# What Time Use Surveys Can (And Cannot) Tell Us About Labor Supply

Cheng Chou*
School of Business
University of Leicester

Ruoyao Shi†
Department of Economics
UC Riverside

This version: January 2021

**Summary**

The American Time Use Survey (ATUS) accurately measures hours worked on a single day. We propose several estimators of elasticities of weekly labor supply in a linear regression model, despite certain impossibility results due to the time specific feature of the ATUS. We recommend the *impute estimator*, a simple modification of the standard two stage least squares estimator, that imputes the dependent variable using daily subsamples, based on our careful investigation of asymptotic and finite sample properties of the estimators under the potential outcome framework. We apply the impute estimator to the ATUS and find substantially different elasticity estimates from the Current Population Survey, especially for married women.

**Keywords:** labor supply, time use surveys, impute estimator, relative asymptotic efficiency, survey methods

**Acknowledgments** Special thanks go to Robert W. Fairlie and Arthur Stone, without whose encouragement and insights there could not have been such a paper. Comments and suggestions from the co-editor Edward J. Vytlacil and three anonymous referees helped improve the paper significantly. The authors thank Jaap Abbring, Joseph Altonji, Moshe Buchinsky, Colin Cameron, Scott Carrell, Juan Moreno Cruz, Julie Berry Cullen, Carlos Dobkin, Daniel Gutknecht, Jinyong Hahn, Zhipeng Liao, Audrey Light, Jesse Matheson, Rosa Matzkin, Maurizio Mazzocco, Kathleen McGarry, Whitney Newey, Rodrigo Pinto, Geert Ridder, and David Rojo Arjona for helpful discussion. Participants of seminars at UCLA, UC Irvine, UC Riverside, University of Leicester, Tilburg University, Southern Methodist University, Canadian Economics Association 2019 Annual Meeting, Econometric Society 2019 Asian Meeting, and Midwest Econometrics Group 2019 Annual Meeting also helped to improve the paper. All remaining errors are the authors'.

*Email: cchou@le.ac.uk.

†Corresponding author. Address: 3136 Sproul Hall, 900 University Avenue, Riverside, CA 92521. Tel: +1-951-827-1494. Email: ruoyao.shi@ucr.edu. The authors do not have conflicts of interest to declare.

# 1 Introduction

Empirical studies of labor supply depend greatly upon data on how much time people spend working. Unfortunately, there is abundant evidence showing that weekly hours worked are poorly measured in frequently used survey data sets such as the Current Population Survey (CPS) and the Panel Study of Income Dynamics (PSID), and that the measurement error is nonclassical (e.g., Bound et al., 1989, 2001), which significantly biases the estimation of labor supply parameters (Barrett and Hamermesh, 2019). Aiming to measure how people allocate their time on market work and non-market activities more accurately, many countries have historical or ongoing time use surveys.[1] Time use surveys typically ask the respondents to record all their activities during a prescribed period in the format of a detailed diary, hence provide much more accurate measure of how individuals allocate their time in that period (e.g., Frazis and Stewart, 2012; Aguiar et al., 2017).[2]

The major conceptual issue in deploying time use surveys to measure labor supply is that time use surveys typically provide information about labor supply for only a few days of a week, but on the contrary the CPS concerns an entire week. For example, the American Time Use Survey (ATUS) records one single day for each respondent, while the Australian and the UK time use surveys currently record two days. To the best of our knowledge, the only exception is the Dutch Time Use Survey (DTUS),[3] in which respondents record their activities for seven consecutive days. If we are interested in weekly labor supply, then ideally we need to observe typical weekly hours worked. The single day observed in the ATUS, albeit randomly sampled, creates a missing data problem.[4] In this paper, we will use the term coined by Barrett and Hamermesh (2019) and refer to this problem as "time specificity". Frazis and Stewart (2012) were the first to investigate the consequences of time specificity problem. In particular, they gave a sufficient condition (eq. (1) of Frazis and Stewart, 2012) for weekly statistics about which inference could be made using time use survey data (e.g., mean of weekly hours).[5] They further pointed out that weekly regression

[1]These countries include Australia, Canada, China, Japan, New Zealand, Pakistan, Russia, the USA and most European countries.

[2]For a review of time use surveys used for studies in other subfields of economics, see Aguiar et al. (2012).

[3]In Dutch, it is called *Het Tijdsbestedingsonderzoek*, but in this paper we call it the DTUS for the consistency with the ATUS.

[4]The hours worked on the non-diary days are missing completely at random and follow the "file-matching" pattern (Little and Rubin, 2019).

[5]Frazis and Stewart (2012) also gave median and variance of weekly hours as counterexamples that do not satisfy

functions can be estimated using time use survey data since they are conditional mean functions.[6] To fix ideas, we will use the weekly labor supply as a leading example in the rest of this paper, and will focus on the ATUS, in which respondents record their time use on *one* randomly chosen day.

The main contributions of this paper are twofold. Methodologically, we propose several intuitive and easy-to-implement estimators that utilize the ATUS daily hours worked to estimate a weekly labor supply equation and carefully investigate their properties. We recommend what we name *impute estimator* based on its superior relative asymptotic efficiency and finite sample stability. In terms of empirical estimates of weekly elasticities of labor supply, we apply our impute estimator to a sample of American workers who participated in both the ATUS and the CPS, uncover multiple interesting empirical findings, and compare these estimates with those obtained from the CPS.

Our recommended impute estimator has the following advantages. First, it is easy to implement – it is a simple modification of the standard two stage least squares (2SLS) estimator, where the same instrumental variables (IVs) are used to impute the (unobserved) dependent variable within daily subsamples, as well as the (endogenous) independent variables with the entire sample. It essentially matches similar individuals based on the *exogenous* IVs only, and uses observed hours of matched individuals as imputed hours for those who were not surveyed by the ATUS on a particular day.[7] Second, it is consistent and asymptotically normal under the same conditions as if the *true* weekly hours worked were observed.[8] This paper is the first to systematically investigate asymptotic as well as finite sample properties of various feasible estimators of linear regression models using time use surveys. Our analysis is made easier and more transparent by employing the potential outcome framework. Third, the impute estimator is also asymptotically more efficient and numerically more stable in small samples, compared to the other estimators. These relative merits of the impute estimator may make a difference in statistical inference, since time use surveys are costly to conduct and have much smaller sample sizes than traditional surveys.

---

such condition.

[6]The scope of Frazis and Stewart (2012)'s paper is wider than ours – it covers both cases where the hours worked are the dependent variable and the independent variable, as well as other issues such as multiple activities, multiple diary days and multiple members in a household, etc.

[7]Aguiar et al. (2017)'s "synthetic time diary" approach shares similar spirits. They used both endogenous and exogenous covariates to match, but we prove that the consistency of all the estimators demands the matching to be based only on the exogenous IVs.

[8]If the true weekly hours worked were observed and the regressors (e.g., wage) are endogenous, then the usual 2SLS estimator only requires valid and relevant IVs to be available. As becomes clear below, the argument generalizes trivially to other time use surveys for more than one day.

In addition to the theory, we utilize the DTUS as a valuable benchmark since it contains accurate diary hours for seven consecutive days. We randomly draw one single day for each individual to imitate the ATUS. This artificial data set permits direct comparison among our proposed estimators and the usual 2SLS estimator, which is infeasible for the ATUS. Through this unique approach, we unambiguously demonstrate the superiority of our impute estimator.

Empirically, we find that the ATUS yields smaller own wage elasticities than the CPS across the board, but the gaps vary among gender and marital groups. Moreover, the ATUS indicates smaller spouse (cross) earning elasticity than the CPS for married women, but larger for married men. Furthermore, the ATUS exhibits weaker elasticity with respect to the number of older kids than the CPS for married women, even though the two surveys result in almost the same elasticity estimates with respect to the number of younger kids.

Empirical studies have found nonclassical measurement errors in many dependent variables (Duncan and Hill, 1985; Bound et al., 2001) including labor supply. But in theoretical econometrics literature, nonclassical measurement errors in dependent variables have drawn far less attention than independent variables (e.g., Hu and Schennach, 2008; Chen et al., 2005; Hu and Sasaki, 2015, 2017). A notable exception is Abrevaya and Hausman (1999). While our impute estimator naturally accommodates endogenous independent variables (e.g., wage), it is unclear, however, whether and how Abrevaya and Hausman (1999)'s estimator could be generalized to this case.

Contemplative readers may wonder: what is the significance of weekly labor supply? Why not estimate monthly, quarterly, or yearly labor supply? The most obvious reason is that the CPS records weekly hours,[9] and we need to aggregate the daily information in the ATUS in order to compare with the CPS. But more importantly, once we bridge the gap between daily hours and weekly hours, then going from weekly hours to longer time frame follows exactly the same logic.

For activities recorded in the ATUS other than working (e.g., Aguiar and Hurst, 2007; Guryan et al., 2008; Aguiar et al., 2017), time specificity problem remains. Time specificity problem also presents itself outside time use surveys, such as recalled food expenditure data (Ahmed et al., 2006; Sousa, 2014; Brzozowski et al., 2017) versus the diary system used in the Expenditure and Food

[9]The CPS asks the respondents how many hours he/she *usually* works per week, and how many hours he/she *actually* worked the week before, both for their main jobs and other jobs. In our empirical studies using CPS data, we used the number of hours per week that the respondents usually work.

Survey (EFS) in the UK.[10]

The rest of the paper is organized as follows. Section 2 gives more information about time use surveys and traditional surveys. In Section 3, we first state two impossibility results regarding the true weekly hours. Then we focus on the estimation of weekly labor supply parameters. We propose several intuitive estimators and recommend the impute estimator based on its superior asymptotic properties. Section 4 demonstrates its superior finite sample properties via simulations using the DTUS as the benchmark. Section 5 applies our impute estimator to the ATUS and compares it with the labor supply elasticity estimates produced by the CPS for the *same* respondents. Section 6 states a few of our comments on the design of time use surveys. Section 7 concludes.

The Supplementary Appendices collect the proofs, additional simulations, additional theoretical and empirical results, as well as various robustness checks of our empirical studies.

## 2 Time Use Surveys

The ATUS randomly draws a subsample of the respondents who just completed their participation in the CPS within the past two to five months.[11] On a randomly chosen day (interview day),[12] the respondents are asked to fill up a diary detailing all their activities minute-by-minute on the previous day (diary day). Adding all the time spent on working by each respondent yields his/her ATUS hours worked for the diary day. Since the respondents of the ATUS had already participated in the CPS, all the data collected by the ATUS and the CPS about them are available for analysis, including demographics and income.[13]

The ATUS has some distinct features that set it apart from commonly used surveys like the CPS. First, the respondents of the ATUS record their activities for only one day (diary day), as opposed to weeks or months. The diary day is completely randomly chosen, with weekends having higher probabilities than weekdays.[14] Second, the ATUS imposes a 24-hour limit on the

[10] The EFS became known as the Living Cost and Food Survey from January 2008.

[11] For the workers who satisfy the criterion for our empirical analysis in Section 5, the number of those who participate in the ATUS account for roughly 1/50 of all the respondents in the CPS.

[12] ATUS User's Guide (2020, Section 3.5) states that "The designated persons are then randomly assigned a day of the week about which to report".

[13] For a more detailed description of the ATUS, see Hamermesh et al. (2005).

[14] ATUS User's Guide (2020, Section 3.5) states that to "ensure good measures of time spent on weekdays and weekend days, ...10 percent of the sample is allocated to each weekday, and 25 percent of the sample is allocated to each weekend day". Weekends are oversampled since they are more informative about people's activities other than work.

time allocated to all recorded activities. These two features are likely to make the ATUS hours a much more accurate measure of the hours worked on a single day. Throughout this paper, we assume that the observed daily hours worked in the ATUS are the true hours worked for the diary day, only to simplify our analysis. We acknowledge that this assumption is almost certainly wrong, and that the incidence and the size of the measurement error in the ATUS daily hours should be carefully examined for any serious empirical research.[15]

On the contrary, the CPS records weekly hours, by asking either how many hours the respondents *usually* work per week or how many hours they *actually* worked in the previous week. While probably less accurate than the ATUS hours, the CPS hours concern a longer time period.

In order to quantify and rectify the consequences of error-ridden hours in the CPS using the more accurate ATUS hours, we have to understand and tackle this time specificity of the two data sources. This time specificity is the crux of this paper. To focus on the consequences of time specificity, we will only include the individuals who participated in both the CPS and the ATUS into our sample for empirical analysis, so that no differences in estimates or efficiency may result from the differences in samples.[16]

Such time specificity of hours between time use surveys and commonly used surveys is not unique to the US, presumably because of high costs of conducting time use surveys. In fact, to the best of our knowledge, the only country that has ongoing time use survey that records activities for an entire week is the Netherlands.[17] The DTUS has been carried out since 1975 and has been published every five years. In the week long diary, the participants record their main activity every ten minutes and a secondary activity that might take place at the same time. The survey randomly draw more than two thousand participants from the Dutch population aged 12 and above since 2006. For the same respondents, the DTUS also contains CPS-type recalled weekly hours and some demographics including age, gender, education and number of children. So the DTUS serves as a particularly precious benchmark against which we can evaluate different estimators. We are

[15]In Supplemental Appendix C, we consider the case where the ATUS hours contain classical measurement error and show that all the theorems in Section 3.2 still hold with minor modification.

[16]In Supplementary Appendix A, we compare the ATUS sample with a much bigger CPS sample and do not find significant difference between the distributions of key variables from the two samples.

[17]The UK time use surveys in 1973, 1974, 1983 and 1984 covered seven days of a week; and more recent time use surveys in the UK cover two days. While diary records for two days still suffer from time specificity problem, they are likely to provide partial information on weekly activity patterns that the ATUS cannot. Readers can refer to IPUMS (2020) for sample characteristics of time use surveys in different countries.

going to base our simulation studies on the DTUS. Unfortunately, the DTUS does not contain detailed information on income, which renders it unsuitable for our empirical analysis involving wage or earnings.[18] But the DTUS contains demographic information which allows us to draw some empirical findings about labor supply along that line.[19]

# 3 Good News and Bad News about Labor Supply

This section has good news and bad news. We start with the bad news in Section 3.1 – that is, what time use surveys cannot tell us about labor supply. Our analysis echoes Frazis and Stewart (2012) that neither the weekly hours worked nor its distribution can be identified using the ATUS type time use survey data. Then we proceed to the good news in Section 3.2 – that is, to provide consistent and relatively efficient estimators of labor supply elasticities and to investigate their properties.

## 3.1 Bad News: Potential Hours and Impossibility Results

The bad news delivered in this subsection is either reiteration or immediate corollaries of the reasoning in Section II of Frazis and Stewart (2012), but we rephrase it under the potential outcome framework. We find it necessary to reiterate Frazis and Stewart (2012)'s insights here because their significance and implications appear not to be fully appreciated in empirical research using time use surveys. In addition, the notation under the potential outcome framework is instrumental for our analysis in Section 3.2, the main theoretical results of this paper.

Let's start with a simple question: how do we recover the distribution of weekly hours worked from the ATUS daily hours data? Since the ATUS diary day is randomly drawn, one may think of the ATUS daily hours as a representative sample of the weekly hours and, therefore, the distribution of weekly hours may be recovered from the distribution of the ATUS daily hours with adjustment for diary day sampling weights.

A small experiment using the DTUS data illustrates that this is a bad idea. In Figure 1, the solid line shows the kernel density of the DTUS weekly hours worked, which is directly observable in the DTUS for each individual. To mimic the ATUS, we randomly choose one day from the

[18] The income variable in the DTUS is only the annual income quartiles.
[19] For a more detailed description of the DTUS, see Fisher et al. (2018).

DTUS as the diary day for each individual, and plot the kernel density of the hours worked on the diary day multiplied by 7. The dashed and the dotted lines in Figure 1 show the kernel densities for two such random experiments. They differ from the DTUS weekly hours significantly.[20]

It turns out that it is just impossible to identify the distribution of weekly hours from daily hours without ad hoc assumptions. Now we introduce notation to facilitate the discussion. Let the individual respondent be indexed by $i = 1, \ldots, n$. Let $H_i^w$ denote the true weekly hours worked by individual $i$. The recalled weekly hours worked $H_i^{CPS}$ in the CPS is an error-ridden measure of $H_i^w$,

$$H_i^{CPS} = H_i^w + e_i. \tag{1}$$

That the measurement error is nonclassical implies that $e_i$ could be correlated with $H_i^w$. Let $t \in \{1, \ldots, 7\}$ denote the days of a week,[21] and let $H_{it}$ denote the true daily hours worked by individual $i$ on day $t$. Naturally, the weekly hours worked equal to the sum of daily hours worked over the week,

$$H_i^w = \sum_{t=1}^{7} H_{it}. \tag{2}$$

Let $t_i$ be the diary day of individual $i$ in the ATUS, then the daily hours worked in ATUS, denoted as $H_i^{ATUS}$, is just $H_{it_i}$. To facilitate our analysis, it helps to write the ATUS daily hours in an alternative way. Let $d_{it} \equiv \mathbb{I}\{t_i = t\}$ be seven diary day dummy variables for each individual $i$.[22] Then

$$H_i^{ATUS} = H_{it_i} = \sum_{t=1}^{7} d_{it} H_{it}. \tag{3}$$

Since for any individual interviewed in the ATUS, one and only one of the seven diary dummies equals 1, we only have an accurate measure of his/her hours worked for a single day of the week, but not for the other six days.[23]

Now it helps to recall the conventional wisdom in the program evaluation literature that even in purely random experiments, neither individual treatment effect nor its distribution in the population

[20] In Supplementary Appendix A, we take the common five-day work schedule into account, and the results are similar.

[21] $t = 1$ indicates Sunday, $t = 2$ indicates Monday, and so on.

[22] The symbol $\equiv$ indicates that the quantity on the left side is defined as the expression on the right side.

[23] Throughout the paper, we assume that the hours worked in time use surveys are the true hours worked for the prescribed period. This is merely for the simplicity of exposition, because all the theoretical results still hold if the measurement error of hours worked in time use surveys is classical. In Supplementary Appendix C, we discuss this in details.

can be identified without ad hoc assumptions on the joint distribution of $(Y_{i1}, Y_{i0})$.[24] Following the program evaluation literature, we call $H_{it}$ "potential hours" of diary day $t$ $(t = 1, \ldots, 7)$.

The following impossibility results naturally follow. First, without ad hoc assumptions, it is impossible to recover individual weekly hours worked $H_i^w$ from what is available in the ATUS. This impossibility result implies some important limitations of the imputed weekly hours – see Remark 2 below for details. Second, the ATUS only contains information regarding the marginal distributions of daily hour worked on a single day, but provides no information about the dependence among $(H_{i1}, \ldots, H_{i7})'$. The latter is required to find out the distribution of weekly hours $H_i^w$. In consequence, the distribution of the weekly hours worked $H_i^w$ (as well as its variance) cannot be recovered using the ATUS daily hours data. Third, computing the standard errors of ATUS based estimators needs extra effort. Without the potential outcome framework, this was not obvious, but the reason will become clearer after we give the standard error formulas for various estimators in Theorem 7.

### 3.2 Good News: Labor Supply Parameters

As Section II.1 of Frazis and Stewart (2012) pointed out, the daily hours worked in the ATUS can be used to produce estimates of the weekly labor supply regression parameters, despite the impossibility results in Section 3.1.[25]. In particular, such parameters include labor supply elasticities, the application we will focus on in the rest of this paper. The reason, rephrased under the potential outcome framework, is as follows. The ATUS closely resembles purely random experiments, since the diary day is completely randomly chosen for each respondent. In random experiments, $E(Y_{i1} - Y_{i0})$, the average treatment effect (ATE) and $E(Y_{i1} - Y_{i0}|X_i = x)$, the conditional average treatment effect (CATE) can be identified and estimated using the data that records either $Y_{i1}$ or $Y_{i0}$ (but not both) for each individual. Since regression equations are essentially conditional mean models, both $E(H_i^w)$ and $E(H_i^w|X_i = x)$, the counterparts of the ATE and the CATE in our scenario, can be identified and estimated. In fact, the labor supply elasticity estimator we recommend later in this section resembles a similarity to the matching regression estimator of the ATE in that it uses

[24] Let $Y_{i1}$, $Y_{i0}$ and $d_i$ denote the outcome if treated, the outcome if not treated and the treatment indicator for individual $i$, respectively, then the observed outcome is $Y_i = d_i Y_{i1} + (1 - d_i) Y_{i0}$. It is well known that the individual treatment effect, defined as $Y_{i1} - Y_{i0}$, cannot be identified.

