there are potential problems with using time use as an independent variable.

Most fundamentally, where the dependent variable is a long-term outcome such as obesity (Pinkston and Stewart 2009) or wages (Hersch 2009), OLS estimates of the coefficients on short-term time-use cannot be interpreted causally. Consider the equation

(5) $Y_i = X_i\beta + t_{id}\gamma + u,$

where Y_i is a long-term outcome (such as obesity) for person i, X_i is a vector of covariates, t_{id} is time-use in an activity of interest (such as exercise) by person i on diary-day d, and u is a residual. By itself we would expect a single day's exercise (t_{id}) to have a negligible effect on obesity (Y_i).

Even though daily time use has little effect on Y_i, it can be considered a proxy for the true variable of interest, t_i, the long-term average time-use of person i. As implied by the decomposition in equation (2), t_{id} will differ from m_i by the residual e_{id} which is uncorrelated with m_i. Thus use of t_{id} rather than m_i is an example of classical measurement error. If there is a single time-use variable in the regression, we have the familiar errors-in-variables result that the coefficient on t_{id} will be attenuated toward zero.

With multiple time-use variables on the RHS, in general nothing can be said about the direction of the bias. Klepper and Leamer (1984) show that in the case where the measurement errors of the different variables are uncorrelated, the true coefficients can be bounded only if (5) and the reverse regressions formed by regressing one of the t_{id} variables on Y and the other t_{id} variables all imply the same signs on the γ vector. The bounds are the regression and reverse-regression coefficients. With time-use variables the situation is even less favorable, as the e_{id} variables for different activities will usually be correlated with each other. In this case, the bounds will be looser than those of Klepper and Leamer. It is straightforward to construct examples where the probability limit of an estimated coefficient will have an incorrect sign even when the regression and all reverse regressions imply the same sign on all time-use variables.

We suggest two solutions. The first is to group the data into cells based on observed characteristics (for example, age, education, and sex) and estimate (5) using cell means rather than individual data. It is important to make sure that cells are sufficiently large and that the day-of-week representation is correct. Faberman (2010) provides an example of this approach.

The second is to use instrumental variables (IV). IV has the added advantage that it also corrects for any reverse causality between Y and t_i. It is important to note that the instruments should predict long-term time-use rather than short-term. For example, Pinkston and Stewart

(2009) use climate information (average temperature and rainfall) as instruments for time spent watching TV in their analysis of obesity, rather than using the weather on the diary day.

Some interesting issues arise when time spent in one activity is an explanatory variable for time spent in another activity (for example, Christian 2009). Consider the equation

$$(6) \qquad t_{id}^{a} = \alpha + \beta t_{id}^{b} + u_{id},$$

where time spent in activity *a* (exercising in Christian's example) is a function of the time spent in activity *b* (commuting in Christian's example) and, for simplicity, no other covariates. From (2), $t_{id}^{a} = m_{i}^{a} + e_{id}^{a}$ and $t_{id}^{b} = m_{i}^{b} + e_{id}^{b}$, so defining β as the population linear regression coefficient we have

$$(7) \quad \beta = \frac{Cov(t_{id}^{a}, t_{id}^{b})}{Var(t_{id}^{b})} = \frac{Cov(m_{i}^{a}, m_{i}^{b}) + Cov(e_{id}^{a}, e_{id}^{b})}{Var(m_{i}^{b}) + Var(e_{id}^{b})} = \frac{Var(m_{i}^{b})\beta_{m} + Var(e_{id}^{b})\beta_{e}}{Var(m_{i}^{b}) + Var(e_{id}^{b})}$$

where $\beta_{m} \equiv \frac{Cov(m_{i}^{a}, m_{i}^{b})}{Var(m_{i}^{b})}$ and $\beta_{e} \equiv \frac{Cov(e_{id}^{a}, e_{id}^{b})}{Var(e_{id}^{b})}$ are the coefficients from regressions of m_{i}^{a} on m_{i}^{b} and of e_{id}^{a} on e_{id}^{b}, respectively. Thus the coefficient estimated by OLS using ATUS-type time-diary data is a weighted average of the coefficient showing the linear relationship of long run variation between the two activities and the coefficient showing the relationship of day-to-day variation. The different coefficients answer different questions—β_m shows the extent to which people who spend more total time commuting spend more or less total time exercising, while β_e shows the extent to which people spend time exercising and commuting on the same days conditional on long-run total time in the activities. Either coefficient may be of interest, but it is hard to imagine what question the weighted average (with the value of the weights unknown) would answer.

Instrumental variables may be used to estimate either β_m or β_e. Christian (2009) uses daily traffic accidents, which are presumably dominated by short-term variation, as an instrument for short-run variation in commuting, thus estimating β_e. Longer-run exogenous factors could be used as instruments for β_m.

Survey Design Issues

As mentioned in the introduction, time-use surveys outside the United States typically interview more than one household member, interview across multiple days, or both. Surveys in South Africa and Thailand interview multiple household members for a single day. The UK, German and Australian time-use surveys collect data on multiple household members for multiple days. In contrast, the ATUS collects data on one household member for one day, which Connelly and Kimmel (2009) describe as a "near-fatal flaw."

Multiple Household Members

At first glance, the availability of data on spouses would seem to allow researchers to estimate how the spouse's time use affects own time use (this is the goal of Connelly and Kimmel 2009, for example). However, we face essentially the same issue as we do when an individual's own time use is on both sides of the regression equation. Consider the following modification of (6):

$$(6') \qquad t_{id}^{H} = \alpha + \beta t_{id}^{W} + u_{id},$$

where here i denotes a household rather than an individual, the H superscript denotes the husband and W denotes the wife. An OLS regression of (6′) will estimate a weighted average of β_m, which in this context is the long-run association between husband's and wife's time in a given activity, and β_e, which measures the association between husband's and wife's deviations from their long-run means for that activity.[5]

As above, IV estimation may be used to estimate either β_m or β_e. But note that it is not necessary to have data on both the husband and the wife to apply IV. Predicted values (either of long-run means or of deviations from means) can be generated from a sample of husbands (wives) and applied in a 2SLS regression of wives' (husbands') time-use. Thus there is little that can be learned from time diary data surveys that collect diaries from multiple household members compared to surveys that collect diaries from only one person per household.[6]

Multiple Days

Many time-use surveys collect data on multiple days from either a single respondent or multiple household respondents. The most common arrangement is for respondents to be interviewed on two days, either consecutive days or one weekday and one weekend day.[7] The 2006 Netherlands survey collected data for seven consecutive days, but this is the only example we are aware of a large-scale survey collecting more than two days per respondent. High response burden and the resulting high non-response rate are obvious deterrents to this type of endeavor.

Given that long-term means and total time-use are already identified in single-diary datasets, surveys that collect multiple days add nothing along this dimension for summary measures of long-run time-use that satisfy equation (1). Further, the periods collected are also

[5] A brief examination of the literature did not reveal many cases of time-use data on couples being used to estimate (6'). Hook (2004) is one example.
[6] Surveys with diaries from multiple household members can answer questions about how parents schedule work and family time within a day.
[7] Australia, Japan, and Norway interview for two consecutive days. Sweden, Turkey, Latvia, Lithuania, Poland, and Denmark interview for one weekday and one weekend day.

typically too short to be considered measures of long-run time use from the viewpoint of estimating statistics that do not satisfy equation (1).

One rationale offered for collecting multiple days is to estimate variability in activities over time. The guidelines to the Harmonised European Time Use Surveys (Eurostat 2009) recommend collecting multiple days and state "Using only one diary day will also be acceptable, but in that case it is impossible to get any idea of intra-personal variation." However, we cannot learn much about intra-personal variability given the design of most multiple-day time-use surveys.

Estimation of parameters describing intra-personal variation in time-use with data collected for two days is only possible if time use can reasonably be assumed to be uncorrelated between the days sampled. Extrapolating equation (2) to two days, d and d', note that the variance of t_{id} is $Var(m_i) + Var(e_{id})$ and that the covariance of time-use between days d and d' is $Var(m_i) + Cov(e_{id}, e_{id'})$. In general, $Cov(e_{id}, e_{id'})$ is unknown unless we can safely assume that it is zero. We can estimate $Var(t_{id})$, $Var(t_{id'})$, and $Cov(t_{id},t_{id'})$ from a two-day-per-person sample, but the underlying parameters $Var(m_i)$, $Var(e_{id})$, $\mathrm{Var}(e_{id'})$, and $Cov(e_{id},e_{id'})$ are not identified.