[25] They also pointed out that the Tobit model leads to inconsistent estimators even though there are many zero-value observations in time use surveys.

the actual ATUS daily hours worked by other individuals with similar exogenous characteristics to impute the six missing daily hours worked for each individual in the ATUS.

One unique feature, however, differentiates the labor supply elasticity estimation problem from the usual treatment effect estimation. Elasticity hinges on not only the mean, but also the partial derivative of the conditional mean function $\partial E(H_i^w|X_i = x)/\partial x$, which would correspond to $\partial E(Y_{i1} - Y_{i0}|X_i = x)/\partial x$ and seems not to have attracted much attention in the treatment effect literature. Because we focus on the partial effect of $X_i$, we find that unlike in the treatment effect literature where the matching estimator aims to impute $Y_{i1}$ or $Y_{i0}$ itself, the characteristics to impute the missing potential hours in our context must be exogenous predictors of the daily hours worked.

### 3.2.1 Model and Estimators

To be concrete, we consider the following equation of weekly labor supply,

$$H_i^w = X_i'\beta + U_i, \qquad i = 1, \ldots, n, \tag{4}$$

where $X_i$ is a $p \times 1$ vector of observable independent variables that affect hours worked with its first element being unit one. The explanatory variables $X_i$, including log wage in particular, tend to be correlated with $U_i$, and hence is often endogenous. Moreover, log wage may also be subject to measurement errors.[26] In response to these problems, we assume that a $q \times 1$ vector of IVs $Z_i$ is available.

The ideal but infeasible case is when the true weekly hours worked $H_i^w$ were observable for each individual. The usual 2SLS estimator is then

$$\hat{\beta}_{wk} = (X'P_zX)^{-1}(X'P_zH^w), \tag{5}$$

where $H^w \equiv (H_1^w, \ldots, H_n^w)'$, $X \equiv (X_1, \ldots, X_n)'$, $Z \equiv (Z_1, \ldots, Z_n)'$ and $P_z \equiv Z(Z'Z)^{-1}Z'$. Since it uses the unobservable true weekly hours worked, we call it *week* estimator and use it as an infeasible benchmark.

[26] See a survey paper by Bound et al. (2001) for details.

Now we consider how to utilize the observed ATUS daily hours. Because the ATUS is designed to survey about a randomly chosen day for each individual, we maintain the following assumption throughout the paper.[27]

**Assumption 1** (Random diary day)**.** *Diary day dummies* $(d_{i1}, \ldots, d_{i7})'$ *are independent from* $(X_i', Z_i', U_i, H_{i1}, \ldots, H_{i7})'$.

We used Pearson's chi-squared test to test the independence between the ATUS diary day and each of the other variables used in this paper.[28] The tests results are in Table A.11 of the Supplementary Appendix A, and they strongly support Assumption 1.

Let $H^{ATUS}$ denote the $n \times 1$ vector of ATUS daily hours. For each $t \in \{1, \ldots, 7\}$, suppose the subsample size for diary day $t$ is $n_t$, let $H_t \equiv (H_{1t}, \ldots, H_{nt})'$, and let $D_t$ denote an $n \times n$ diagonal matrix with elements $d_{it}$ $(i = 1, \ldots, n)$. What $D_t$ does is just to select the subsample for diary day $t$. Equation (3) can be re-written in such matrix notation as $H^{ATUS} = \sum_{t=1}^{7} D_t H_t$.

Since the diary day is chosen randomly, it appears natural to expect the day-to-day variation of daily hours worked within a week to cancel out in large samples if we pool all diary days together. This intuition leads to what we call *pool* estimator,

$$\hat{\beta}_{pool} \equiv (X' P_z X)^{-1} X' P_z \Big( \sum_{t=1}^{7} r_{nt} D_t H_t \Big). \tag{6}$$

In eq. (6), $r_{nt} \equiv n/n_t$ adjusts for the sampling probability of the diary days. If every day gets 1/7 probability of being sampled, then the pool estimator is equivalent to a simple 2SLS using the pooled ATUS daily hours multiplied by seven.

**Remark 1** (Adjusting for diary day sampling probabilities in $\hat{\beta}_{pool}$)**.** *Most empirical research adjusts for the sampling weights of* individuals*. But eq.* (6) *implies that for time use surveys, it is important to also adjust for the diary day sampling probabilities with* $r_{nt}$ *when pooling the observations, otherwise Theorem 2 below implies that* $\hat{\beta}_{pool}$ *will be inconsistent. The distinction between these two types of adjustment will be discussed further in Remark 12.*

[27] ATUS User's Guide (2020, Section 3.5) states that "10 percent of the sample is allocated to each weekday, and 25 percent of the sample is allocated to each weekend day."

[28] Most of our regressors are categorical variables, for which chi-squared test can be directly used. We bin the continuous variables, like hourly wage, according to their deciles before applying the chi-squared tests for them.

The second intuitive estimator relies on the disaggregation of the weekly labor supply model into a number of daily labor supply models; that is, for $t = 1, \ldots, 7$,

$$H_{it} = X_i'\beta_t + U_{it}, \tag{7}$$

where $E(U_{it}) = 0$. Then the parameters $\beta$ in the weekly labor supply model can be re-written as $\beta = \sum_{t=1}^{7} \beta_t$. Therefore, it appears to be a logical attempt to estimate $\beta$ using what we call *day* estimator, defined as

$$\hat{\beta}_{day} \equiv \sum_{t=1}^{7} \hat{\beta}_t = \sum_{t=1}^{7} (X'P_{zt}X)^{-1}X'P_{zt}H_t, \tag{8}$$

where for each $t \in \{1, \ldots, 7\}$, $\hat{\beta}_t$ is simply the usual 2SLS estimator of $\beta_t$ using only the subsample for diary day $t$, and $P_{zt} = (D_tZ)(Z'D_tZ)^{-1}(D_tZ)'$.

Later we are going to show that both the pool estimator and the day estimator are consistent under mild conditions. However, neither of them is ideal in terms of efficiency and robustness against small sample sampling errors. Instead, we propose a third feasible estimator, which deviates from the infeasible benchmark $\hat{\beta}_{wk}$ as little as possible, and we will show that the third estimator outperforms the first two in terms of asymptotic efficiency and small sample robustness.

In light of eq. (2) and the definition of $P_z$,[29] the infeasible 2SLS estimator $\hat{\beta}_{wk}$ can be re-written as

$$\hat{\beta}_{wk} = (X'P_zX)^{-1}X'P_z \sum_{t=1}^{7} H_t = (X'P_zX)^{-1}X'P_zZ \sum_{t=1}^{7} \left(Z'Z/n\right)^{-1} \left(Z'H_t/n\right).$$

By the law of large numbers, we know that $\left(Z'Z/n\right)^{-1} \left(Z'H_t/n\right) \xrightarrow{p.} [E(Z_iZ_i')]^{-1}E(Z_iH_{it})$. Assumption 1 implies that in this expression, the unconditional means equal to the conditional means, i.e., $[E(Z_iZ_i')]^{-1}E(Z_iH_{it}) = [E(Z_iZ_i' \mid d_{it} = 1)]^{-1}E(Z_iH_{it} \mid d_{it} = 1)$. As a result, we can use the subsample for diary day $t$, instead of the entire sample, to estimate the two conditional means for each $t$. Replace the last part of $\hat{\beta}_{wk}$ by its diary day $t$ counterpart, we get

$$\hat{\beta}_{im} \equiv (X'P_zX)^{-1}X'P_zZ \sum_{t=1}^{7} \left(Z'D_tZ/n_t\right)^{-1} \left(Z'D_tH_t/n_t\right). \tag{9}$$

We call this estimator *impute* estimator. In practice, impute estimator is easy to compute using

[29] Moreover, $P_z$ is an idempotent matrix, i.e., $P_zP_z = P_z$.

the ATUS data by the following steps:

1. ("$X$ first stage") Regress $X_i$ on $Z_i$ using the entire sample and take the fitted values $\hat{X}_i$;

2. ("$H$ first stage") For each diary day $t$, regress $H_i^{ATUS}$ (i.e., $H_{it_i}$) on $Z_i$ using the subsample $d_{it} = 1$ to get $\hat{\alpha}_t$, and impute the weekly hours worked by $\hat{H}_i^w = \sum_{t=1}^7 \hat{H}_{it} = \sum_{t=1}^7 Z_i'\hat{\alpha}_t$ for the entire sample;

3. ("Second stage") Regress $\hat{H}_i^w$ on $\hat{X}_i$ using the entire sample and get $\hat{\beta}_{im}$.

Compared to the usual 2SLS estimator, this estimator adds one more simple step in the middle where the values of the unobservable weekly hours $H_i^w$ is imputed based on the IV.

In the "$H$ first stage", if the hours worked by individual $i$ on day $t$ is not observed, the impute estimator essentially matches individual $i$ with those respondents in the diary day group $t$ who have similar values of $Z_i$ with her, and uses their hours worked as the imputed hours for individual $i$. This is similar to the "synthetic time diary" method employed by Aguiar et al. (2017). It also resembles the matching estimator in the treatment effect literature, except that here we make it clear that the basis for matching has to be exogenous IVs $Z_i$, and cannot be endogenous regressors $X_i$ in the weekly labor supply eq. (4).

**Remark 2** (Limitation of imputed weekly hours $\hat{H}_i^w$)**.** *It might be tempting to think of $\hat{H}_i^w$ as "predicted" weekly hours worked for worker $i$, and to use $\hat{H}_i^w$ to impute other variables. For example, one might propose to impute hourly wage rate by $I_i/\hat{H}_i^w$ for weekly paid workers, where $I_i$ is weekly earning of worker $i$. Unfortunately, our earlier impossibility result indicates that such use of $\hat{H}_i^w$, in general, is wrong. $\hat{H}_i^w$ is merely an intermediate that facilitates efficient estimation of $\beta$. In addition, our analysis emphasizes that imputation of $\hat{H}_i^w$ should only be based on the instruments $Z_i$, but not endogenous $X_i$. Even though in many cases the latter may deliver a better "predicted" weekly hours, it results in bias in $\beta$ estimates.*

**Remark 3** (Exogenous $X_i$)**.** *If $X_i$ are exogenous (hence $X_i$ are their own IVs), then $\hat{\beta}_{wk} = (X'X)^{-1}(X'H^w)$ simply becomes the OLS estimator for model eq. (4). It is easy to verify that in this case $\hat{\beta}_{day}$ is numerically identical to the impute estimator $\hat{\beta}_{im}$. The two differ if $X_i$ are endogenous.*

**Remark 4** (Classical measurement error in the ATUS)**.** *We acknowledge that time use surveys are not error free. Let $H_{it}$ be the true hours worked on day $t$, and let $H_{it}^{ATUS} = H_{it} + e_{it}^{ATUS}$ be the ATUS hours if respondent $i$ was interviewed for his/her hours worked on day $t$. In Supplementary Appendix C, we show that when $e_{it}^{ATUS}$ is classical measurement error, all the theoretical results that we will elaborate in Section 3.2.2 continue to hold, with only a same small adjustment term added to the asymptotic variances of all feasible estimators.*

### 3.2.2 Large Sample Properties

In this section, we will show that all proposed feasible estimators for the ATUS are consistent under the same conditions for the consistency of the usual 2SLS estimator, as if the true weekly hours worked were observed. In addition, we will show that the impute estimator has superior relative asymptotic efficiency. The proofs for all the theorems in this section are provided in Supplementary Appendix B. We maintain the following three assumptions throughout the paper.

**Assumption 2** (Random sample)**.** *For any $i \in \{1, \ldots, n\}$, the vector $(H_{i1}, \ldots, H_{i7}, X_i', Z_i', d_{i1}, \ldots, d_{i7})'$ is randomly drawn from the population.*

**Assumption 3** (Valid and relevant instrumental variables in weekly equation)**.** *Assume that $E(U_i Z_i) = 0$, $\operatorname{rank} E(Z_i Z_i') = q$ $(q \geq p)$, and $\operatorname{rank} E(Z_i X_i') = p$.*

**Assumption 4** (Diary day sampling probability)**.** *Assume that each day of a week has a positive probability of being sampled. That is, $0 < \Pr(d_{it} = 1) < 1$ for each day $t \in \{1, \ldots, 7\}$.*

**Theorem 1** (Identification)**.** *Under Assumptions 1 to 4, the unknown parameters $\beta$ are identified using the ATUS data.*

**Theorem 2** (Consistency)**.** *Under Assumptions 1 to 4, we have that $\hat{\beta}_{wk}$, $\hat{\beta}_{im}$, $\hat{\beta}_{pool}$, and $\hat{\beta}_{day}$ all converge to $\beta$ in probability as $n \to \infty$.*

**Remark 5** (Weak conditions for consistency of $\hat{\beta}_{day}$)**.** *We need to point out that all the estimators we consider, including the day estimator, are consistent under the weaker assumption that $E(U_i Z_i) = 0$ (Assumption 3), instead of the stronger $E(U_{it} Z_i) = 0$ (Assumption 5 below). That is, the IV only need to be valid for the weekly labor supply equation, and not necessarily so for each* *daily ones. Even if each daily 2SLS estimator $\hat{\beta}_t$ might be inconsistent for $\beta_t$, the day estimator $\hat{\beta}_{day}$ still is.*

The 2SLS estimator based on the CPS recalled weekly hours, on the other hand, is in general inconsistent. This is again a well known consequence of the nonclassical measurement error $e_i$ defined in eq. (1).[30]

To derive the asymptotic distributions, it helps to consider the "$H$ first stage" where the potential daily hours $H_{it}$ are regressed on the IV $Z_i$:

$$H_{it} = Z_i'\alpha_t + V_{it}, \tag{10}$$

and let $V_t = (V_{1t}, \dots, V_{nt})'$ denote the vector of projection residuals. By construction, $E(V_{it}) = 0$ and $E(Z_i V_{it}) = 0$.

To show the relative asymptotic efficiency of the estimators (feasible as well as the infeasible benchmark), we first give the following two theorems (Theorems 3 to 4). We need to emphasize that they are only to show the relative asymptotic efficiency of the estimators and should not be used to compute the standard errors of the feasible estimators, since they involve variables that are not observed in the data. Feasible standard error formulas will be based on the two theorems that follow (Theorems 5 to 6). Define $A \equiv BC^{-1}B'$ with $B \equiv E(X_i Z_i')$ and $C \equiv E(Z_i Z_i')$, and let $r_t = 1/\Pr(d_{it} = 1)$.

**Theorem 3** (Infeasible Benchmark)**.** *Under Assumptions 1 to 4, we have $\sqrt{n}(\hat{\beta}_{wk} - \beta) \xrightarrow{d} \mathcal{N}(0, \Omega_{wk})$, with*

$$\Omega_{wk} \equiv A^{-1}BC^{-1}E(Z_i U_i^2 Z_i')C^{-1}B'A^{-1}. \tag{11}$$

**Theorem 4** (Relative Asymptotic Efficiency)**.** *Under Assumptions 1 to 4, we have the following asymptotic normality results:*

*(i) $\sqrt{n}(\hat{\beta}_{im} - \beta) \xrightarrow{d} \mathcal{N}(0, \Omega_{im})$, with $\Omega_{im} = \Omega_{wk} + \Omega_{im-wk}$, where*

$$\Omega_{im-wk} \equiv A^{-1}BC^{-1}\Big[\sum_{t=1}^{7}(r_t - 1)E(Z_i V_{it}^2 Z_i') - 2\sum_{1 \le t < \tau \le 7} E(Z_i V_{it} V_{i\tau} Z_i')\Big]C^{-1}B'A^{-1}, \tag{12}$$

[30]The probability limit of $\hat{\beta}_{wk}^{CPS} \equiv (X'P_z X)^{-1}(X'P_z H^{CPS})$, the 2SLS estimator based on the CPS weekly hours, is $\beta + (E(X_i Z_i')[E(Z_i Z_i')]^{-1}E(Z_i X_i'))^{-1}E(X_i Z_i')[E(Z_i Z_i')]^{-1}E(Z_i e_i)$, and the latter term is in general not zero, since $E(Z_i e_i) \neq 0$ for the nonclassical measurement error $e_i$.

*and $\Omega_{wk}$ is the infeasible asymptotic variance-covariance matrix of the infeasible benchmark estimator $\hat{\beta}_{wk}$ defined in Theorem 3;*

*(ii)* $\sqrt{n}(\hat{\beta}_{pool} - \beta) \xrightarrow{d} \mathcal{N}(0, \Omega_{pool})$*, with* $\Omega_{pool} = \Omega_{im} + \Omega_{pool-im}$*, where*

$$\Omega_{pool-im} \equiv A^{-1}BC^{-1}\Big[\sum_{t=1}^{7}(r_t - 1)E(Z_i\alpha_t' Z_i Z_i' \alpha_t Z_i') - 2\sum_{1\leq t<\tau\leq 7} E(Z_i\alpha_t' Z_i Z_i'\alpha_\tau Z_i')\Big]C^{-1}B'A^{-1}; \quad (13)$$

*(iii)* $\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk}) \xrightarrow{d} \mathcal{N}(0, \Omega_{im} - \Omega_{wk})$*, hence* $n(\hat{\beta}_{im} - \hat{\beta}_{wk})'(\Omega_{im} - \Omega_{wk})^{-1}(\hat{\beta}_{im} - \hat{\beta}_{wk}) \sim \chi^2(p)$.[31]

As is clearly shown in the proof, both $\Omega_{im-wk}$ and $\Omega_{pool-im}$ are variance-covariance matrices of some random vectors (hence positive definite), which in turn implies that $\Omega_{pool} \geq \Omega_{im} \geq \Omega_{wk}$. Although the pool estimator is intuitive and consistent, the impute estimator is better in terms of asymptotic efficiency.

**Remark 6** (Relative efficiency of $\hat{\beta}_{wk}$)**.** *It is not surprising that the infeasible estimator $\hat{\beta}_{wk}$ is asymptotically the most efficient should the true weekly hours worked be observed. The efficiency difference between $\hat{\beta}_{wk}$ and $\hat{\beta}_{im}$ results from the fact that $\hat{\beta}_{im}$ only utilizes diary day subsamples to impute $\hat{H}_{it}$ (and sum $\hat{H}_{it}$ to get $\hat{H}_i^w$), while $\hat{\beta}_{wk}$ directly imputes $\hat{H}_i^w$ using the entire sample.*[32] *For this same reason, the efficiency loss of $\hat{\beta}_{im}$ compared to $\hat{\beta}_{wk}$ depends on the correlations among the daily hours for the same individual.*[33] *This can be seen from the second term in the square brackets in the expression of $\Omega_{im-wk}$ in Theorem 4(i).*

**Remark 7** (Relative efficiency of $\hat{\beta}_{im}$)**.** *The asymptotic efficiency gain of $\hat{\beta}_{im}$ compared to $\hat{\beta}_{pool}$ might be less expected. But is also very intuitive – to impute $\hat{H}_{it}$, $\hat{\beta}_{im}$ uses only data on $H_{it}$, the relevant diary day observations. On the contrary, $\hat{\beta}_{pool}$ uses data on both $H_{it}$ and $H_{i\tau}$ ($\tau \neq t$), and $H_{i\tau}$ observations merely add noise, which results in a less efficient estimator. An even less obvious point is that the size of the efficiency gap depends on the diary day sampling weights. In*

[31] Here we do not provide the asymptotic variance for $\hat{\beta}_{day}$, but we will provide asymptotic variance for $\hat{\beta}_{day}$ after imposing Assumption 5. Assumption 3 only guarantees that the IVs are valid for the weekly labor supply equation, but not necessarily for the daily ones, so $\hat{\beta}_t$ might be inconsistent for $\beta_t$ for some $t$. The asymptotic distribution of IV estimators, when the IVs are invalid, is very complicated in general (for details, see Kiviet and Niemczyk, 2009).