For most existing surveys that sample two days, independence is unlikely for two reasons. First, the days sampled are typically close to each other in time, which may result in either positive or negative covariances depending on the activity. Busy periods at work that require longer hours may extend over several days and cause positive covariances for market work. On the other hand, household tasks may only be necessary periodically, so if observed on one day they would be unlikely to be observed on surrounding days, with a consequent negative covariance.

The second, more subtle, reason is that some surveys intentionally sample two different days of the week, for example one weekday and one weekend-day. With this type of sampling, even for days that are widely separated from each other, time-use may be correlated if persons have regular weekly schedules that vary across the population.

We first examine what can be estimated if activities on the sampled days are uncorrelated. As before, we decompose time use on a diary day into its components. Specifically, we redefine *d* to be a day of the week (*d*=1,…,7), let *w* denote a specific week, and rewrite the decomposition in equation (2) as follows (again omitting covariates for simplicity):

$$(2'') \qquad \begin{aligned} t_{idw} &= m_i + e_{id} \\ &= M + \dot{m}_i + \dot{N}_d + \dot{n}_{id} + e_{idw} \end{aligned}$$

where *M* is the population mean time spent in the activity; $\dot{m}_i (= m_i - M)$ is the deviation of individual *i*'s time use from the population mean, *M*; $\dot{N}_d$ is the population mean deviation from *M* on a particular day of the week *d*; and $\dot{n}_{id}$ is the average deviation from the person *i* mean time m_i on day of the week *d* net of the population day effect ($\dot{n}_{id} = \bar{t}_i - M - \dot{m}_i - \dot{N}_d$), and e_{idw} is the residual from the day-of-the-week mean. Note that by construction, $\sum_{i=1}^{I} \dot{m}_i = 0$, $\sum_{d=1}^{7} \dot{N}_d = 0$, and $\sum_{d=1}^{7} \dot{n}_d = 0$.

It is plausible to assume that $Cov(e_{idw}, e_{id'w'}) = 0$ for two weeks, *w* and *w′*, that are sufficiently widely separated in time. Thus the covariance between the widely separated days, *d* and *d′*, is: $Cov(t_{idw}, t_{id'w'}) = E((\dot{m}_i + \dot{n}_{id} + e_{idw})(\dot{m}_i + \dot{n}_{id'} + e_{id'w'})) = E(\dot{m}_i^2 + \dot{n}_{id}\dot{n}_{id'})$ $= Var(\dot{m}_i) + Cov(\dot{n}_{id}, \dot{n}_{id'})$.

Consider sampling two days at random from two weeks that are widely-separated in time each with probability 1/7. The average covariance between pairs of days is

$$E_{d,d'}E(\dot{n}_{id}\dot{n}_{id'}) = E\left(\left(\frac{1}{7}\dot{n}_{i1} + \frac{1}{7}\dot{n}_{i2} + \ldots + \frac{1}{7}\dot{n}_{i7}\right)\left(\frac{1}{7}\dot{n}_{i1} + \frac{1}{7}\dot{n}_{i2} + \ldots + \frac{1}{7}\dot{n}_{i7}\right)\right) = E\left(\frac{(\sum_{d=1}^{7}\dot{n}_{id})^2}{7^2}\right) = 0.$$

Thus for randomly-selected pairs of days in widely separated weeks, the covariance of time-use after subtracting day-of-week effects is: $Cov(t_{idw} - \dot{N}_d, t_{id'w'} - \dot{N}_{d'}) = Var(\dot{m}_i)$, which means that day-to-day variation in time use can be disentangled from long-run variation.

Now consider a sampling scheme that is similar to those used in some existing surveys where one weekend-day (*d*= 1,2) and one weekday (*d′*=3,…,7) are chosen at random. We have:

$$E_{d,d'}E(n_{id}n_{id'}) = E\left(\frac{n_{i1}\sum_{d'=3}^{7} n_{id'} + n_{i2}\sum_{d'=3}^{7} n_{id'}}{2\times 5}\right) \neq 0,$$

so long-run variation in time-use is no longer identified.

If the survey design is such that time use on the days sampled is uncorrelated between days, one can identify both β_m and β_e in equation (7). In the regression

$$(8) \qquad t_{id}^{A} - t_{id'}^{A} = \alpha_{\delta} + \beta_{\delta}(t_{id}^{B} - t_{id'}^{B}) + v_{idd'},$$

the coefficient β_δ will equal β_e in this case. As noted above, the $Var(\dot{m}_i)$ and $Var(e_{id})$ parameters (or rather $Var(e_{idw} + \dot{n}_{id})$, in the notation of (2″)) can also be identified in this case, so β_m can be identified by estimating (5) and substituting β_e, $Var(\dot{m}_i)$ and $Var(e_{id})$ into (6).

If time use on the days sampled is correlated, the regression in equation (8) will not identify β_e:

$$\beta_\delta = \frac{Cov(t_{id}^A - t_{id'}^A, t_{id}^B - t_{id'}^B)}{Var(t_{id}^B - t_{id'}^B)} = \frac{Cov(e_{id}^A - e_{id'}^A, e_{id}^B - e_{id'}^B)}{Var(e_{id}^B - e_{id'}^B)} = \frac{Cov(e_{id}^A, e_{id}^B) - Cov(e_{id}^A, e_{id'}^B)}{Var(e_{id}^B) - Cov(e_{id}^B, e_{id'}^B)},$$

(using the decomposition (2) instead of (2") for simplicity). The non-zero cross-day covariances $Cov(e_{id}^A, e_{id'}^B)$ and $Cov(e_{id}^B, e_{id'}^B)$ prevent identification of β_m and β_e.

Even with time-use correlated between days, collecting multiple days will tighten the upper bound of the long-run variance of t_i described in (3). This bound will be tighter (on a per-day basis) with more days surveyed per respondent.

Multiple Household Members on Multiple Days

Several time-use surveys obtain data from multiple household members on multiple days. The issues arising from this design are similar to those in previous sections. The short- and long-run regression coefficients β_e and β_m in (6′) can be identified by the equivalent of (8) if time-use on the days sampled is uncorrelated between days, but not otherwise.

How Prevalent are these Issues in the Literature?

To answer this question, we conducted a brief review of the literature to see how often these issues arise in published papers. To keep the review to a manageable size, we focused on 26 papers that use ATUS data, were written by researchers outside of BLS, and were published in economics journals or conference volumes.[8] We examined three issues. First, when it was long-run time use that was of interest, did the authors report any statistics that did not satisfy equation (1)? Second, did the authors only report time use for weekends and weekdays even though they were interested in long-run time use? And third, in cases where time-use was an

[8] The papers are Aguiar and Hurst (2007a,b), Burda and Hamermesh (2010), Burda, Hamermesh, and Weil (2008), Connelly and Kimmel (2009), Connolly (2008), Davis and You (2010), Folbre and Yoon (2007, 2008), Guryan, Hurst, and Kearny (2008), Hamermesh (2007, 2008a, 2008b, 2010) , Hamermesh, Myers and Pocock (2008), Hersch (2009), Kalenkoski, Ribar, and Stratton (2007), Kimmel and Connelly (2007), Krueger (2007), Landefeld, Fraumeni, and Vojtech,(2009), Meyer and Sullivan (2008), Price (2008), Vernon (2009), Ward-Batts (2008), Zick and Bryant (2008), Zick, Bryant, and Srisukhumbowornchai (2008).

independent variable, did the authors use aggregation or instrumental variables to mitigate the mismeasurement of long-run time-use?

The results confirm that these three issues frequently arise in the literature. Out of the 26 papers we reviewed, 15 had one or more issues. Of the 22 papers concerned with long-run time-use as a dependent variable, 6 reported statistics that do not satisfy (1) (in every case, the authors reported percentiles). Four papers out of the same 22 did not report totals where relevant. Of the 7 papers where time-use was an independent variable, 5 did not use any methods to mitigate measurement error. In many cases the suspect analysis is only a minor part of the overall paper.[9] But it is clear that time-diary data are frequently used as if they were an accurate measure of the respondent's long-run time use.

Conclusion

Time diary data provide researchers with a detailed look at how people spend their time over a 24-hour period. However, the short reference period combined with the large amount of day-to-day variation in time use implies that any given time diary is a poor indicator of that individual's long-run time use. Although short-run time use is sometimes of interest, for most policy-related questions it is long-run time use that is relevant. This mismatch between the reference period of the data (a person-day) and the period of interest (a person-month or person-year) has important implications for the analysis of time diary data. These implications have not been fully appreciated by many researchers.