[32] Note that $\hat{\beta}_{wk} = (X'P_zX)^{-1}(X'P_zH^w) = (X'P_zP_zX)^{-1}(X'P_zP_zH^w) = (\hat{X}'\hat{X})^{-1}(\hat{X}'\hat{H}^w)$ since $P_z$ is an idempotent matrix.

[33] Or more precisely, the correlations among the residuals after projecting the daily hours on the IVs.

*the extreme case where there is no variation in daily hours (i.e.,* $H_{i1} = \cdots = H_{i7}$*, and hence* $E(Z_i V_{it} V_{it} Z_i') = E(Z_i V_{it} V_{i\tau} Z_i')$ *and* $E(Z_i \alpha_t' Z_i Z_i' \alpha_t Z_i') = E(Z_i \alpha_t' Z_i Z_i' \alpha_\tau Z_i')$ *for all* $t, \tau = 1, \ldots, 7$*), one might think that it does not matter which day gets surveyed, and hence the pool estimator (subject to sampling weights adjustment) suffices. However, part (ii) of Theorem 4 shows that for* $\Omega_{pool-im}$ *to be zero, we need equal sampling weights so that* $r_t = \lim_{n\to\infty} n/n_t = 7$ *for* $t = 1, \ldots, 7$*. Otherwise* $\Omega_{pool} > \Omega_{im} > \Omega_{wk}$ *remains. This means that, given the sampling weights of the ATUS diary days (i.e.,* $r_1 = r_7 = 4$ *and* $r_2 = \cdots r_6 = 10$*), the impute estimator will be more efficient than the pool estimator even when there is no variation in daily hours.*

**Remark 8** (Hausman test between the CPS and the ATUS)**.** *Part (iii) of Theorem 4 indicates that we can test the presence of nonclassical measurement errors in the recalled weekly hours in the CPS using the Hausman test. Under the null hypothesis of no nonclassical measurement errors, the 2SLS based on the recalled weekly hours in the CPS will be consistent and as efficient as the week estimator* $\hat{\beta}_{wk}$*; while under the alternative, such 2SLS will be biased. In both cases, the impute estimator* $\hat{\beta}_{im}$ *is consistent but less efficient.*

Even though Theorem 4 clearly ranks the estimators in terms of asymptotic relative efficiency, it is not very informative about how to compute the standard errors of the estimators. The reason is that in Theorem 4, both $E(Z_i U_i^2 Z_i')$ in $\Omega_{wk}$ and $E(Z_i V_{it} V_{i\tau} Z_i')$ $(1 \le t < \tau \le 7)$ in $\Omega_{im-wk}$ make it seem that one needs to observe the same individuals on different days in order to estimate $\Omega_{im}$, and the ATUS is inadequate in this regard. Fortunately, the asymptotic variances of $\hat{\beta}_{im}$ and $\hat{\beta}_{pool}$ can be computed without first deriving that for the infeasible $\hat{\beta}_{wk}$, which leads to straightforward formulas for the standard errors of the feasible estimators. Such results are summarized in the following theorem.

**Theorem 5** (Asymptotic Normality I)**.** *Under Assumptions 1 to 4, we have the following asymptotic normality results:*

*(i)* $\sqrt{n}(\hat{\beta}_{im} - \beta) \xrightarrow{d.} \mathcal{N}(0, \Omega_{im})$*, with*

$$\Omega_{im} \equiv A^{-1}BC^{-1}\left\{\sum_{t=1}^{7} r_t E(Z_i V_{it}^2 Z_i') + E\left[Z_i\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right)^2 Z_i'\right]\right.$$
$$\left. +2\sum_{t=1}^{7} E\left[Z_i V_{it}\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right)Z_i'\right]\right\} C^{-1}B'A^{-1}, \qquad (14)$$

*and note that $\Omega_{im}$ in eq. (14) equals to that given in Theorem 4(i);*

*(ii)* $\sqrt{n}(\hat{\beta}_{pool} - \beta) \xrightarrow{d} \mathcal{N}(0, \Omega_{pool})$*, with*

$$\Omega_{pool} \equiv A^{-1}BC^{-1}\left\{\sum_{t=1}^{7} r_t E\left(Z_i V_{it}^2 Z_i'\right) + \sum_{t=1}^{7} r_t E(Z_i \alpha_t' Z_i Z_i' \alpha_t Z_i') \right.$$
$$+ E\left(Z_i \beta' X_i X_i' \beta Z_i'\right) - 2\sum_{t=1}^{7} E\left[Z_i \alpha_t' Z_i X_i' \beta Z_i'\right]$$
$$\left. +2\sum_{t=1}^{7} E\left[Z_i V_{it}\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right) Z_i'\right]\right\} C^{-1}B'A^{-1}, \qquad (15)$$

*and note that $\Omega_{pool}$ in eq. (15) equals to that given in Theorem 4(ii).*

To derive the asymptotic normality for the day estimator, it is necessary to make an additional assumption.

**Assumption 5** (Instrumental variables in daily equations)**.** *Assume that $E(U_{it}Z_i) = 0$ for all $t = 1, \ldots, 7$, that is, the instrumental variables are valid in the daily labor supply equations.*

**Theorem 6** (Asymptotic Normality II)**.** *Under Assumptions 1 to 5, we have:*

*(i)* $\sqrt{n}(\hat{\beta}_{day} - \beta) \xrightarrow{d} \mathcal{N}(0, \Omega_{day})$*, with*

$$\Omega_{day} \equiv A^{-1}BC^{-1}\left[\sum_{t=1}^{7} r_t E(Z_i U_{it}^2 Z_i')\right] C^{-1}B'A^{-1}; \qquad (16)$$

*(ii) The asymptotic efficiency gap between the asymptotic variances of $\hat{\beta}_{day}$ and $\hat{\beta}_{im}$ is*

$$\Omega_{day} - \Omega_{im} = A^{-1}BC^{-1}\left[\sum_{t=1}^{7}(r_t - 1)\,\mathrm{E}[Z_i(U_{it} + V_{it})(U_{it} - V_{it})Z_i'] \right.$$
$$\left. - \sum_{t\neq\tau}\mathrm{E}[Z_i(U_{it} + V_{it})(U_{i\tau} - V_{i\tau})Z_i']\right] C^{-1}B'A^{-1}. \qquad (17)$$

Theorem 6(ii) reveals that there is no general efficiency ranking between $\hat{\beta}_{day}$ and the other two feasible estimators. Contrary to Theorem 2, the asymptotic normality of the day estimator $\hat{\beta}_{day}$ does require a stronger condition than the other estimators, i.e., Assumption 5. The reason can be seen from eq. (16), where $U_{it}$, the error term in the daily labor supply equation plays a central role.

In addition, $U_{it}$ in eq. (16) cannot be consistently estimated if Assumption 5 fails to hold since $\beta_t$ are not consistently estimable in this case, then the standard error of $\hat{\beta}_{day}$ will not be feasible to compute using the ATUS data.

**Remark 9** (Stronger conditions for asymptotic normality of $\hat{\beta}_{day}$)**.** *The distinction whether Assumption 5 is assumed could be consequential in certain contexts. For example, Goldin (2014, pp. 1091) found that "firms …disproportionately reward individuals who labored long hours and worked particular hours", and this is responsible for a noticeable proportion of gender gap in pay. In other words, comparing two workers who have the same unobserved factors that determine the weekly hours worked (i.e., the same $U_i$), the one who works a regular schedule (or can meet clients during particular periods, or can work when everybody else does, etc.) tends to be paid with a higher hourly wage than the one who works a flexible schedule. The association between flexible schedule (i.e., uneven allocation of $U_{it}$ among seven days) and lower wage (i.e., $X_i$) in turn implies that an IV that is valid for the weekly labor supply equation is well likely to be invalid for the daily equations.*

**Remark 10** (Relative efficiency of $\hat{\beta}_{day}$)**.** *In Supplementary Appendix B (following the proof of Theorem 6, we show that the sign of $\Omega_{day} - \Omega_{im}$ is not definite (might be positive or negative), indicating that there is no fixed asymptotic efficiency ranking between $\hat{\beta}_{day}$ and $\hat{\beta}_{im}$. The sign of $\Omega_{day} - \Omega_{im}$ depends on a number of factors, which will be detailed in Supplementary Appendix B. One can easily construct examples where $\Omega_{day} \geq \Omega_{im}$ or the opposite.*[34]

**Remark 11** (A Variation of $\hat{\beta}_{day}$)**.** *An anonymous referee suggested a variation of the day estimator, which we denote as $\tilde{\beta}_{day}$ and is computed using the ATUS data by the following steps:*

1. *("X first stage") Regress $X_i$ on $Z_i$ using the entire sample and take the fitted values $\hat{X}_i$;*
2. *("Daily second stage") For each diary day $t$, regress $H_i^{ATUS}$ (i.e., $H_{it_i}$) on $\hat{X}_i$ using the subsample $d_{it} = 1$ to get $\tilde{\beta}_t$, and let $\tilde{\beta}_{day} \equiv \sum_{t=1}^{7} \tilde{\beta}_t$.*

*This estimator is appealing for two evident reasons: first, compared to the impute estimator, $\tilde{\beta}_{day}$ does not need to impute $\hat{H}_i^w$ and saves computational burden; second, compared to the day estimator, $\tilde{\beta}_{day}$ imputes $\hat{X}_i$ using the entire sample and should be more stable in small samples.*

[34]Inspired by an anonymous referee, we conducted simple simulation experiments to demonstrate both cases. These results are not reported but available upon request.

*In Supplementary Appendix B, we will prove that $\tilde{\beta}_{day}$ is asymptotically equivalent to $\hat{\beta}_{im}$ under Assumptions 1 to 5.*

*Since $\tilde{\beta}_{day}$ consists of seven daily elasticity estimators $\tilde{\beta}_t$, its asymptotic normality requires Assumption 5, which we explained in Remark 9 might be too restrictive in some cases. For this reason, we believe that the impute estimator $\hat{\beta}_{im}$ is slightly more versatile. If a researcher is convinced that Assumption 5 is plausible in her particular study, then $\tilde{\beta}_{day}$ is simpler to compute and performs equally well as $\hat{\beta}_{im}$ in large samples.*[35]

The asymptotic variance formulas in Theorem 5 and Theorem 6 lead to easy-to-compute standard errors for $\hat{\beta}_{im}$, $\hat{\beta}_{pool}$ and $\hat{\beta}_{day}$. To introduce them, we need some notation.

Let $A_n \equiv n^{-1}\sum_{i=1}^n \hat{X}_i\hat{X}_i'$, $B_n \equiv n^{-1}\sum_{i=1}^n X_iZ_i'$ and $C_n \equiv n^{-1}\sum_{i=1}^n Z_iZ_i'$. Let $\hat{\alpha}_t$ be the OLS estimates of $\alpha_t$ in the "$H$ first stage" eq. (10) using the subsample for diary day $t$, and let $\hat{V}_{it}$ denote the residuals. Let $\hat{U}_{it} = H_{it} - X_i'\hat{\beta}_t$ denote the residuals of the daily labor supply eq. (7) using the subsample for diary day $t$.

Using this notation, we define

$$
\begin{aligned}
\hat{\Omega}_{im} \equiv A_n^{-1}B_nC_n^{-1}\Bigg\{&\sum_{t=1}^{7} r_{nt}\left(\frac{1}{n_t}\sum_{i=1}^{n} d_{it}Z_i\hat{V}_{it}^2Z_i'\right) + \left[\frac{1}{n}\sum_{i=1}^{n} Z_i\left(Z_i'\sum_{t=1}^{7}\hat{\alpha}_t - X_i'\hat{\beta}_{im}\right)^2 Z_i'\right] \\
&+2\sum_{t=1}^{7}\left[\frac{1}{n_t}\sum_{i=1}^{n} d_{it}Z_i\hat{V}_{it}\left(Z_i'\sum_{s=1}^{7}\hat{\alpha}_s - X_i'\hat{\beta}_{im}\right)Z_i'\right]\Bigg\} C_n^{-1}B_n'A_n^{-1}, \qquad (18) \\
\hat{\Omega}_{pool} \equiv A_n^{-1}B_nC_n^{-1}\Bigg\{&\sum_{t=1}^{7} r_{nt}\left(\frac{1}{n_t}\sum_{i=1}^{n} d_{it}Z_i\hat{V}_{it}^2Z_i'\right) + \sum_{t=1}^{7} r_{nt}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\hat{\alpha}_t'Z_iZ_i'\hat{\alpha}_tZ_i'\right) \\
&+\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\hat{\beta}_{pool}'X_iX_i'\hat{\beta}_{pool}Z_i'\right) - 2\sum_{t=1}^{7}\left[\frac{1}{n}\sum_{i=1}^{n} Z_i\hat{\alpha}_t'Z_iX_i'\hat{\beta}_{pool}Z_i'\right] \\
&+2\sum_{t=1}^{7}\left[\frac{1}{n_t}\sum_{i=1}^{n} d_{it}Z_i\hat{V}_{it}\left(Z_i'\sum_{s=1}^{7}\hat{\alpha}_s - X_i'\hat{\beta}_{im}\right)Z_i'\right]\Bigg\} C_n^{-1}B_n'A_n^{-1}, \qquad (19) \\
\hat{\Omega}_{day} \equiv A_n^{-1}B_nC_n^{-1}&\left[\sum_{t=1}^{7} r_{nt}\left(\frac{1}{n_t}\sum_{i=1}^{n} d_{it}Z_i\hat{U}_{it}^2Z_i'\right)\right] C_n^{-1}B_n'A_n^{-1}. \qquad (20)
\end{aligned}
$$

**Theorem 7** (Standard errors)**.** *Under Assumptions 1 to 4, we have the following results: (i) $\hat{\Omega}_{im} \xrightarrow{p.} \Omega_{im}$; (ii) $\hat{\Omega}_{pool} \xrightarrow{p.} \Omega_{pool}$. If in addition we assume Assumption 5 holds, then we also*

[35] In the DTUS-based simulations in Section 4.1, where Assumption 5 is satisfied, the performance of $\tilde{\beta}_{day}$ is almost identical to $\hat{\beta}_{im}$.

*have $\hat{\Omega}_{day} \xrightarrow{p.} \Omega_{day}$.*

**Remark 12** (Standard error of $\hat{\beta}_{pool}$)**.** *One may be inclined to compute the standard error of the pool estimator $\hat{\beta}_{pool}$ using stratification formula (for example, eq. (20.8) in Wooldridge, 2010), provided that the sampling weights are adjusted for.*[36] *But we need to point out that eq.* (6) *is conceptually and mathematically different from adjusting for the weights in stratified sampling designs, where $r_{nt}$, the inverse of the sampling weight enters both the numerator and the denominator of the estimator, while $r_{nt}$ enters our $\hat{\beta}_{pool}$ only in the numerator.*

## 4 Lessons from the Dutch Time Use Survey

The sample we use in this section consists of individuals from the DTUS (see Fisher et al., 2018, for details) aged between 25 and 54 surveyed in 1980, 1985, 1990, 1995, 2000 and 2005, whose recalled hours and recorded diary hours are both positive. The entire sample contains 6,567 individual-year records.

### 4.1 Simulations

Based on the DTUS data, we design a simulation study to compare the finite sample performance of the estimators discussed previously. The nice thing about the DTUS is that it contains CPS-type recalled weekly hours, as well as daily diary hours for an entire week. As a result, we are able to compute the week estimator $\hat{\beta}_{wk}$, which would have been impossible for the ATUS.

Given the daily hours worked $H_{it}^{DTUS}$ ($t = 1, \dots, 7$) in the DTUS, we generate a single endogenous regressor $\tilde{X}_i$ and a single instrumental variable $\tilde{Z}_i$ such that eq. (7) is satisfied with $X_i = (1, \tilde{X}_i)'$, $Z_i = (1, \tilde{Z}_i)'$, $\mathrm{Corr}(U_{it}, Z_i) = 0$ for $t = 1, \dots, 7$. In particular, let $H^{DTUS}$ denote the $n \times 7$ matrix with elements $H_{it}^{DTUS}$, and let $T_1, \dots, T_7$ be the principal components of $H^{DTUS}$. We set $\tilde{Z}_i$ to be the first principal component of $H^{DTUS}$, i.e., $\tilde{Z} = T_1$. To introduce the endogeneity, we generate an $n \times 7$ matrix of independent random variables from $N(0, 2)$,[37] denoted by $V$. Then we set $H_{it} = H_{it}^{DTUS} + V_{it}$ and $\tilde{X}_i = \tilde{Z}_i + \rho \sum_{t=1}^{7} V_{it}$ for $i = 1, \dots, n$ and $t = 1, \dots 7$. The true

[36] In fact, to the best of our knowledge, the current literature using the ATUS is not explicit about whether and how the diary day sampling probabilities are adjusted for (see, for example, Frazis and Stewart, 2012; Barrett and Hamermesh, 2019). We do not want to speculate how the standard error is computed in the literature, and here we only base our discussion on the formula of $\hat{\beta}_{pool}$ in eq. (6).

[37] Such that $\mathrm{Var}(U_i) \approx \mathrm{Var}(T_1)$ in the exogenous regressor case.

parameters $\beta_t$ are therefore just the weights in $H_t$ $(t = 1, \ldots, 7)$ associated with the first principal component. The true value of $\beta$ in eq. (4) is 2.2694.[38] By varying $\rho$, we vary $\text{Corr}(\tilde{X}_i, U_i)$, the degree of endogeneity of the regressor $\tilde{X}_i$. When $\rho = 0$, the regressor is exogenous, and we try other values of $\rho$ such that $\text{Corr}(\tilde{X}_i, U_i) \in \{0, 0.25, 0.5, 0.75\}$. Note that as $\rho$ increases, the strength of the IV also decreases. For the above values of $\rho$, $\text{Corr}(\tilde{X}_i, \tilde{Z}_i)$ equals 1, 0.95, 0.80 and 0.43, respectively.

To evaluate the finite sample performance of the various estimators considered in Section 3, we randomly draw a subsample of size $n \in \{250, 500, 1000, 2500\}$. Then we generate fictitious ATUS-type samples by randomly choosing only one day for each individual in the drawn subsamples using the diary day sampling weights of the ATUS.[39] We repeat the experiment $10,000$ times, and Table 1 reports the mean squared errors (MSE), squared biases and variances for all estimators.