In general, when time use is the dependent variable, researchers can make inferences about the long-run time use of individuals from a sample of person-days only when the value of

[9] For example, in one section of a substantial paper Aguilar and Hurst (2007b) discuss changes in percentiles of leisure, but otherwise analyze means. Price (2008) uses median regressions as a robustness check for his OLS regressions.

the statistic is invariant to the amount of day-to-day variation in the data. For example, researchers can estimate mean (per person) time use for a population of individuals, but it is impossible to estimate how much time the median person spends in an activity.

Several recent studies have included time use as an explanatory variable in regression equations. Long-run outcomes such as obesity can only be affected by time-use over a long period of time. But, as we have seen, time use on a randomly chosen diary day is not a good indicator of long-run time use. This implies that time-diary data measure the true variable of analytic interest with considerable error even if the respondent has made no recall errors. Using instrumental variables (IV) or entering time use as an aggregate (for example, cell means) are two ways of dealing with this measurement error.

Lastly, our findings have implications for survey design. Time use researchers have long recommended that surveys collect time diaries from every person in the household and collect them for multiple days. Collecting time diaries from everyone in the household would seem to allow researchers to learn about the intra-household allocation of household work and leisure. But we have shown that users of single-day data from multiple household members face the problem of disentangling the day-to-day covariance of activities from the long-run covariance, which implies that there is little that one can learn by collecting diaries from everyone in the household that cannot be learned from single-person-per-household diaries. Collecting multiple diaries from respondents can be potentially valuable, but (given practical limits on the number of days sampled) only if the days are sampled so as to ensure that the activities on those days are independent. Without independence, between-day covariance makes it impossible to identify the within- and between-person variance.

The increasing availability of time-diary data has allowed analysts to research a wide variety of questions whose investigation was previously limited by data that was either nonexistent or subject to substantial recall bias. We hope that future researchers will take account of the short-term nature of the data in their research designs while continuing to take advantage of the greater accuracy that time diaries offer.

References

Aguiar, Mark and Erik Hurst (2007a) "Life-Cycle Prices and Production" *American Economic Review* 97(5), pp. 1533-1559.

Aguiar, Mark and Erik Hurst (2007b) "Measuring Trends in Leisure: The allocation of Time over Five Decades." *Quarterly Journal of Economics* 122(3), pp.969-1006.

Burda, Michael C. and Daniel S. Hamermesh (2010) "Unemployment, market work and household production." *Economics Letters* 107, pp. 131-133.

Burda, Michael C., Daniel S. Hamermesh and Phillipe Weil (2008) "The Distribution of Total Work in the EU and US." In *Working Hours and Job Sharing in the EU and USA: Are Europeans Lazy? Or Americans Crazy?* Tito Boeri, Michael C. Burda, and Francis Kramarz, ed. New York: Oxford University Press, pp 13-91.

Christian, Thomas J. (2009) "Opportunity Costs Surrounding Exercise and Dietary Behaviors: Quantifying Trade-offs Between Commuting Time and Health-Related Activities." Unpublished working paper.

Connelly, Rachel and Jean Kimmel (2009) "Spousal influences on parents' non-market time choices." *Review of Economics of the Household* 7(4), pp. 361-394.

Connolly, Marie (2008) "Here Comes the Rain Again: Weather and the Intertemporal Substitution of Leisure." *Journal of Labor Economics* 26(1), pp. 73-100.

Davis, George C., and Wen You (2010) "The time cost of food at home: general and food stamp participant profiles" *Applied Economics* 42(20), pp. 2537-2552.

Eurostat (2009) *Harmonised European time use surveys 2008 guidelines*. Eurostat Methodologies and Working Papers. Luxembourg: Office for Official Publications of the European Communities, p. 9.

Faberman, R. Jason (2010) "Revisiting The Role Of Home Production In Life-Cycle Labor Supply." Working Paper No. 10-3, Research Department, Federal Reserve Bank of Philadelphia.

Folbre, Nancy and Jayoung Yoon (2007) "What is Child Care? Lessons from time-use surveys of Major English-speaking Countries." *Review of Economics of the Household* 5(3), pp. 223-248.

Folbre, Nancy and Jayoung Yoon (2008) "The Value of Unpaid Child Care in the United States in 2003," in *How Do We Spend Our Time? Evidence from the American Time Use Survey*, Jean Kimmel, ed. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, pp. 31-56.