Some patterns are apparent. First, the usual 2SLS estimator using the CPS-type recalled weekly hours, $\hat{\beta}_{re}$, has the largest MSE in almost all parameterizations, which is roughly ten times larger than the maximum among all the other estimators. The large MSE is nearly entirely driven by the large bias, which is in turn a result of nonclassical measurement error in the CPS-type recalled weekly hours. Below we will illustrate this nonclassical measurement error using the DTUS data in Figure 2. Second, for almost all parameterizations, the biases of all the estimators based on the diary hours are negligible, and the differences in the performance of $\hat{\beta}_{wk}$, $\hat{\beta}_{im}$, $\hat{\beta}_{pool}$ and $\hat{\beta}_{day}$ reside in efficiency and robustness. Third, since the infeasible week estimator $\hat{\beta}_{wk}$ uses the diaries of an entire week, it is much more efficient than the others. This verifies the result of Theorem 4(i). Fourth, the impute estimator $\hat{\beta}_{im}$ is more efficient than $\hat{\beta}_{pool}$ and $\hat{\beta}_{day}$ in all parameterizations. Again, this verifies the result of Theorem 4(ii). Fifth, when the regressor is exogenous, $\hat{\beta}_{im}$ and $\hat{\beta}_{day}$ perform equally well. Sixth, the MSE, biases and variances of $\tilde{\beta}_{day}$ discussed in Remark 11 is almost identical (not reported) to those of $\hat{\beta}_{im}$, since Assumption 5 is satisfied here. This is because, as mentioned in Remark 3, the two estimators are numerically identical in this case. Last but not least, the day estimator $\hat{\beta}_{day}$ appears to be unstable, especially when the sample size is smaller and when the IV is weaker. The reason is that $\hat{\beta}_{day}$ relies on the daily 2SLS estimators $\hat{\beta}_t$. When the sample size is small, the effective sample size for each day gets even smaller, and taking

[38] $\beta_1 = 0.0007$, $\beta_2 = 0.4379$, $\beta_3 = 0.4554$, $\beta_4 = 0.4576$, $\beta_5 = 0.4528$, $\beta_6 = 0.4304$ and $\beta_7 = 0.0346$.

[39] That is, the probability of being drawn is 0.25 for $t = 1$ (Sundays) and $t = 7$ (Saturdays), and 0.1 for the others.

the inverse of the sample average matrices magnifies the sampling errors substantially.[40]

### 4.2 Labor Supply Elasticity Estimates

In this section, we illustrate the empirical impacts on the labor supply elasticity estimates of both nonclassical measurement errors and time specificity using the DTUS.

Figure 2 shows the measurement error in the recalled weekly hours worked in the Dutch data. The "measurement error" in Figure 2 equals the recalled weekly hours worked minus the weekly hours worked from the seven-day diaries in the DTUS.[41] If the recalled hours worked do not have nonclassical measurement error, then the measurement error in Figure 2 would be uncorrelated with the weekly hours from the seven-day diaries. Panel A of Figure 2 suggests the opposite: the measurement error in the recalled hours is negatively correlated with the hours from time use survey. Its kernel density (panel B of Figure 2) suggests that more people overstate the recalled hours worked than understate. The negative correlation between the measurement error and the true hours worked coincides with the observation made by Bound et al. (1989) about the PSID.

We estimate the labor supply elasticities using the following model,

$$H_i^w = \beta_0 + \beta_1 kids_i + \beta_2 edu_i + \beta_3' X_i + U_i, \tag{21}$$

where $kids_i$ is the number of kids aged under 18, $edu_i$ includes two dummy variables, one for completing secondary education and the other for higher than secondary education, and $X_i$ is a vector of control variables, including age, age-squared, a dummy of working in private sector, an urban area dummy, and year dummies.

Table 2 shows the effects of the number of children and education on labor supply. We use both the recalled weekly hours worked ($\hat{\beta}_{re}$) and the seven-day diary hours ($\hat{\beta}_{wk}$) as the dependent variable. We also randomly draw one day for each respondent, then apply our impute estimator ($\hat{\beta}_{im}$). For both married men and married women, $\hat{\beta}_{re}$ are considerably different from $\hat{\beta}_{wk}$ and $\hat{\beta}_{im}$,

[40] We also conduct the same simulations based only on the five weekdays in the DTUS. The results are qualitatively the same and are reported in Table A.1 in Supplementary Appendix A.

[41] In the DTUS, the recalled weekly hours combine three questions in the survey: (1) hours worked in the previous week; (2) usual weekly hours worked in the previous year; or (3) the seven-day diary hours. The answer to the next question will be used if the respondents are unable to answer the previous question(s), but it is not indicated the answer to which question was actually used for each individual. Probably due to this, many respondents in the DTUS exhibit "zero measurement errors" in the recalled weekly hours.

with different signs when significant. In the meantime, the latter two always have the same signs, even though the magnitudes may differ. We conduct joint Hausman tests for the three coefficients in the table between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$, and between $\hat{\beta}_{wk}$ and $\hat{\beta}_{im}$. For both married women and married men, the Hausman tests reject the null hypotheses $\hat{\beta}_{re} = \hat{\beta}_{im}$ but do not reject $\hat{\beta}_{wk} = \hat{\beta}_{im}$.

Based on the time use survey hours, both $\hat{\beta}_{wk}$ and $\hat{\beta}_{im}$ indicate that the effects on married women's labor supply are significantly negative for more children and significantly positive for higher education. The recalled hours, on the other hand, produce $\hat{\beta}_{re}$ estimates that are too noisy to draw a conclusion.

# 5 Comparing Labor Supply Elasticity Estimates Using the ATUS and the CPS

In this section, we compare the labor supply elasticity estimates resulted from the CPS recalled weekly hours and the ATUS daily diary hours.

## 5.1 Empirical Sample and Summary Statistics

The data are from the 2003–2017 ATUS (Hofferth et al., 2018). As mentioned in Section 2, the ATUS sample is randomly drawn from the outgoing rotation group of the CPS respondents.[42] Therefore, for every respondent in the ATUS, we have their answers to all CPS questions as well. The sample used for our empirical analysis consists of hourly paid workers[43] aged between of 25 and 54, whose wage rate is positive, and spouse earnings (if married) and total usual weekly hours worked at current job reported in the last CPS interview are observed. The age restriction is to avoid complications of schooling and retirement decisions. The hourly wage rate was trimmed at percentiles 1 from below and 99 from above. After the trimming, the hourly wage in the sample ranges from \$5.2 to \$67.8 for men and from \$3.6 to \$63.1 for women (2017 US dollars).

We argue that the discrepancies between the ATUS sample and the CPS sample are small, and the reasons are as follows. First, if the respondent changed job (or changed employment status) since the last CPS interview, then her answers to related CPS questions are updated at the time of the ATUS interview, and we use the updated CPS hours whenever applied. This eliminates the

[42] The ATUS is conducted, on average, two to five months after their participation in the CPS.

[43] We exclude salaried workers because their hourly wage rate is much harder to measure. In a typical survey, the hourly wage for salaried workers is total earnings during a particular period divided by the hours worked in that period.

discrepancy due to job or employment status change. Second, we include only the respondents who answered both the CPS and the ATUS questions for themselves. This removes the discrepancies due to someone speculating someone else's CPS hours.[44] Third, we verify that those respondents who made their way into the ATUS sample are representative of the larger CPS sample,[45] even though the response rate of the ATUS might seem low to sharp eyes.[46]

Panel A of Table 3 provides means and standard deviations of the hours worked and hourly wage rate, computed using both the CPS and the ATUS for the same respondents in our empirical analysis sample. The CPS weekly hours worked we use is the number of hours per week that the respondent usually works at his/her current job at the reported hourly wage rate.[47] Here we calculate a lower bound in the following way. It is reasonable to assume that the correlation between the hours worked by the same person in two days, $H_{it}$ and $H_{it'}$, is nonnegative. By $H_i^w = \sum_{t=1}^{7} H_{it}$, we have $\text{Var}(H_i^w) \geq \sum_{t=1}^{7} \text{Var}(H_{it})$, where $\text{Var}(H_{it})$ can be readily estimated by the sample variance of hours worked on day $t$ in the ATUS. According to Table 3, men work slightly more hours than women regardless of marital status and data source; married men work slightly more hours than unmarried men, but married women work less than unmarried women; and for both genders, the married have higher hourly wage rates than the unmarried.

### 5.2 Labor Supply Elasticities

We estimate the labor supply elasticities using the following linear regression model,

$$H_i^w = \beta_0 + \beta_1 \ln w_i + \beta_2 y_i^{sp} + \beta_3 kid5_i + \beta_4 kid18_i + \beta_5' X_i + U_i, \tag{22}$$

[44] Every respondent in the ATUS records time diary for themselves, but the household head might answer the CPS questions on behalf of other household members.

[45] We compare our empirical analysis sample (from the ATUS) to two larger CPS samples. The first is the 2003–2017 CPS sample (regardless of whether the respondent took part in the ATUS or not) after applying the same criterion (age, trimming of hourly wage, etc.) used for our empirical analysis sample, and the other is the entire 2003-2017 CPS sample of hourly paid workers (all age, no trimming of hourly wage, etc.). In Supplementary Appendix A (submitted together with this paper), Table A.3 tabulates the summary statistics of many key variables for all these three samples, and all of the them are essentially the same across the three samples. Moreover, Table A.4 in Supplementary Appendix A reports the summary statistics of weekly hours and the weekly labor supply elasticity estimates based on the first larger CPS sample, and they are very close to the CPS based estimates reported in Table 3 of this paper.

[46] The average response rate of the ATUS is roughly 50%, and that of the CPS is higher than 80%.

[47] By the potential outcome argument, the standard deviation of the ATUS imputed weekly hours worked is impossible to compute without ad hoc assumptions. This was first pointed out by Frazis and Stewart (2012).

where $\ln w_i$ is the natural log of hourly wage, $y_i^{sp}$ is the usual weekly earnings of $i$'s spouse ($y_i^{sp} = 0$ for unmarried worker), $kid5_i$ is the number of children aged below 5, $kid18_i$ is the number of children aged between 5 and 18, and $X_i$ is a vector of control variables, including age, age-squared, two education dummies, eight Census division dummies, a metropolitan area dummy, race dummies, year dummies, occupation dummies and industry dummies. For hourly paid workers, ATUS directly asks them the hourly wage rate. Note that our sample only consists of those individuals who participated in both the ATUS and the CPS. As a result, despite the use of different measures of hours worked, $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$ are built on the same sample size.

In order to safeguard against the potential classical measurement error problem in wage and spouse weekly earnings, we use wage deciles and spouse earning deciles as IVs.[48] The reported estimates of elasticities here are evaluated at the respective sample mean hours worked per week.

Panel B of Table 3 shows the estimation results by gender and marital status. For each explanatory variable of interest, we report both $\hat{\beta}_{re}$, which is based on the CPS recalled weekly hours and $\hat{\beta}_{im}$, which is based on the ATUS daily hours and our proposed imputation method. For the CPS based $\hat{\beta}_{re}$, the standard errors are just the usual 2SLS standard errors. For the ATUS based $\hat{\beta}_{im}$, however, we report the standard errors computed using eq. (18). We conduct joint Hausman tests of the coefficients of the variables appearing in Panel B between the CPS and the ATUS. The $p$-values are smaller than 0.1 for unmarried men and married women.[49]

Both the CPS and the ATUS indicate that women's labor supply is more wage-elastic than that of men, with the labor supply of married women having the largest wage elasticity (0.1589 and 0.1048 respectively). Compared to the CPS, the ATUS results in smaller own wage elasticities across the board, and this agrees with what Barrett and Hamermesh (2019) found.[50] From the CPS to the ATUS, the reduction of own wage elasticities for men exceeds that for women, raising the relative own wage elasticities for women.[51] For married women, the ATUS yields much smaller

[48]We report OLS estimates in Table A.9 of the Supplementary Appendix A, and the results are almost the same, both qualitatively and quantitatively. Our choice of IVs follows the suggestion by Juhn and Murphy (1997) and Blau and Kahn (2007). The reason that wage (or spouse earning) decile serves as a valid IV for wage (or spouse earning) in the presence of classical measurement errors is that we believe that the variation in the measurement error is not big enough to alter the decile grouping for a substantial proportion of respondents.

[49]The empirical findings in Table 3 are very robust to choice of IVs, subsamples, or definition of "work" activities. Robustness checks are provided in Supplementary Appendix A.

[50]Note that this pattern cannot be explained by respondents bunching their recalled weekly hours at 40, the usual suspect of nonclassical measurement errors, which alone will result in lower elasticities from the CPS than from the ATUS.

[51]Heckman (1993) argues that an important reason that married women display higher own wage elasticity than

cross earning elasticity than implied by the CPS ($-0.0579$ v.s. $-0.0943$). For married men, the CPS indicates that their labor supply is non-elastic with respect to spouse earnings ($-0.0019$), consistent with previous findings in the literature (e.g., Blau and Kahn, 2007); notwithstanding, the ATUS produces a much higher cross earning elasticity, and it is comparable with that of married women ($-0.0347$).[52]

Our estimates based on the CPS are on par with those in the literature. Across roughly twenty estimates surveyed by Blundell and MaCurdy (1999), the median own wage labor supply elasticity is 0.08 for men and 0.78 for married women. For cross wage elasticities and conditional on having positive hours, Devereux (2004) reports around $-0.06$ for men and around $-0.5$ for women in the 1980s. For married women, Blau and Kahn (2007) document robust and substantial decline in married women's labor supply elasticities from 1980s to 2000s. Their own wage elasticity fell from roughly 0.77 in 1980 to roughly 0.36 in 2000. Their cross wage elasticity decreased from around $-0.33$ in 1980 to around $-0.19$ in 2000. Since our sample covers the years 2003–2017, the fact that all of our CPS based estimates have smaller absolute values than those for 2000 in Blau and Kahn (2007) is consistent with the decline in the responsiveness of married women's labor supply.[53]

Panel B of Table 3 also gives interesting elasticity estimates with respect to number of kids. For married women, both surveys lead to very large and almost identical elasticities with respect to number of younger kids for married women ($-0.0897$ and $-0.0858$); and yet the ATUS yields a much smaller elasticity with respect to number of older kids than the CPS ($-0.012$ v.s. $-0.0287$). For married men, the ATUS implies more elastic labor supply with respect to numbers of kids than the CPS as well.

It is worth mentioning that there are many possible sources that result in different elasticity estimates between the CPS and the ATUS (see Section 5 of Bound et al., 2001, for example), and the mean-reverting error[54] is only one of them. In fact, if the mean-reverting error was the only reason for different estimates between the CPS and the ATUS, then all the ATUS elasticity

men and unmarried women is because their labor force participation decision is more wage-elastic, and that basing the estimation only on those who work essentially compares married women's higher extensive margin elasticity with other groups' intensive margin elasticities. Due to the lack of good instruments, we don't correct for the sample selection bias and acknowledge that our elasticity estimates are hybrid of both margins.

[52]The Hausman test of this single coefficient rejects the null hypothesis of equal coefficients between the CPS and the ATUS.

[53]In addition, Blau and Kahn (2007) report men's own wage elasticities around 0.1 without notable time trend.

[54]That is, people who work more hours tend to under-report, and those who work fewer hours tend to over-report.

estimates (with respect to own wage, spouse earning, and number of kids) would all have had larger absolute values than their CPS counterparts. This contradicts our empirical findings in Table 3, where the ATUS indicates more elastic labor supply with respect to some regressors for some groups and the opposite in other cases. The same patterns are very robust across various robustness checks we conduct.[55]

## 6 Comments on Time Use Survey Design

For $\hat{\beta}_{im}$ to have the same precision as $\hat{\beta}_{wk}$, how much larger the sample size have to be? To get a rough idea, let's assume homoskedasticity so that $\Omega_{im-wk}$ and $\Omega_{wk}$ simplify to $\Omega_{im-wk} = \left[\sum_{t=1}^{7}(r_t - 1)E(V_{it}^2) - 2\sum_{1\le t<\tau\le 7} E(V_{it}V_{i\tau})\right]A^{-1}$ and $\Omega_{wk} = E(U_i^2)A^{-1}$. Using the DTUS data,[56] and we get $\hat{\mathrm{E}}(U_i^2) = 146$ and $\sum_{t=1}^{7}(r_t - 1)\hat{\mathrm{E}}(V_{it}^2) - 2\sum_{1\le t<\tau\le 7} \hat{\mathrm{E}}(V_{it}V_{i\tau}) = 409$. Hence the estimates of the asymptotic variances of $\hat{\beta}_{wk}$ and $\hat{\beta}_{im}$ are $\widehat{\mathrm{Var}}(\hat{\beta}_{wk}) = n^{-1}146A^{-1}$ and $\widehat{\mathrm{Var}}(\hat{\beta}_{im}) = n^{-1}(146 + 409)A^{-1}$.

If the correlation coefficients among the impute residuals of hours worked across different days in the ATUS are the same as in the DTUS, then such back-of-envelop calculation implies that compared to a survey that records the respondents' activities for an entire week and enables the use of the week estimator $\hat{\beta}_{wk}$, the number of respondents surveyed in the ATUS has to be roughly 3.8 times in order to get an impute estimator $\hat{\beta}_{im}$ with the same precision. For survey designers, this implies that if the average costs of following the same individuals for seven consecutive days is higher than 3.8 times of interviewing them for one day, then the latter is justified from the efficiency point of view.[57]

It is still worthwhile to do the former, at least in a smaller pilot sample. Knowledge about the correlation among the daily hours can help determine the sampling scheme that gives rise to the most efficient impute estimator. The reason is that $\Omega_{im-wk}$ in eq. (12) as well as $\Omega_{im}$ in eq. (14) depends on the diary day sampling probabilities $1/r_t$ $(t = 1, \ldots, 7)$. If efficiency of $\hat{\beta}_{im}$ is our primary concern, then we can minimize $\Omega_{im-wk}$ (or equivalently, $\Omega_{im}$) by choosing $r_t$ subject to the constraint $\sum_{t=1}^{7} 1/r_t = 1$. The optimal sampling probabilities are $1/r_t = \sigma_t/\sum_{s=1}^{7}\sigma_s$, where

[55] Reported in Table A.7 to Table A.10 in Supplementary Appendix A.

[56] The variables in $Z_i$ are the same as in Section 4.2, and the diary day sampling weights are in accordance with the ATUS weights.

[57] In addition, the time use survey hours become less reliable as the period of survey gets longer.

$\sigma_t^2 \equiv E(V_{it}^2)$ and we assumed homoskedasticity for simplicity. That is, more weights should be given to the days on which the hours worked exhibit larger variation among the population.

## 7 Conclusion

In this paper, we propose several intuitive estimators of weekly labor supply parameters using daily hours data in time use surveys and recommend the impute estimator on the ground of efficiency and robustness after carefully examining their asymptotic and finite sample properties. The impute estimator is a simple modification of the usual 2SLS estimator, which imputes the dependent variable (within each diary subsample) as well as the independent variables using the instruments. We then proceed to illustrate the finite sample properties of all the estimators we consider using the DTUS data, which track the respondents' activities for an entire week, and hence is a valuable benchmark. Multiple empirical findings are also drawn from the DTUS data. Finally, we compare the estimated labor supply elasticities using the ATUS impute estimator and that using the CPS recalled hours, and we are able to get a number of interesting empirical findings that are new in the labor economics literature.

## References


Abrevaya, J. and Hausman, J. A. (1999). Semiparametric estimation with mismeasured dependent variables: An application to duration models for unemployment spells. *Annales d'Économie et de Statistique*, (55/56):243–275.

Aguiar, M., Bils, M., Charles, K. K., and Hurst, E. (2017). Leisure luxuries and the labor supply of young men. Technical report, NBER.

Aguiar, M. and Hurst, E. (2007). Measuring trends in leisure: The allocation of time over five decades. *The Quarterly Journal of Economics*, 122(3):969–1006.

Aguiar, M., Hurst, E., and Karabarbounis, L. (2012). Recent developments in the economics of time use. *Annual Review of Economics*, 4(1):373–397.

Ahmed, N., Brzozowski, M., and Crossley, T. F. (2006). Measurement errors in recall food consumption data. IFS Working Papers WP06/21.

ATUS User's Guide (2020). American time use survey user's guide: Understanding atus 2003 to 2019. Technical report, U.S. Census Bureau.

Barrett, G. F. and Hamermesh, D. S. (2019). Labor supply elasticities: Overcoming nonclassical measurement error using more accurate hours data. *Journal of Human Resources*, 54(1):255–265.

Blau, F. D. and Kahn, L. M. (2007). Changes in the labor supply behavior of married women: 1980–2000. *Journal of Labor Economics*, 25(3):393–438.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In *Handbook of Labor Economics*, volume 3, pages 1559–1695. Elsevier.

Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. In *Handbook of Econometrics*, volume 5, pages 3705–3843. Elsevier.

Bound, J., Brown, C. C., Duncan, G., and Rodgers, W. L. (1989). Measurement error in cross-sectional and longitudinal labor market surveys: Results from two validation studies. Working Paper 2884, NBER.

Brzozowski, M., Crossley, T. F., and Winter, J. K. (2017). A comparison of recall and diary food expenditure data. *Food Policy*, 72:53–61.

Chen, X., Hong, H., and Tamer, E. (2005). Measurement error models with auxiliary data. *The Review of Economic Studies*, 72(2):343–366.

Devereux, P. J. (2004). Changes in relative wages and family labor supply. *Journal of Human Resources*, 39(3):698–722.

Duncan, G. J. and Hill, D. H. (1985). An investigation of the extent and consequences of measurement error in labor-economic survey data. *Journal of Labor Economics*, 3(4):508–532.

Fisher, K., Gershuny, J., Flood, S. M., Roman, J. G., and Hofferth, S. L. (2018). Multinational time use study extract system: Version 1.2. Technical report, Minneapolis, MN: IPUMS, https://doi.org/10.18128/D062.V1.2.

Frazis, H. and Stewart, J. (2012). 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.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.

Guryan, J., Hurst, E., and Kearney, M. (2008). Parental education and parental time with children. *Journal of Economic Perspectives*, 22(3):23–46.

Hamermesh, D. S., Frazis, H., and Stewart, J. (2005). Data watch: The american time use survey. *Journal of Economic Perspectives*, 19(1):221–232.

Heckman, J. J. (1993). What has been learned about labor supply in the past twenty years? *American Economic Review*, 83(2):116–121.

Hofferth, S. L., Flood, S. M., and Sobek, M. (2018). American time use survey data extract builder: Version 2.7. Technical report, College Park, MD: University of Maryland and Minneapolis, MN: IPUMS, https://doi.org/10.18128/D060.V2.7.

Hu, Y. and Sasaki, Y. (2015). Closed-form estimation of nonparametric models with non-classical measurement errors. *Journal of Econometrics*, 185(2):392–408.

Hu, Y. and Sasaki, Y. (2017). Identification of paired nonseparable measurement error models. *Econometric Theory*, 33(4):955–979.

Hu, Y. and Schennach, S. M. (2008). Instrumental variable treatment of nonclassical measurement error models. *Econometrica*, 76(1):195–216.

IPUMS (2020). Mtus samples characteristics. `https://www.mtusdata.org/mtus/samples.shtml`.

Juhn, C. and Murphy, K. M. (1997). Wage inequality and family labor supply. *Journal of Labor Economics*, 15(1, Part 1):72–97.

Kiviet, J. F. and Niemczyk, J. (2009). On the limiting and empirical distribution of iv estimators when some of the instruments are invalid. *UvA Econometrics Discussion Paper*, 2006/02.

Little, R. J. A. and Rubin, D. B. (2019). *Statistical Analysis with Missing Data*, volume 793. John Wiley & Sons.

Sousa, J. (2014). Estimation of price elasticities of demand for alcohol in the united kingdom. HMRC Working Paper 16.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

## Table 1: Simulations Based on the Dutch Time Use Survey (DTUS)

| Corr($\tilde{X}_i, U_i$) / Corr($\tilde{X}_i, \tilde{Z}_i$) | | Panel A: $n = 250$ | | | | | Panel B: $n = 500$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\hat{\beta}_{re}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{pool}$ | $\hat{\beta}_{day}$ | $\hat{\beta}_{re}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{pool}$ | $\hat{\beta}_{day}$ |
| 0 / 1 | MSE | 1.255 | 0.004 | 0.051 | 0.122 | 0.051 | 1.241 | 0.002 | 0.023 | 0.061 | 0.023 |
| | Bias$^2$ | 1.232 | 0.000 | 0.000 | 0.000 | 0.000 | 1.230 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.023 | 0.004 | 0.051 | 0.122 | 0.051 | 0.011 | 0.002 | 0.023 | 0.061 | 0.023 |
| 0.25 / 0.95 | MSE | 1.245 | 0.002 | 0.049 | 0.125 | 0.049 | 1.240 | 0.001 | 0.022 | 0.061 | 0.022 |
| | Bias$^2$ | 1.222 | 0.000 | 0.000 | 0.000 | 0.000 | 1.228 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.024 | 0.002 | 0.049 | 0.125 | 0.049 | 0.012 | 0.001 | 0.022 | 0.061 | 0.022 |
| 0.5 / 0.80 | MSE | 1.248 | 0.005 | 0.052 | 0.126 | 0.070 | 1.245 | 0.003 | 0.023 | 0.059 | 0.028 |
| | Bias$^2$ | 1.222 | 0.000 | 0.000 | 0.000 | 0.001 | 1.231 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.027 | 0.005 | 0.052 | 0.126 | 0.069 | 0.013 | 0.003 | 0.023 | 0.059 | 0.028 |
| 0.75/ 0.43 | MSE | 1.234 | 0.079 | 0.127 | 0.207 | 639.291 | 1.231 | 0.036 | 0.058 | 0.098 | 12.010 |
| | Bias$^2$ | 1.182 | 0.001 | 0.001 | 0.002 | 0.326 | 1.206 | 0.000 | 0.000 | 0.000 | 0.091 |
| | Var | 0.052 | 0.077 | 0.126 | 0.205 | 638.965 | 0.025 | 0.036 | 0.058 | 0.098 | 11.919 |
| Corr($\tilde{X}_i, U_i$) / Corr($\tilde{X}_i, \tilde{Z}_i$) | | Panel C: $n = 1000$ | | | | | Panel D: $n = 2500$ | | | | |
| | | $\hat{\beta}_{re}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{pool}$ | $\hat{\beta}_{day}$ | $\hat{\beta}_{re}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{pool}$ | $\hat{\beta}_{day}$ |
| 0 / 1 | MSE | 1.235 | 0.001 | 0.011 | 0.030 | 0.011 | 1.230 | 0.000 | 0.004 | 0.012 | 0.004 |
| | Bias$^2$ | 1.229 | 0.000 | 0.000 | 0.000 | 0.000 | 1.228 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.006 | 0.001 | 0.011 | 0.030 | 0.011 | 0.002 | 0.000 | 0.004 | 0.012 | 0.004 |
| 0.25 / 0.95 | MSE | 1.128 | 0.000 | 0.011 | 0.029 | 0.010 | 1.231 | 0.000 | 0.004 | 0.012 | 0.004 |
| | Bias$^2$ | 1.232 | 0.000 | 0.000 | 0.000 | 0.000 | 1.229 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.006 | 0.000 | 0.011 | 0.029 | 0.010 | 0.002 | 0.000 | 0.004 | 0.012 | 0.004 |
| 0.5 / 0.80 | MSE | 1.233 | 0.001 | 0.011 | 0.030 | 0.013 | 1.230 | 0.001 | 0.004 | 0.012 | 0.005 |
| | Bias$^2$ | 1.226 | 0.000 | 0.000 | 0.000 | 0.000 | 1.228 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.007 | 0.001 | 0.011 | 0.030 | 0.013 | 0.003 | 0.001 | 0.004 | 0.012 | 0.005 |
| 0.75/ 0.43 | MSE | 1.229 | 0.017 | 0.027 | 0.047 | 0.087 | 1.229 | 0.007 | 0.011 | 0.018 | 0.021 |
| | Bias$^2$ | 1.217 | 0.000 | 0.000 | 0.000 | 0.011 | 1.224 | 0.000 | 0.000 | 0.000 | 0.001 |
| | Var | 0.012 | 0.017 | 0.027 | 0.047 | 0.076 | 0.005 | 0.007 | 0.011 | 0.018 | 0.020 |

[1] This table compares finite sample performance of various estimators using the DTUS data. 10,000 random samples of different sizes are drawn from the original DTUS sample of 6,567 individual-year records.

[2] The two numbers in the first column represent: (i) correlation coefficient between regressor $\tilde{X}_i$ and error term $U_i$ (degree of endogeneity); (ii) correlation coefficient between regressor $\tilde{X}_i$ and IV $\tilde{Z}_i$ (strength of IV). Both are adjusted by changing the parameter $\rho$ in the simulation setup.

[3] $\hat{\beta}_{re}$ is the 2SLS estimator using the error-ridden recalled weekly hours worked in the DTUS. $\hat{\beta}_{re}$ exhibits large bias.

[4] $\hat{\beta}_{wk}$ is the 2SLS estimator given in eq. (5), which uses the accurate weekly hours worked in the DTUS and serves as an infeasible benchmark for the three estimators based on the ATUS. $\hat{\beta}_{wk}$ has virtually no bias and the smallest variance.

[5] For each individual in the DTUS, we randomly draw one from the seven days using the diary day sampling probabilities of the ATUS, thus obtained samples that imitate the ATUS, and we apply $\hat{\beta}_{im}$, $\hat{\beta}_{pool}$ and $\hat{\beta}_{day}$ to them in order to evaluate their performance.

[6] $\hat{\beta}_{im}$ has virtually no bias and the smallest variance among the three, followed by $\hat{\beta}_{pool}$.

[7] $\hat{\beta}_{day}$ is numerically equivalent to $\hat{\beta}_{im}$ when $\tilde{X}_i$ is exogenous. When $\tilde{X}_i$ is endogenous, however, $\hat{\beta}_{day}$ could display notable bias and enormous variance, especially when the sample size is smaller (and hence each day subsample is even smaller).

[8] $\tilde{\beta}_{day}$ introduced in Remark 11 performs almost identically to $\hat{\beta}_{im}$, but we do not report it here to avoid repetition.

Table 2: Weekly Labor Supply Elasticity Estimates (Hundredths): the DTUS

| | Married Men | | | Married Women | | |
|---|---|---|---|---|---|---|
| | $\hat{\beta}_{re}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{re}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ |
| $n$ of kids aged $< 18$ | 0.93 | 0.39 | 0.22 | 0.02 | $-16.81$ | $-21.02$ |
| | (0.41) | (0.58) | (1.16) | (0.80) | (1.72) | (3.34) |
| Educ: completed 2ndry | 2.14 | $-1.16$ | $-7.43$ | $-2.12$ | 11.88 | 9.79 |
| | (1.12) | (1.59) | (2.99) | (2.09) | (4.47) | (8.79) |
| Educ: above 2ndry | 4.13 | $-2.06$ | $-5.59$ | $-0.86$ | 22.68 | 21.53 |
| | (1.19) | (1.68) | (3.22) | (2.48) | (5.32) | (10.51) |
| P value of joint Hausman test | 0.00 | 0.11 | | 0.00 | 0.53 | |
| $n$ of Obs. | 1746 | 1746 | 1746 | 835 | 835 | 835 |
| $R$ squared[5] | 0.06 | 0.03 | 0.07 | 0.18 | 0.39 | 0.26 |

[1] The other control variables are age, age-squared, a dummy of working in private sector, an urban area dummy, and year dummies.

[2] $\hat{\beta}_{re}$ uses the recalled weekly hours; $\hat{\beta}_{wk}$ uses the true diary weekly hours; $\hat{\beta}_{im}$ uses the sample where only one day is randomly chosen for each individual using the ATUS diary day sampling weights.

[3] Standard errors are in parentheses.

[4] We conduct the joint Hausman tests (i.e., the coefficients associated with the three regressors in the table) regarding whether there are significant differences between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$, and between $\hat{\beta}_{wk}$ and $\hat{\beta}_{im}$, respectively.

[5] The $R$ squared for impute estimator is the average $R$ squared of the seven linear regression of daily hours worked $H_{it} = X_i'\beta_t + U_{it}$ for $t = 1, \ldots, 7$.

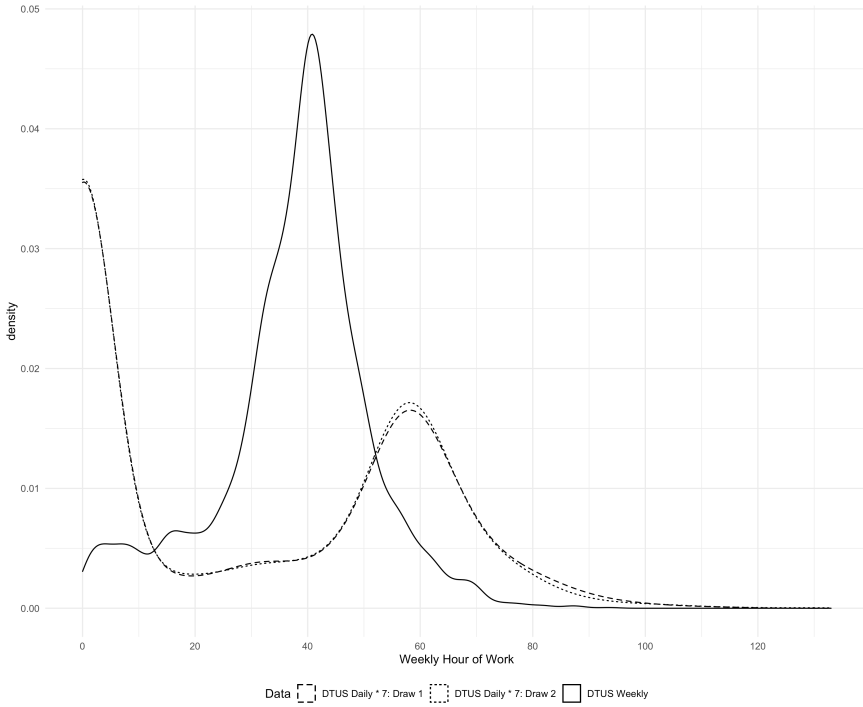
Table 3: Weekly Labor Supply Elasticity Estimates: the CPS and the ATUS

| Panel A: Mean and std dev of hours and wage | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| CPS Usual Weekly Hours Worked[1] | 39.63 | 38.42 | 32.50 | 35.52 |
| s.d. | (6.13) | (7.26) | (10.43) | (8.63) |
| ATUS Hours Worked on Diary Day | 4.70 | 4.74 | 3.56 | 4.18 |
| s.d. | (4.55) | (4.44) | (4.00) | (4.21) |
| ATUS Imputed Weekly Hours Worked | 41.27 | 40.38 | 31.96 | 36.18 |
| s.d. (lower bound)[2] | (9.57) | (9.79) | (9.26) | (9.68) |
| Hourly Wage (2017 US dollars) | 21.88 | 18.65 | 18.70 | 16.56 |

| Panel B: Elasticities (hundredths)[3] | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| Wage (CPS) | 5.39 | 11.38 | 15.89 | 11.72 |
| | (0.89) | (1.06) | (1.26) | (1.07) |
| Wage (ATUS) | 1.47 | 4.71 | 10.48 | 8.14 |
| | (3.36) | (3.25) | (3.32) | (3.30) |
| Spouse weekly earnings (CPS) | −0.19 | | −9.43 | |
| | (0.41) | | (0.77) | |
| Spouse weekly earnings (ATUS) | −3.47 | | −5.79 | |
| | (1.62) | | (2.12) | |
| Num. of kids age < 5 (CPS) | −0.80 | | −8.58 | |
| | (0.48) | | (0.82) | |
| Num. of kids age < 5 (ATUS) | −1.08 | | −8.97 | |
| | (1.92) | | (2.11) | |
| Num. of kids ages 5–18 (CPS) | −0.00 | | −2.87 | |
| | (0.26) | | (0.42) | |
| Num. of kids ages 5–18 (ATUS) | −0.44 | | −1.20 | |
| | (1.12) | | (1.18) | |
| $R$ squared (CPS) | 0.08 | 0.15 | 0.22 | 0.15 |
| $R$ squared (ATUS)[6] | 0.16 | 0.24 | 0.17 | 0.17 |
| $p$ value of joint Hausman test | 0.25 | 0.05 | 0.06 | 0.28 |
| $n$ of obs. | 3889 | 3816 | 5602 | 5731 |

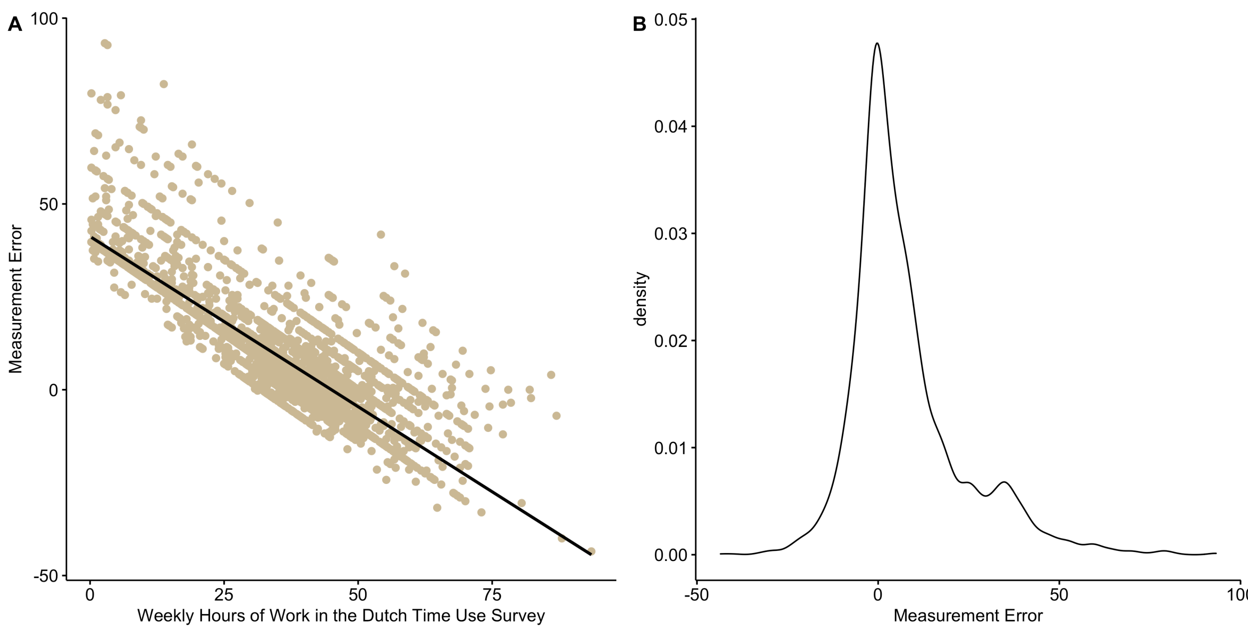
[1] This is the number of hours per week that the respondent usually works at his/her current job at the reported hourly wage rate.
[2] See footnote 47 in the paper for more details.
[3] The estimates based on the CPS recalled weekly hours are $\hat{\beta}_{re}$; the estimates based on the ATUS diary day hours are $\hat{\beta}_{im}$.
[4] The standard errors are in parentheses.
[5] The elasticities are evaluated at the respective mean hours worked in each data source.
[6] The $R$ squared for impute estimator is the average $R$ squared of the seven linear regression of daily hours worked $H_{it} = X_i'\beta_t + U_{it}$ for $t = 1, \ldots, 7$.
[7] For each sample group, we conduct joint Hausman tests regarding whether there are significant differences between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$.
[8] The other control variables are age, age-squared, two education dummies, eight Census division dummies, a metropolitan area dummy, race dummies, year dummies, occupation dummies and industry dummies.

Figure 1: DTUS Weekly Hours vs. Randomly Drawn Daily Hours $\times 7$




*Note*: The DTUS sample used here is pooled across the years 1985, 1990, 1995, 2000, and 2005. The sample includes only full-time workers aged between 25 and 54 at the time of interview. We used the default sample weight of the DTUS, which makes the weighted frequencies of the diaries within each age and sex group are evenly distributed in a week.