Frazis, Harley and Jay Stewart (2009) "How Does Household Production Affect Measured Income Inequality?" Forthcoming in *Journal of Population Economics*.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney (2008) "Parental Education and Parental Time with Children," *Journal of Economic Perspectives* 22(3), pp. 23-46.

Hamermesh, Daniel S. (2007) "Time to Eat: Household Production Under Increasing Income Inequality." *Journal of Agricultural Economics* 89(4), pp. 852-863.

Hamermesh, Daniel S. (2008a) “Direct estimates of household production.” *Economics Letters* 98, pp. 31-34.

Hamermesh, Daniel S. (2008b) “The Time of Our Lives.” in *How Do We Spend Our Time? Evidence from the American Time Use Survey*, Jean Kimmel, ed. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, pp. 11-30.

Hamermesh, Daniel S. (2010) “Incentives, time use and BMI: The roles of eating, grazing and goods.” *Economics and Human Biology* 8, pp. 2-15.

Hamermesh, Daniel S., Caitlin Knowles Myers, Mark L. Pocock (2008) “Cues for Timing and Coordination: Latitude, Letterman, and Longitude.” *Journal of Labor Economics* 26(2), pp. 223-246.

Hersch, Joni (2009) “Home production and wages: evidence from the American Time Use Survey.” *Review of Economics of the Household* 7, pp. 159-178.

Hook, Jennifer (2004) “Reconsidering the Division of Household Labor: Incorporating Volunteer Work and Informal Support.” *Journal of Marriage and the Family* 66(1), pp. 101-117.

Kalenkoski, Charlene, David Ribar, and Leslie Stratton (2007) “The Effect of Family Structure on Parents' Child Care Time in the US and the UK.” *Review of Economics of the Household* 5(4), pp. 353-384.

Kimmel, Jean and Rachel Connelly (2007) “Mothers’ Time Choices: Caregiving, Leisure, Home Production, and Paid Work.” *Journal of Human Resources* 42(3), pp. 643-681.

Klepper, Steven and Edward E. Leamer, Edward E. (1984) "Consistent Sets of Estimates for Regressions with Errors in All Variables." *Econometrica*. 52(1), pp. 163-183.

Krueger, Alan B. (2007) “Are We Having More Fun Yet? Categorizing and Evaluating Changes in Time Allocation.” *Brookings Papers on Economic Activity* 2007(2), pp. 193-215.

Landefeld, J. Steven, Barbara M. Fraumeni, and Cindy M. Vojtech (2009) “Accounting for Household Production: A Prototype Satellite Account Using the American Time Use Survey”. *Review of Income and Wealth* 55(2), pp. 205-225.

Meyer, Bruce and James X. Sullivan (2008) “Changes in the Consumption, Income, and Well-Being of Single Mother Headed Families.” *American Economic Review* 98(5), pp. 2221-2241.

Pinkston, Joshua, and Jay C. Stewart (2009) “How Does Time Use Affect the Probability of Becoming Obese?” Unpublished manuscript.

Price, Joseph (2008) “Parent-Child Quality Time: Does Birth Order Matter?” *Journal of Human Resources* 43(1), pp. 240-265.

Stewart, Jay C. (2009) “Tobit or not Tobit?” BLS Working Paper No. 432.

Stewart, Jay (2010) "The Timing of Maternal Work and Time with Children," *Industrial & Labor Relations Review*. 64(1), pp. 181-200

Vernon, Victoria (2010) “Marriage: for love, for money ... and for time?” *Review of Economics of the Household*. Forthcoming.

Ward-Batts, Jennifer (2008) "Household Production, Consumption, and Retirement," in *How Do We Spend Our Time? Evidence from the American Time Use Survey*, Jean Kimmel, ed. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, pp. 81-108.

Zick, Cathleen D. and W. Keith Bryant (2008) "Does Housework Continue to Narrow the Income Gap?" in *How Do We Spend Our Time? Evidence from the American Time Use Survey*, Jean Kimmel, ed. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, pp. 57-80.

Zick, Cathleen D., W. Keith Bryant, and Sivithee Srisukhumbowornchai (2008) "Does housework matter anyore? The shifting impact of housework on economic inequality." *Review of Economics of the Household* 6(1), pp. 1-28.