Figure 2: Measurement Errors in the DTUS Recalled Weekly Hours Worked




Panel A (left): scatter plot of the measurement errors in recalled weekly hours worked vs. the DTUS weekly hours worked. Panel B (right): kernel density of the measurement errors. In both, the measurement errors are obtained by subtracting the DTUS weekly hours worked from the recalled weekly hours worked for the same individuals.

# (NOT FOR PUBLICATION)

Supplementary Appendices for

*What Time Use Surveys Can (And Cannot) Tell Us About Labor Supply*

Cheng Chou[1] Ruoyao Shi[2]

January 2021

In Supplementary Appendix A, we report additional simulations, empirical analyses and robustness checks. In Supplementary Appendix B, we provide the proofs of the theorems and related results in Section 3.2 of our main paper, Chou and Shi (2020). In Supplementary Appendix C, we show the consequences of classical measurement errors in the ATUS.

## A Additional Simulations, Empirical Results and Robustness Checks

In this appendix, we show additional simulation results, additional empirical results and various robustness checks that complement our main paper, Chou and Shi (2020).

### A.1 Density Plots Based Only on Weekdays in the DTUS

In Figure 1 of the main paper, the ATUS-type daily hours exhibit bimodal distributions since most people work very little hours on weekends, if at all.[3] Figure A.1 shows the results of a similar experiment which takes the common five-day work schedule into account. We only keep those individuals whose diary days are the workdays, and then multiple their ATUS-type daily hours by 5. As is shown in Figure A.1, even though the DTUS weekly hours and the scaled ATUS-type daily hours have similar mode, their distributions differ notably, especially toward the left end. This again highlights the impossibility results in Section 3.1 of the main paper.

### A.2 Simulations Based Only on Weekdays in the DTUS

Table A.1 reports the results of simulation experiments that are very similar to those in Table 1. For Table A.1, we only use the daily hours worked in the DTUS for the weekdays. The regressors $X_i$

[1]Cheng Chou: School of Business, University of Leicester, UK. Email: cchou@le.ac.uk.

[2]Ruoyao Shi: Department of Economics, UC Riverside, USA. Email: ruoyao.shi@ucr.edu.

[3]According to the U.S. Bureau of Labor Statistics, in 2017, 89% of full-time workers worked on an average weekday, compared with 32.6% on an average weekend day.

and the IVs $Z_i$ are generated from the $n \times 5$ matrix with elements $H_{it}^{DTUS}$ $(t = 2, \ldots, 6)$, denoted by $H^{DTUS,5}$, using the same design described in Section 4.1. To generate fictitious ATUS-type samples, we randomly choose only one day from Monday to Friday for each individual using equal sampling weights.

Just like in Table 1, the week estimator $\hat{\beta}_{wk}$ is our infeasible benchmark, which has virtually no biases and the smallest variances. The efficiency gain of the impute estimator $\hat{\beta}_{im}$ relative to the pool estimator $\hat{\beta}_{pool}$ and the day estimator $\hat{\beta}_{day}$ becomes less pronounced. This is likely due to the fact that the first principal component of $H^{DTUS}$ captures the dichotomy between weekdays and weekends, and once that is removed, the daily variation of hours worked drops dramatically.[4] Besides, the ATUS assigns equal sampling weights to the weekdays. As we explained in Remark 7 in Chou and Shi (2020), if $H_{i2} = \cdots = H_{i6}$ and $r_2 = \cdots = r_6$, then $\Omega_{pool-im} = 0$ and there will be no difference in the asymptotic efficiency between $\hat{\beta}_{im}$ and $\hat{\beta}_{pool}$. Our additional simulation results here verify our theoretical prediction in the main paper.

## A.3 Coefficient Estimates in the DTUS Weekly Labor Supply Regression

In Table 2 of the main paper, we report the weekly labor supply elasticity estimates using the DTUS. Table A.2 reports the coefficient estimates in the weekly labor supply regression equation shown in eq. (4), and the elasticity estimates reported in Table 2 are evaluated at the sample mean hours.

## A.4 Coefficient Estimates in the ATUS Weekly Labor Supply Regression

In Table 3 of the main paper, we report the weekly labor supply elasticity estimates using the ATUS. Table A.5 reports the coefficient estimates in the weekly labor supply regression equation shown in eq. (22), and the elasticity estimates reported in Table 3 are evaluated at respective sample means based on these coefficients and the sample mean hours.

[4] Indeed, the first principal component of $H^{DTUS,5}$ assigns the weights $\beta_1 = 0.4389$, $\beta_2 = 0.4560$, $\beta_3 = 0.4580$, $\beta_4 = 0.4531$ and $\beta_5 = 0.4294$ to its columns, which correspond to Monday to Friday, respectively; i.e., each weekday contributes roughly equally to the first principal component.

## A.5 Representativeness of the ATUS Sample

The ATUS is designed to be a random subsample of those who recently complete their participation in the CPS. We compare the ATUS sample against the CPS sample. Sample means and sample standard deviations of the key variables used in the empirical studies are reported in Table A.3. The ATUS sample (first column) is the one used in the empirical studies in our main paper. The CPS sample (middle column) is the entire CPS 2003-2017 sample after the same sample selection criterion (hourly paid workers aged between of 25 and 54, whose wage rate is positive, and spouse earnings (if married) and total usual weekly hours worked at all jobs reported in the CPS are observed The entire CPS sample (last column) includes the respondents whose hourly wage or spouse weekly earnings is missing. None of the key variable summary statistics differ significantly among the three samples.

The elasticity estimates in Table 3 of the main paper are based on the sample in the first column of Table A.3. Using the sample of second column of Table A.3, we estimate the labor supply elasticities similar to the main paper. We report such estimates in Table A.4. Comparing them with the CPS results in Table 3 in the main paper, we find no notable differences.

Therefore, it is safe to conclude that the ATUS sample is a representative subsample of the CPS, which implies that the differences between the ATUS and the CPS elasticity estimates are more likely due to the nonclassical measurement errors in the CPS than due to the composition of the ATUS sample.

Moreover, the ATUS sample does not exhibit strong seasonal fluctuations over a year, whether as a whole or within each occupation. In Table A.6, we categorize the ATUS sample into different occupations and months. First, the entire ATUS sample is very balanced over a year, with people surveyed in all months having roughly equal proportions. Second, within each occupation, the ATUS also surveys approximately same numbers of people in every month. Third, among the nine occupation categories, not a single occupation bears overwhelming weights. So the empirical results in the main paper are not likely to be driven by anomaly in a single occupation or a single month.

### A.6 Robustness Checks of the Empirical Results in Section 5

In Section 5 of the main paper, we estimate labor supply elasticities using the ATUS daily hours and compare the estimates with those obtained using the CPS recalled weekly hours. The ATUS estimates reported in Table 3 of the main paper uses the "work" hours on all jobs (activity code: 050100) for all the occupations in the ATUS.

In this section, we conduct four robustness checks. The first robustness check, reported in Table A.7, restricts to the three occupations with the most observations; they are computer and mathematical science, healthcare support, and office and administrative support occupations. The second robustness check, reported in Table A.8, uses "work" and "work-related" hours (activity codes: 050100 and 050200) for all the occupations in the ATUS.[5] The third robustness check, reported in Table A.9, estimates the elasticities using the OLS, without correcting the potential measurement issues in own hourly wage and spouse weekly earnings (using their respective decile as IVs). Comparing Tables A.7 to A.9 here with Table 3 of the main paper, we see that none of the estimates change much, neither qualitatively nor quantitatively.

The fourth robustness check, reported in Table A.10, uses survey year-month group indicators as IVs.[6] Angrist (1991) proposes the use of group classification variable that is independent from the error term as IV. He also proves that the resulting 2SLS estimator is a generalization of the Wald estimator in the treatment effect literature that is frequently used in binary treatment and binary IV cases. The identification power of such 2SLS estimators comes from the variation in group means, and it requires that the individual deviation from group means to be uncorrelated with the IVs. Since we have no reason to believe that the error term in the weekly labor supply eq. (4) is systematically correlated with survey year or survey month, the survey year-month dummies satisfy the exclusion restriction. On the other hand, the correlation between survey year (or survey month) and log wage (or spouse earnings) is probably weak, which may lead to inflated standard errors and sizable finite sample bias. Compare Table A.10 with Table 3 in the main paper, the standard errors of the elasticity estimates (Panel B) rise remarkably. Among those elasticity estimates which remain significant – CPS own wage for all groups, CPS spouse earning and older kids for married

[5] Examples of work-related activities here include attending social events, attending sporting events, and eating or drinking with bosses, co-workers or clients, etc.

[6] Our sample contains respondents in 15 years (2003-2017), which together with 12 months result in 180 group indicators.

women, CPS and ATUS younger kids for married women – neither sign nor magnitude changes much. This shows that our labor supply elasticity estimates are not very sensitive to the choice of IVs.

# B Proofs of the Theorems in Section 3.2

*Proof of Theorem 1.* First we show the identification of $\beta$ if $H_i^w$ were observed, as it will be instructive for our discussion based on the ATUS data $H_i^{ATUS}$. If the true weekly hours worked $H_i^w$ were observed, then the identification of the $p$-dimensional parameter vector $\beta$ is just the usual argument for 2SLS (i.e., generalized method of moments) estimators. Formally, $\beta$ is identified if the following $q$-dimensional moment conditions

$$E(Z_iU_i) = E[Z_i(H_i^w - X_i'\beta)] = 0 \Longleftrightarrow E(Z_iH_i^w) = E(Z_iX_i')\beta \tag{B.1}$$

have a unique solution of $\beta$, which is true if $q \geq p$, and the rank of the $q \times p$ matrix $E(Z_iX_i')$ is $p$ (i.e., Assumption 3). Provided that $E(Z_iZ_i')$ is nonsingular (part of Assumption 3), eq. (B.1) is equivalent to

$$E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iH_i^w) = E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iX_i')\beta, \tag{B.2}$$

and

$$\beta = \left(E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iX_i')\right)^{-1} E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iH_i^w) \tag{B.3}$$

is *the unique solution* of eq. (B.2). $\hat{\beta}_{wk}$ is to replace the expectations in eq. (B.3) by respective sample means.

Next we consider the case where only $H_i^{ATUS} = \sum_{t=1}^{7} d_{it}H_{it}$ is observed. The identification of $\beta$ is still based on the same moment conditions in eq. (B.1), but the only problem now is that the ATUS data are not informative about the term $E(Z_iH_i^w)$ in eq. (B.3). Since the expression of $\beta$ in eq. (B.3) is the *unique solution* of eq. (B.2), the identification of $\beta$ will be proved if we can find equivalent expressions of eq. (B.3) that have sample counterparts in the ATUS data. The rest of our proof shows that. Under the potential outcome framework, we have

$$
\begin{aligned}
\beta &= \left(E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iX_i')\right)^{-1} E(X_iZ_i')[E(Z_iZ_i')]^{-1}\sum_{t=1}^{7}E(Z_iH_{it}) && \text{(B.4)}\\
&= \left(E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iX_i')\right)^{-1} E(X_iZ_i')\sum_{t=1}^{7}[E(Z_iZ_i'|d_{it}=1)]^{-1}E(Z_iH_{it}|d_{it}=1) && \text{(B.5)}\\
&= \left(E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iX_i')\right)^{-1} E(X_iZ_i')[E(Z_iZ_i')]^{-1}\sum_{t=1}^{7}E(r_{nt}d_{it})E(Z_iH_{it})\\
&= \left(E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iX_i')\right)^{-1} E(X_iZ_i')[E(Z_iZ_i')]^{-1}\sum_{t=1}^{7}E(r_{nt}d_{it}Z_iH_{it})\\
&= \left(E(X_iZ_i')[E(Z_iZ_i')]^{-1}E(Z_iX_i')\right)^{-1} E(X_iZ_i')[E(Z_iZ_i')]^{-1}\sum_{t=1}^{7}E(r_{nt}Z_iH_{it}|d_{it}=1) && \text{(B.6)}\\
&= \sum_{t=1}^{7}\left(E(X_iZ_i'|d_{it}=1)[E(Z_iZ_i'|d_{it}=1)]^{-1}E(Z_iX_i'|d_{it}=1)\right)^{-1}\\
&\qquad \times E(X_iZ_i'|d_{it}=1)[E(Z_iZ_i'|d_{it}=1)]^{-1}E(Z_iH_{it}|d_{it}=1), && \text{(B.7)}
\end{aligned}
$$

where eq. (B.4) holds by the definition of $H_i^w$, eqs. (B.5) to (B.7) hold by Assumption 1 and that $E(r_{nt}d_{it}) = 1$. Equation (B.5) is the population counterpart of $\hat{\beta}_{im}$, eq. (B.6) is the population counterpart of $\hat{\beta}_{pool}$, and eq. (B.7) is the population counterpart of $\hat{\beta}_{day}$, all of which are now estimable using the ATUS data. □

*Proof of Theorem 2.* First, we show the consistency of $\hat{\beta}_{wk}$:

$$\hat{\beta}_{wk} - \beta = A_n^{-1}X'P_zU = A_n^{-1}B_nC_n^{-1}(Z'U/n) \xrightarrow{p.} A^{-1}BC^{-1}E(Z_iU_i) = 0.$$

In fact, this is a standard result for instrumental variable estimators.

Second, we show the consistency of $\hat{\beta}_{im}$. Consider the difference $(\hat{\beta}_{im} - \hat{\beta}_{wk})$ using their definitions:

$$
\begin{aligned}
\hat{\beta}_{im} - \hat{\beta}_{wk} &= (X'P_zX)^{-1}X'P_z\left[\sum_{t=1}^{7}Z(Z'D_tZ)^{-1}Z'D_tH_t - H^w\right]\\
&= (X'P_zX)^{-1}X'P_z\left[\sum_{t=1}^{7}Z(Z'D_tZ)^{-1}Z'D_tH_t - P_z\sum_{t=1}^{7}H_t\right]
\end{aligned}
$$

$$= \sum_{t=1}^{7} (X'P_zX)^{-1}X'P_zZ[(Z'D_tZ)^{-1}Z'D_tH_t - (Z'Z)^{-1}Z'H_t]$$
$$= \sum_{t=1}^{7} (X'P_zX)^{-1}X'Z[(Z'D_tZ)^{-1}Z'D_tH_t - (Z'Z)^{-1}Z'H_t].$$

Using the linear projection eq. (10), we have

$$\hat{\beta}_{im} - \hat{\beta}_{wk} = \sum_{t=1}^{7} A_n^{-1}B_n \left[ \left( \frac{1}{n_t} Z'D_tZ \right)^{-1} \frac{1}{n_t} Z'D_tV_t - \left( \frac{1}{n} Z'Z \right)^{-1} \frac{1}{n} Z'V_t \right]. \qquad \text{(B.8)}$$

Define

$$C_{n_t} = Z'D_tZ/n_t.$$

Following from the law of large numbers, $A$, $B$ and $C$ are the probability limit of $A_n$, $B_n$, and $C_n$ (also $C_{n_t}$) as $n \to \infty$, respectively. By the definition of $A_n$, $B_n$, $C_n$ and $C_{n_t}$, we have

$$\begin{aligned} \hat{\beta}_{im} - \hat{\beta}_{wk} &= \sum_{t=1}^{7} A_n^{-1}B_n \left[ C_{n_t}^{-1} \frac{1}{n_t} Z'D_tV_t - C_n^{-1} \frac{1}{n} Z'V_t \right] \\ &\xrightarrow{p.} \sum_{t=1}^{7} A^{-1}BC^{-1}[E(Z_i d_{it} V_{it}) - E(Z_i V_{it})] \\ &= \sum_{t=1}^{7} A^{-1}BC^{-1}[E(Z_i V_{it})E(d_{it}) - E(Z_i V_{it})] \\ &= 0, \end{aligned} \qquad \text{(B.9)}$$

because $E(Z_iV_{it}) = 0$. Since $\hat{\beta}_{wk} \xrightarrow{p.} \beta$ and $\hat{\beta}_{im} - \hat{\beta}_{wk} \xrightarrow{p.} 0$, we conclude that $\hat{\beta}_{im} \xrightarrow{p.} \beta$.

Third, we show the consistency of $\hat{\beta}_{pool}$. By the definition of $A_n$, $B_n$, $C_n$ and $C_{n_t}$, we have

$$\begin{aligned}
\hat{\beta}_{pool} - \hat{\beta}_{wk} &= \sum_{t=1}^{7} A_n^{-1} B_n C_n^{-1} \frac{Z'(r_{nt} D_t - I) H_t}{n} \\
&\xrightarrow{p.} A^{-1} B C^{-1} \sum_{t=1}^{7} \frac{Z'(r_t D_t - I) H_t}{n} \\
&\xrightarrow{p.} A^{-1} B C^{-1} \sum_{t=1}^{7} E((r_t d_{it} - 1) Z_i H_{it}) \\
&= A^{-1} B C^{-1} \sum_{t=1}^{7} E(r_t d_{it} - 1) E(Z_i H_{it}) \\
&= 0,
\end{aligned} \tag{B.10}$$

where the second line holds because $r_{nt} \xrightarrow{p.} r_t$, and the last equality holds since $E(r_t d_{it} - 1) = 0$. Combined with the result that $\hat{\beta}_{wk} \xrightarrow{p.} \beta$, this implies that $\hat{\beta}_{pool} \xrightarrow{p.} \beta$.

Fourth, we show the consistency of $\hat{\beta}_{day}$. The weekly labor supply equation in eq. (4) can be re-written as the sum of seven daily labor supply equations in eq. (7), with

$$\beta = \sum_{t=1}^{7} \beta_t \qquad \text{and} \qquad U_i = \sum_{t=1}^{7} U_{it}.$$

We then can re-write the day estimator as

$$\begin{aligned}
\hat{\beta}_{day} &= \sum_{t=1}^{7} (X' P_{zt} X)^{-1} X' P_{zt} H_t \\
&= \sum_{t=1}^{7} (X' P_{zt} X)^{-1} X' P_{zt} (X \beta_t + U_t) \\
&= \sum_{t=1}^{7} \beta_t + \sum_{t=1}^{7} (X' P_{zt} X)^{-1} X' P_{zt} U_t \\
&= \beta + \sum_{t=1}^{7} (X' P_{zt} X)^{-1} X' P_{zt} U_t.
\end{aligned} \tag{B.11}$$

Simply by the law of large numbers, continuous mapping theorem, and the definition of $P_{zt}$, we have

$$\begin{aligned}
\hat{\beta}_{day} - \beta &= \sum_{t=1}^{7} (X' P_{zt} X)^{-1} X' P_{zt} U_t \\
&= \sum_{t=1}^{7} \left( \frac{X' P_{zt} X}{n_t} \right)^{-1} \frac{X' D_t Z}{n_t} \left( \frac{Z' D_t Z}{n_t} \right)^{-1} \frac{Z' D_t U_t}{n_t} \\
&\xrightarrow{p.} \sum_{t=1}^{7} A^{-1} B C^{-1} E(Z_i U_{it}) \\
&= A^{-1} B C^{-1} E\Big[ Z_i \sum_{t=1}^{7} U_{it} \Big] \\
&= A^{-1} B C^{-1} E(Z_i U_i) \\
&= 0.
\end{aligned} \tag{B.12}$$

This completes the proof. □

*Proof of Theorem 3.* We have

$$\sqrt{n}(\hat{\beta}_{wk} - \beta) = A^{-1} \frac{1}{\sqrt{n}} X' P_z U + o_p(1),$$

which is asymptotically normal with mean zero and variance

$$\Omega_{wk} = A^{-1} B C^{-1} E(U_i^2 Z_i Z_i') C^{-1} B' A^{-1},$$

This completes the proof of Theorem 3. Again, this is a standard result for instrumental variable estimators. □

*Proof of Theorem 4.* To show (i), we consider the decomposition

$$\sqrt{n}(\hat{\beta}_{im} - \beta) = \sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk}) + \sqrt{n}(\hat{\beta}_{wk} - \beta).$$

Since the asymptotic variance of $\sqrt{n}(\hat{\beta}_{wk} - \beta)$ is given by Theorem 3, the key to finding the asymptotic distribution of $\sqrt{n}(\hat{\beta}_{im} - \beta)$ is therefore to compute the asymptotic variance of $\sqrt{n}(\hat{\beta}_{im} -$

$\hat{\beta}_{wk})$ and $\sqrt{n}(\hat{\beta}_{wk} - \beta)$, as well as their asymptotic covariance. Recall that eq. (B.8) implies

$$\begin{aligned}\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk}) &= \sum_{t=1}^{7} A_n^{-1} B_n \sqrt{n} \left[ \left( \frac{1}{n_t} Z' D_t Z \right)^{-1} \frac{n}{n_t} \frac{1}{n} Z' D_t V_t - \left( \frac{1}{n} Z' Z \right)^{-1} \frac{1}{n} Z' V_t \right] \\ &= \sum_{t=1}^{7} A_n^{-1} B_n \left[ C_{n_t}^{-1} r_{nt} \frac{1}{\sqrt{n}} Z' D_t V_t - C_n^{-1} \frac{1}{\sqrt{n}} Z' V_t \right].\end{aligned} \tag{B.13}$$

Because $n^{-1/2} Z' D_t V_t = O_p(1)$ and $n^{-1/2} Z' V_t = O_p(1)$, we have

$$\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk}) = A^{-1} B C^{-1} \sum_{t=1}^{7} \frac{1}{\sqrt{n}} Z' (r_t D_t - I_n) V_t + o_p(1). \tag{B.14}$$

The key is then the asymptotic distribution of

$$\sum_{t=1}^{7} \frac{1}{\sqrt{n}} Z' (r_t D_t - I_n) V_t = \sum_{t=1}^{7} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (r_t d_{it} - 1) Z_i V_{it}.$$

Because $d_{it} \perp\!\!\!\perp (Z, H_t)$ and $E(r_t d_{it} - 1) = 0$, we have that $\mathrm{E}[(r_t d_{it} - 1) Z_i V_{it}] = 0$. Moreover, we have

$$\mathrm{E}\big[(r_t d_{it} - 1) Z_i V_{it} V_{i\tau} Z_i' (r_\tau d_{i\tau} - 1)\big] = \mathrm{E}[(r_t d_{it} - 1)(r_\tau d_{i\tau} - 1)] E\big(Z_i V_{it} V_{i\tau} Z_i'\big).$$

It can be shown that

$$\mathrm{E}[(r_t d_{it} - 1)(r_\tau d_{i\tau} - 1)] = \begin{cases} r_t - 1, & t = \tau, \\ -1, & t \neq \tau. \end{cases} \tag{B.15}$$

We hence have

$$\mathrm{Var}((r_t d_{it} - 1) Z_i V_{it}) = (r_t - 1) E(Z_i V_{it} V_{it} Z_i'),$$

and for $t \neq \tau$,

$$\mathrm{Cov}((r_t d_{it} - 1) Z_i V_{it}, (r_\tau d_{i\tau} - 1) Z_i V_{i\tau}) = -E(Z_i V_{it} V_{i\tau} Z_i').$$

From eq. (B.14), we conclude that $\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk})$ is asymptotically normal with mean zero and variance

$$\Omega_{im-wk} \equiv A^{-1}BC^{-1}\Big[\sum_{t=1}^{7}(r_t-1)E(Z_iV_{it}V_{it}Z_i') - 2\sum_{1\leq t<\tau\leq 7}E(Z_iV_{it}V_{i\tau}Z_i')\Big]C^{-1}B'A^{-1};$$

We then proceed to compute the covariance between $\sqrt{n}(\hat{\beta}_{im}-\hat{\beta}_{wk})$ and $\sqrt{n}(\hat{\beta}_{wk}-\beta)$. Note that we have shown $E\Big(\sqrt{n}(\hat{\beta}_{im}-\hat{\beta}_{wk})\Big) = o_p(1)$ and $E\Big(\sqrt{n}(\hat{\beta}_{wk}-\beta)\Big) = o_p(1)$. In addition, we have

$$\begin{aligned}
&E\Big(\sqrt{n}(\hat{\beta}_{im}-\hat{\beta}_{wk})\sqrt{n}(\hat{\beta}_{wk}-\beta)\Big)\\
&= A^{-1}BC^{-1}E\left(\sum_{t=1}^{7}n^{-1}Z'(r_tD_t-I_n)V_tU'P_zX\right)A^{-1}+o_p(1)\\
&= A^{-1}BC^{-1}\sum_{t=1}^{7}E\big(n^{-1}Z'(r_tD_t-I_n)V_tU'P_zX\big)A^{-1}+o_p(1)\\
&= A^{-1}BC^{-1}\sum_{t=1}^{7}E\big(n^{-1}Z'E((r_tD_t-I_n)V_tU'P_zX \mid Z)\big)A^{-1}+o_p(1)\\
&= A^{-1}BC^{-1}\sum_{t=1}^{7}E\big(n^{-1}Z'E(r_tD_t-I_n)E(V_tU'P_zX \mid Z)\big)A^{-1}+o_p(1),
\end{aligned}$$

where the last equality holds because the diary day is completely random, i.e., $d_{it}$ (and hence $D_t$) is independent from everything else. This, combined with

$$E(r_tD_t-I_n)=0$$

implies

$$E\Big(\sqrt{n}(\hat{\beta}_{im}-\hat{\beta}_{wk})\sqrt{n}(\hat{\beta}_{wk}-\beta)\Big) = o_p(1).$$

As a result,

$$\mathrm{Cov}\left(\sqrt{n}(\hat{\beta}_{im}-\hat{\beta}_{wk}),\sqrt{n}(\hat{\beta}_{wk}-\beta)\right) = o_p(1).$$

We conclude that the asymptotic variance of the impute estimator equals

$$\Omega_{im} = \Omega_{wk} + \Omega_{im-wk},$$

This completes the proof of (i).

To show (ii), we follow similar steps as for (i). We decompose

$$\sqrt{n}(\hat{\beta}_{pool} - \beta) = \sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im}) + \sqrt{n}(\hat{\beta}_{im} - \beta),$$

where we only need to find the asymptotic variance of $\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im})$ and the asymptotic covariance between the two terms. First, we have

$$\begin{aligned}\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im}) &= \sqrt{n}(X'P_zX)^{-1}X'Z\sum_{t=1}^{7}\big[(Z'Z)^{-1}r_{nt}Z'D_tH_t - (Z'D_tZ)^{-1}Z'D_tH_t\big] \\ &= A_n^{-1}B_n\sum_{t=1}^{7}(C_n^{-1} - C_{n_t}^{-1})\frac{1}{\sqrt{n}}r_{nt}Z'D_tH_t.\end{aligned}$$

In light of the linear projection eq. (10) of $H_t$, we have

$$\begin{aligned}\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im}) &= A_n^{-1}B_n\sum_{t=1}^{7}(C_n^{-1} - C_{n_t}^{-1})\frac{1}{\sqrt{n}}r_{nt}Z'D_t\big(Z\alpha_t + V_t\big) \\ &= A_n^{-1}B_n\sum_{t=1}^{7}(C_n^{-1} - C_{n_t}^{-1})\frac{1}{\sqrt{n}}r_{nt}Z'D_tZ\alpha_t + o_p(1) \\ &= A_n^{-1}B_n\sum_{t=1}^{7}\left(C_n^{-1}\frac{1}{\sqrt{n}}Z'r_{nt}D_tZ\alpha_t - \sqrt{n}\alpha_t\right) + o_p(1) \\ &= A_n^{-1}B_n\sum_{t=1}^{7}\left(C_n^{-1}\frac{1}{\sqrt{n}}Z'r_{nt}D_tZ\alpha_t - \sqrt{n}C_n^{-1}\frac{Z'Z}{n}\alpha_t\right) + o_p(1) \\ &= A_n^{-1}B_nC_n^{-1}\sum_{t=1}^{7}\left(\frac{1}{\sqrt{n}}Z'r_{nt}D_tZ\alpha_t - \frac{1}{\sqrt{n}}Z'Z\alpha_t\right) + o_p(1) \\ &= A^{-1}BC^{-1}\sum_{t=1}^{7}\frac{1}{\sqrt{n}}Z'(r_tD_t - I_n)Z\alpha_t + o_p(1), \qquad \text{(B.16)}\end{aligned}$$

where the second equality holds since $C_n^{-1} - C_{n_t}^{-1} = o_p(1)$, $n^{-1/2}r_{nt}Z'D_tV_t = O_p(1)$, and $C_{n_t}^{-1}Z'D_tZ/n_t = I_n$, and the last equality holds by the definition of $C_n$ and $C_{n_t}$. It follows straightforward that $\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im})$ is asymptotically normal with some asymptotic variance $\Omega_{pool-im}$. To calculate $\Omega_{pool-im}$, let

$$\delta_{it} = (r_td_{it} - 1)Z_i\alpha_t'Z_i,$$

and rewrite

$$\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im}) = A^{-1}BC^{-1}\sum_{t=1}^{7}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\delta_{it} + o_p(1).$$

Using eq. (B.15), we can show that

$$\mathrm{Var}(\delta_{it}) = (r_t - 1)E(Z_i\alpha_t' Z_i Z_i' \alpha_t Z_i'),$$

and

$$\mathrm{Cov}(\delta_{it}, \delta_{i\tau}) = -E(Z_i\alpha_t' Z_i Z_i' \alpha_\tau' Z_i').$$

As a result,

$$\Omega_{pool-im} = A^{-1}BC^{-1}\Big[\sum_{t=1}^{7}(r_t - 1)E(Z_i\alpha_t' Z_i Z_i' \alpha_t Z_i') - 2\sum_{1\le t<\tau\le 7}E(Z_i\alpha_t' Z_i Z_i' \alpha_\tau' Z_i')\Big]C^{-1}B'A^{-1}. \tag{B.17}$$

Second, we consider the asymptotic covariance between $\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im})$ and $\sqrt{n}(\hat{\beta}_{im} - \beta)$. By the definition of $V_{i\tau}$ in the linear projection eq. (10), $Z_i$ and $V_{i\tau}$ ($\tau = 1, \ldots, 7$) are orthogonal with each other. This implies that for any $1 \le t \le \tau \le 7$,

$$\mathrm{Cov}((r_t d_{it} - 1)Z_i\alpha_t' Z_i, (r_\tau d_{i\tau} - 1)Z_i V_{i\tau}) = 0.$$

This further implies that $\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im})$ and $\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk})$ are asymptotically uncorrelated. Furthermore, using the same argument as in the proof of (i), one can show that $\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im})$ and $\sqrt{n}(\hat{\beta}_{wk} - \beta)$ are asymptotically uncorrelated. Together they imply that $\sqrt{n}(\hat{\beta}_{pool} - \hat{\beta}_{im})$ and $\sqrt{n}(\hat{\beta}_{im} - \beta)$ are asymptotically uncorrelated.

To summarize, we have shown that the asymptotic variance of $\sqrt{n}(\hat{\beta}_{pool} - \beta)$ equals to

$$\Omega_{pool} = \Omega_{pool-im} + \Omega_{im}.$$

Note that since $\Omega_{pool}$ is positive definite, it implies that $\hat{\beta}_{im}$ is asymptotically more efficient than $\hat{\beta}_{pool}$. This completes the proof of (ii).

Part (iii) follows from writing $\text{Var}(\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk}))$ as the following sum,

$$\text{Var}(\sqrt{n}(\hat{\beta}_{im} - \beta)) + \text{Var}(\sqrt{n}(\hat{\beta}_{wk} - \beta)) - 2\,\text{Cov}(\sqrt{n}(\hat{\beta}_{im} - \beta), \sqrt{n}(\hat{\beta}_{wk} - \beta)).$$

Because we have shown $E(\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk})\sqrt{n}(\hat{\beta}_{wk} - \beta)) = o_p(1)$, we have that

$$E(\sqrt{n}(\hat{\beta}_{im} - \beta)\sqrt{n}(\hat{\beta}_{wk} - \beta)) = \text{Var}(\sqrt{n}(\hat{\beta}_{wk} - \beta)) + o_p(1).$$

We hence conclude that $\text{Var}(\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk})) = \text{Var}(\sqrt{n}(\hat{\beta}_{im} - \beta)) - \text{Var}(\sqrt{n}(\hat{\beta}_{wk} - \beta))$. The rest of part (iii) follows immediately. □

*Proof of Theorem 5.* To prove (i), first note that by the definition of $U_i$ and the "$H$ first stage", we have

$$U_i \equiv H_i^w - X_i'\beta = \sum_{t=1}^{7} H_{it} - X_i'\beta = \sum_{t=1}^{7}(Z_i'\alpha_t + V_{it}) - X_i'\beta = \sum_{t=1}^{7} V_{it} + Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta. \quad \text{(B.18)}$$

Therefore, we have

$$\begin{aligned} E(U_i^2 Z_i Z_i') =& E\left[\left(\sum_{t=1}^{7} V_{it}\right)^2 Z_i Z_i'\right] + E\left[\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right)^2 Z_i Z_i'\right] \\ &+ 2E\left[\left(\sum_{t=1}^{7} V_{it}\right)\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right) Z_i Z_i'\right] \\ =& \sum_{t=1}^{7} E(V_{it}^2 Z_i Z_i') + 2\sum_{1\le t<\tau\le 7} E(V_{it}V_{i\tau}Z_i Z_i') \\ &+ E\left[\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right)^2 Z_i Z_i'\right] + 2E\left[\left(\sum_{t=1}^{7} V_{it}\right)\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right) Z_i Z_i'\right]. \end{aligned} \quad \text{(B.19)}$$

We can then replace $E(U_i^2 Z_i Z_i')$ in the middle of $\Omega_{wk}$ in eq. (11) by eq. (B.19). Part (i) follows by adding $\Omega_{wk}$ and $\Omega_{im-wk}$ together, which are given in eq. (11) and eq. (12), respectively. Since $\Omega_{im-wk}$ involves terms like $E(Z_i V_{it} V_{i\tau} Z_i')$, it may seem at a glance that $\Omega_{im}$ depends on the correlations among $V_{it}$ and $V_{i\tau}$ for $t \neq \tau$. But the proof here shows that these terms from $\Omega_{wk}$ and $\Omega_{im-wk}$ cancel with each other.

Part (ii) can be proven by the same argument as for part (i), i.e., by expanding the term $E\left[\left(Z_i'\sum_{t=1}^7 \alpha_t - X_i'\beta\right)^2 Z_iZ_i'\right]$ in $\Omega_{im}$ and adding it together with $\Omega_{pool-im}$ in eq. (13). $\square$

*Proof of Theorem 6.* Part (i). For every $t = 1, \dots, 7$, it follows from a standard result for instrumental variable estimators that

$$\sqrt{n_t}(\hat{\beta}_t - \beta_t) \xrightarrow{d.} N(0, A^{-1}BC^{-1}E(U_{it}^2 Z_i Z_i')C^{-1}B'A^{-1}),$$

which implies that if we normalize by $\sqrt{n}$ instead of $\sqrt{n_t}$, we have

$$\sqrt{n}(\hat{\beta}_t - \beta_t) \xrightarrow{d.} N(0, r_t A^{-1}BC^{-1}E(U_{it}^2 Z_i Z_i')C^{-1}B'A^{-1}).$$

Moreover, note that $\hat{\beta}_t$ only uses the data on those individuals whose diary day is $t$. Since the individuals are drawn independently, $\hat{\beta}_t$ is independent of $\hat{\beta}_\tau$ for any $t \neq \tau$. This implies that the asymptotic variance of the day estimator $\hat{\beta}_{day}$ is

$$\Omega_{day} = A^{-1}BC^{-1}\left[\sum_{t=1}^{7} r_t E(U_{it}^2 Z_i Z_i')\right] C^{-1}B'A^{-1}.$$

This proves eq. (16).

To prove part (ii), we first derive an alternative expression for $\Omega_{day}$. Similar to eq. (B.18), we can decompose $U_{it}$ in a similar manner:

$$U_{it} \equiv H_{it} - X_i'\beta_t = V_{it} + \left(Z_i'\alpha_t - X_i'\beta_t\right),$$

which implies that

$$E(U_{it}^2 Z_i Z_i') = E(V_{it}^2 Z_i Z_i') + E\left[\left(Z_i'\alpha_t - X_i'\beta_t\right)^2 Z_iZ_i'\right] + 2E\left[V_{it}\left(Z_i'\alpha_t - X_i'\beta_t\right) Z_iZ_i'\right],$$

which combined with eq. (16) in turn implies that

$$\Omega_{day} = A^{-1}BC^{-1}\left\{\sum_{t=1}^{7} r_t E(V_{it}^2 Z_i Z_i') + \sum_{t=1}^{7} r_t E\left[\left(Z_i'\alpha_t - X_i'\beta_t\right)^2 Z_iZ_i'\right]\right.$$

$$+2\sum_{t=1}^{7} r_t E\left[V_{it}\left(Z_i'\alpha_t - X_i'\beta_t\right) Z_i Z_i'\right]\Bigg\} C^{-1}B'A^{-1}. \tag{B.20}$$

Subtracting $\Omega_{im}$ in eq. (14) from $\Omega_{day}$ in eq. (B.20), we have

$$\Omega_{day} - \Omega_{im} = A^{-1}BC^{-1}(\Omega^a_{day-im} + \Omega^b_{day-im})C^{-1}B'A^{-1},$$

where

$$\Omega^a_{day-im} \equiv \sum_{t=1}^{7} r_t \,\mathrm{E}[(Z_i'\alpha_t - X_i'\beta_t)^2 Z_i Z_i'] - \mathrm{E}\left[\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right)^2 Z_i Z_i'\right],$$
$$\Omega^b_{day-im} \equiv 2\sum_{t=1}^{7} r_t \,\mathrm{E}[V_{it}(Z_i'\alpha_t - X_i'\beta_t) Z_i Z_i'] - 2\,\mathrm{E}\left[\left(\sum_{t=1}^{7} V_{it}\right)\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right) Z_i Z_i'\right].$$

We will show that $\Omega^a_{day-im}$ is a variance-covariance matrix, $\Omega^b_{day-im}$ is a cross-covariance matrix, and their sum is also a cross-covariance matrix. Whether or not $\Omega^a_{day-im} + \Omega^b_{day-im}$ is positive definite depends on the covariance between $(U_{i1}, \ldots, U_{i7})'$ and $(V_{i1}, \ldots, V_{i7})'$.

The proof relies on two observations:

$$\beta = \sum_{t=1}^{7} \beta_t \qquad \text{and} \qquad Z_i'\alpha_t - X_i'\beta_t = Z_i'\alpha_t - H_{it} + H_{it} - X_i'\beta_t = U_{it} - V_{it}.$$

Because we will repeatedly use $U_{it} - V_{it}$, we denote $\eta_{it} \equiv U_{it} - V_{it}$. Using these two observations, we first can write $\Omega^a_{day-im}$ as follows,

$$\begin{aligned}
\Omega^a_{day-im} &= \sum_{t=1}^{7} E(\eta_{it}^2 Z_i Z_i') + \sum_{t=1}^{7}(r_t - 1)E(\eta_{it}^2 Z_i Z_i') - \mathrm{E}\left[\left(\sum_{t=1}^{7}\eta_{it}\right)^2 Z_i Z_i'\right] \\
&= \sum_{t=1}^{7} E(\eta_{it}^2 Z_i Z_i') + \sum_{t=1}^{7}(r_t - 1)E(\eta_{it}^2 Z_i Z_i') - \sum_{t=1}^{7} E(\eta_{it}^2 Z_i Z_i') - 2\sum_{1\le t<\tau\le 7} E(\eta_{it}\eta_{i\tau} Z_i Z_i') \\
&= \sum_{t=1}^{7}(r_t - 1)E(\eta_{it}^2 Z_i Z_i') - 2\sum_{1\le t<\tau\le 7} E(\eta_{it}\eta_{i\tau} Z_i Z_i') \\
&= \mathrm{E}\left[\left(\sum_{t=1}^{7}(r_t d_{it} - 1)\eta_{it} Z_i\right)\left(\sum_{t=1}^{7}(r_t d_{it} - 1)\eta_{it} Z_i'\right)\right],
\end{aligned} \tag{B.21}$$

where the last equality holds by Assumption 1 and the following equalities:

$$E[(r_t d_{it}-1)^2] = E(r_t^2 d_{it}^2) + 1 - 2E(r_t d_{it}) = E(r_t^2 d_{it}) + 1 - 2 = r_t - 1 = r_t - 1 \tag{B.22}$$

$$E[(r_t d_{it}-1)(r_\tau d_{i\tau}-1)] = E(r_t r_\tau d_{it} d_{i\tau}) - E(r_t d_{it}) - E(r_\tau d_{i\tau}) + 1 = -1. \tag{B.23}$$

Similarly, we have

$$\begin{aligned}
\frac{1}{2}\Omega^b_{day-im} &= \sum_{t=1}^{7} E(V_{it}\eta_{it}Z_iZ_i') + \sum_{t=1}^{7}(r_t-1)E(V_{it}\eta_{it}Z_iZ_i') - \mathrm{E}\Big[\Big(\sum_{t=1}^{7} V_{it}\Big)\Big(\sum_{t=1}^{7}\eta_{it}\Big)Z_iZ_i'\Big] \\
&= \sum_{t=1}^{7} E(V_{it}\eta_{it}Z_iZ_i') + \sum_{t=1}^{7}(r_t-1)E(V_{it}\eta_{it}Z_iZ_i') - \sum_{t=1}^{7} E(V_{it}\eta_{it}Z_iZ_i') - \sum_{t\neq\tau}\mathrm{E}\Big[V_{it}\eta_{i\tau}Z_iZ_i'\Big] \\
&= \sum_{t=1}^{7}(r_t-1)E(V_{it}\eta_{it}Z_iZ_i') - \sum_{t\neq\tau}\mathrm{E}\Big[V_{it}\eta_{i\tau}Z_iZ_i'\Big] \\
&= \mathrm{E}\Big[\Big(\sum_{t=1}^{7}(r_t d_{it}-1)V_{it}Z_i\Big)\Big(\sum_{t=1}^{7}(r_t d_{it}-1)\eta_{it}Z_i'\Big)\Big] \\
&= \mathrm{Cov}\Big(\sum_{t=1}^{7}(r_t d_{it}-1)V_{it}Z_i, \sum_{t=1}^{7}(r_t d_{it}-1)\eta_{it}Z_i\Big),
\end{aligned} \tag{B.24}$$

where the fourth equality holds again by Assumption 1, eq. (B.22) and eq. (B.23); the last equality holds since $Z_i$ are IVs which are uncorrelated with the zero mean $\eta_{it}$.

Next, we derive $\Omega^a_{day-im} + \Omega^b_{day-im}$ using eq. (B.21) and eq. (B.24). Note that $\eta_{it} = U_{it} - V_{it}$, hence $\eta_{it} + 2V_{it} = U_{it} + V_{it}$. We have

$$\begin{aligned}
\Omega^a_{day-im} + 2\Big(\frac{1}{2}\Omega^b_{day-im}\Big) &= \mathrm{E}\Big[\Big(\sum_{t=1}^{7}(r_t d_{it}-1)\eta_{it}Z_i\Big)\Big(\sum_{t=1}^{7}(r_t d_{it}-1)\eta_{it}Z_i'\Big)\Big] \\
&\quad + \mathrm{E}\Big[\Big(\sum_{t=1}^{7}(r_t d_{it}-1)2V_{it}Z_i\Big)\Big(\sum_{t=1}^{7}(r_t d_{it}-1)\eta_{it}Z_i'\Big)\Big] \\
&= \mathrm{E}\Big[\Big(\sum_{t=1}^{7}(r_t d_{it}-1)(U_{it}+V_{it})Z_i\Big)\Big(\sum_{t=1}^{7}(r_t d_{it}-1)(U_{it}-V_{it})Z_i'\Big)\Big] \\
&= \mathrm{Cov}\Big(\Big(\sum_{t=1}^{7}(r_t d_{it}-1)(U_{it}+V_{it})Z_i\Big), \Big(\sum_{t=1}^{7}(r_t d_{it}-1)(U_{it}-V_{it})Z_i\Big)\Big).
\end{aligned}$$

Again, by Assumption 1, eq. (B.22) and eq. (B.23), we can expand the covariance term in the last line and conclude that

$$\Omega_{day} - \Omega_{im} = A^{-1}BC^{-1}\Big[\sum_{t=1}^{7}(r_t - 1)E((U_{it} + V_{it})(U_{it} - V_{it})Z_iZ_i') \\ - \sum_{t\neq\tau} E((U_{it} + V_{it})(U_{i\tau} - V_{i\tau})Z_iZ_i')\Big]C^{-1}B'A^{-1}.$$

This completes the proof of Theorem 6. $\square$

**Remark 10** (Relative efficiency of $\hat{\beta}_{day}$ (cont'd))**.** *To demonstrate that the sign of $\Omega_{day} - \Omega_{im}$ in Theorem 6 is indeterminate in general, we note that under homoskedasticity and fixed effect assumptions, the difference between the asymptotic variances of $\hat{\beta}_{day}$ and $\hat{\beta}_{im}$ in eq. (17) can be simplified to*

$$\Omega_{day}-\Omega_{im} = \Big[-\sum_{t=1}^{7}(r_t-1)(2\beta_t'E(e_ic_i)+\beta_t'E(e_ie_i')\beta_t)+\sum_{t\neq\tau}(2\beta_\tau'E(e_ic_i)+\beta_\tau'E(e_ie_i')\beta_t)\Big]A^{-1}, \quad \text{(B.25)}$$

*where $c_i$ is the fixed effect defined below, and $e_i$ is the error term in the first stage regression of $X_i$ on IVs $Z_i$. The term $-\beta_t'E(e_ie_i')\beta_t$ is non-positive; but for $t \neq \tau$, the terms $-2\beta_t'E(e_ic_i)$, $2\beta_\tau'E(e_ic_i)$ and $\beta_\tau'E(e_ie_i')\beta_t$ could be positive or negative and their absolute values might be larger or smaller than that of the former. So whether or not $\hat{\beta}_{day}$ is asymptotically more efficient than $\hat{\beta}_{im}$ is indeterminate and it depends on the sign of $\beta_t$ $(t = 1, \ldots, 7)$ and the correlation between $e_i$ and $c_i$. Inspired by an anonymous referee, we conducted simple simulation experiments to demonstrate both $\Omega_{day} - \Omega_{im} > 0$ case and the opposite case. These results are not reported but available upon request.*

*In the rest of this remark, we will prove eq.* (B.25)*. First, assume homoskedasticity so that we can move $Z_iZ_i'$ out and rewrite $E((U_{it} + V_{it})(U_{i\tau} - V_{i\tau})Z_iZ_i') = E((U_{it} + V_{it})(U_{i\tau} - V_{i\tau}))E(Z_iZ_i')$ $(t = \tau$ or $t \neq \tau)$.*

*Recall the daily regression models $(t = 1, \ldots, 7)$*

$$H_{it} = X_i'\beta_t + U_{it},$$

*as well as the reduced form equations for $X_i$ and $H_{it}$*

$$X_i = \pi' Z_i + e_i,$$
$$H_{it} = (\pi' Z_i + e_i)'\beta_t + U_{it} = Z_i' \underbrace{\pi\beta_t}_{\alpha_t} + \underbrace{e_i'\beta_t + U_{it}}_{V_{it}},$$

*where we know that $E(Z_i e_i) = 0$, $E(Z_i U_{it}) = 0$ so $E(Z_i V_{it}) = 0$, but $E(e_i U_{it}) \neq 0$. In order to capture dependence among daily hours worked determined by unobserved factors, we postulate a common fixed effect structure*

$$U_{it} = c_i + \xi_{it},$$

*which in turn implies that $V_{it} = e_i'\beta_t + c_i + \xi_{it}$. So for any $t, \tau = 1, \ldots, 7$, we have*

$$\begin{aligned} E((U_{it} + V_{it})(U_{i\tau} - V_{i\tau})) &= -E(\beta_\tau' e_i (2U_{it} + e_i'\beta_t)) \\ &= -2E(\beta_\tau' e_i c_i) - 2E(\beta_\tau' e_i \xi_{it}) - E(\beta_\tau' e_i e_i' \beta_t) \\ &= -2\beta_\tau' E(e_i c_i) - \beta_\tau' E(e_i e_i')\beta_t, \end{aligned}$$

*where the last equality holds because $E(e_i c_i) \neq 0$ and $E(e_i \xi_{it}) = 0$ since the fixed effect might be correlated with the endogenous regressors $X_i$ but is uncorrelated with the idiosyncratic errors $\xi_{it}$. Plugging the last expression into the formula of $\Omega_{day} - \Omega_{im}$ in Theorem 6, we immediately get eq.* (B.25) *under homoskedasticity.*

*Proof of Theorem 7.* The result holds by the consistency of the estimators (Theorem 2), the law of large numbers and the continuous mapping theorem. The proof is standard and therefore is omitted here. □

*Proof of Remark 11.* Now we prove that $\tilde{\beta}_{day}$, the variation of the day estimator described in Remark 11, is asymptotically equivalent to the impute estimator under Assumptions 1 to 5.

Formally,

$$\tilde{\beta}_{day} \equiv \sum_{t=1}^{7} (X' P_z D_t P_z X)^{-1} X' P_z D_t H_t. \tag{B.26}$$

Our proof proceeds in three steps: first, we obtain the expression of $\sqrt{n}(\tilde{\beta}_{day} - \beta)$; second, we derive an asymptotically equivalent expression of $\sqrt{n}(\tilde{\beta}_{day} - \beta)$ by replacing some sample averages in the first step with their probability limits; third, we derive an asymptotically equivalent expression of $\sqrt{n}(\hat{\beta}_{im} - \beta)$ under Assumption 5 and show that it is the same as that in the second step.

First, recall that $H_t = X\beta_t + U_t$ and $P_z = Z(Z'Z)^{-1}Z'$, and note the decomposition

$$X = P_z X + (I - P_z)X, \tag{B.27}$$

so based on eq. (B.26), we get

$$\begin{aligned}
\tilde{\beta}_{day} &= \sum_{t=1}^{7}(X'P_zD_tP_zX)^{-1}X'P_zD_tP_zX\beta_t + \sum_{t=1}^{7}(X'P_zD_tP_zX)^{-1}X'P_zD_t[(I-P_z)X\beta_t + U_t] \\
&= \sum_{t=1}^{7}\beta_t + \sum_{t=1}^{7}(X'P_zD_tP_zX)^{-1}X'P_zD_t[(I-P_z)X\beta_t + U_t] \\
\Longrightarrow \sqrt{n}(\tilde{\beta}_{day} - \beta) &= \sqrt{n}\sum_{t=1}^{7}(X'P_zD_tP_zX)^{-1}X'P_zD_t[(I-P_z)X\beta_t + U_t] \\
&= \sqrt{n}\sum_{t=1}^{7}(X'P_zD_tP_zX)^{-1}X'Z(Z'Z)^{-1}Z'D_t[(I-P_z)X\beta_t + U_t] \\
&= \sum_{t=1}^{7}\left(\frac{X'P_zD_tP_zX}{n_t}\right)^{-1}\frac{X'Z}{n}\left(\frac{Z'Z}{n}\right)^{-1}\sqrt{\frac{n}{n_t}}\frac{1}{\sqrt{n_t}}Z'D_t[(I-P_z)X\beta_t + U_t],
\end{aligned} \tag{B.28}$$

since $\beta = \sum_{t=1}^{7}\beta_t$.

Second, note that $\frac{1}{n_t}Z'D_t(I-P_z)X\beta_t \xrightarrow{p.} 0$ because $(I-P_z)X$ is the vector of "$X$" first stage residuals (by regressing $X$ on $Z$) and by construction is uncorrelated with $Z$ for each diary day, since the diary day is completely random; in addition, $\frac{1}{n_t}Z'D_tU_t \xrightarrow{p.} 0$ if Assumption 5 holds (i.e., $E(Z_iU_{it}) = 0$). Based on these, a proper central limit theorem implies that $\frac{1}{\sqrt{n_t}}Z'D_t[(I-P_z)X\beta_t + U_t] \xrightarrow{d.} \mathcal{N}(0,\Sigma)$ with some positive definite matrix $\Sigma$. This further implies that in eq. (B.28), if we replace the terms in front of $\frac{1}{\sqrt{n_t}}Z'D_t[(I-P_z)X\beta_t+U_t]$ with their respective probability limits, the asymptotic distribution of $\sqrt{n}(\tilde{\beta}_{day} - \beta)$ won't be altered. As a result, we get

$$\sqrt{n}(\tilde{\beta}_{day} - \beta) = A^{-1}BC^{-1}\sum_{t=1}^{7}\sqrt{r_t}\frac{1}{\sqrt{n_t}}Z'D_t[(I-P_z)X\beta_t + U_t] + o_p(1). \tag{B.29}$$

Third, recall that $H_t = X\beta_t + U_t$ and $P_z = Z(Z'Z)^{-1}Z'$, and use the decomposition in eq. (B.27), we can rewrite $\hat{\beta}_{im}$ as follows:

$$\begin{aligned}
\hat{\beta}_{im} &= (X'P_zX)^{-1}X'P_z\sum_{t=1}^{7}Z(Z'D_tZ)^{-1}Z'D_tH_t \\
&= (X'P_zX)^{-1}X'P_z\sum_{t=1}^{7}Z(Z'D_tZ)^{-1}Z'D_tP_zX\beta_t \\
&\quad + (X'P_zX)^{-1}X'P_z\sum_{t=1}^{7}Z(Z'D_tZ)^{-1}Z'D_t[(I-P_z)X\beta_t+U_t] \\
&= \beta + (X'P_zX)^{-1}X'P_z\sum_{t=1}^{7}Z(Z'D_tZ)^{-1}Z'D_t[(I-P_z)X\beta_t+U_t] \\
\Longrightarrow \sqrt{n}(\hat{\beta}_{im}-\beta) &= \left(\frac{X'P_zX}{n}\right)^{-1}\frac{X'Z}{n}\sum_{t=1}^{7}\left(\frac{Z'D_tZ}{n_t}\right)^{-1}\sqrt{r_{nt}}\frac{1}{\sqrt{n_t}}Z'D_t[(I-P_z)X\beta_t+U_t].
\end{aligned} \tag{B.30}$$

When Assumption 5 hold, we can again replace the terms in front of $\frac{1}{\sqrt{n_t}}Z'D_t[(I-P_z)X\beta_t+U_t]$ in eq. (B.30) with their respective probability limits, without altering the asymptotic distribution of $\sqrt{n}(\hat{\beta}_{im}-\beta)$. As a result, we get

$$\sqrt{n}(\hat{\beta}_{im}-\beta) = A^{-1}BC^{-1}\sum_{t=1}^{7}\sqrt{r_t}\frac{1}{\sqrt{n_t}}Z'D_t[(I-P_z)X\beta_t+U_t]+o_p(1),$$

which is the same as eq. (B.29). This completes the proof of the asymptotic equivalence of $\tilde{\beta}_{day}$ and $\hat{\beta}_{im}$. □

# C When the ATUS Hours Have Classical Measurement Error

In this appendix, we provide detailed discussion about the consequence when the ATUS hours contain classical measurement error $e_{it}^{ATUS}$. To summarize: (i) the weekly labor supply elasticities $\beta$ are still identified; (ii) the estimators are still consistent and asymptotically normal; (iii) the asymptotic variance of the infeasible $\hat{\beta}_{wk}$ remains unchanged since it does not use the ATUS hours; (iv) the asymptotic variances of the feasible estimators all increase by $\sum_{t=1}^{7} r_t \operatorname{Var}(e_{it}^{ATUS})A^{-1}$. As a result, the asymptotic efficiency ranking among the estimators remains unchanged.

Let $H_{it}^{ATUS}$ denote the recorded hours worked on day $t$ by respondent $i$, and let $H_{it}$ denote the true hours worked on that day. On top of the assumptions in our main paper, the following assumption about the measurement error $e_{it}^{ATUS} = H_{it}^{ATUS} - H_{it}$ is maintained throughout this section.

**Assumption C1** (Classical measurement error in the ATUS)**.** *For all* $t = 1, \ldots, 7$*, we assume that* $E(e_{it}^{ATUS}) = 0$ *and* $e_{it}^{ATUS} \perp\!\!\!\perp (d_{i1}, \ldots, d_{i7}, Z_i', U_i)'$*.*

With Assumption C1, we can rewrite eq. (7) (main model) and eq. (10) (first stage) as follows,

$$H_{it}^{ATUS} = H_{it} + e_{it}^{ATUS} = X_i'\beta_t + \underbrace{U_{it} + e_{it}^{ATUS}}_{\equiv \tilde{U}_{it}},$$
$$H_{it}^{ATUS} = Z_i'\alpha_t + \underbrace{V_{it} + e_{it}^{ATUS}}_{\equiv \tilde{V}_{it}}.$$

For our purpose, $\tilde{U}_{it}$ differs from $U_{it}$ only by bringing larger variance (so does $\tilde{V}_{it}$ from $V_{it}$). So the statistical properties of the estimators in our main paper remain. We elaborate this point in what follows.

## C.1 Identification

The measurement error $e_{it}^{ATUS}$ does not enter the true weekly hours worked $H^w$, so the identification of $\beta$ still results from eq. (B.3) if the ATUS contains measurement errors.

For the feasible estimators based on the ATUS data, the identification of $\beta$ follows the same argument as in the proof of Theorem 1; that is, we only need to find the counterparts of eq. (B.5), eq. (B.6) and eq. (B.7) in the presence of classical measurement errors in the ATUS hours. By Assumption 1 and Assumption C1, we have

$$\begin{aligned}
E(Z_i H_{it}^{AUTS} | d_{it} = 1) &= E(Z_i H_{it} | d_{it} = 1) + E(Z_i e_{it}^{AUTS} | d_{it} = 1) \\
&= E(Z_i H_{it} | d_{it} = 1) + E(Z_i e_{it}^{AUTS}) \\
&= E(Z_i H_{it} | d_{it} = 1) + E(Z_i) E(e_{it}^{AUTS}) \\
&= E(Z_i H_{it} | d_{it} = 1), \qquad \text{(C.1)} \\
E(r_{nt} Z_i H_{it}^{AUTS} | d_{it} = 1) &= E(r_{nt} Z_i H_{it} | d_{it} = 1) + E(r_{nt} Z_i e_{it}^{AUTS} | d_{it} = 1) \\
&= E(r_{nt} Z_i H_{it} | d_{it} = 1) + E(r_{nt} Z_i e_{it}^{AUTS})
\end{aligned}$$

$$= E(r_{nt}Z_iH_{it}|d_{it} = 1) + E(r_{nt}Z_i)E(e_{it}^{AUTS})$$

$$= E(r_{nt}Z_iH_{it}|d_{it} = 1). \tag{C.2}$$

Plugging eq. (C.1) into eq. (B.5) and eq. (B.7) and plugging eq. (C.2) into eq. (B.6), we see that the identification of $\beta$ still holds when the ATUS contains classical measurement errors.

## C.2 Consistency

First, the infeasible estimator $\hat{\beta}_{wk}$ is not affected by the measurement error in the ATUS, and is still consistent. To see the consistency of other estimators when the ATUS contains classical measurement error, we only need to slightly modify eqs. (B.9) to (B.11), which were the key steps in establishing the consistency without measurement error. With measurement error, eq. (B.9) becomes

$$\begin{aligned}
\hat{\beta}_{im} - \hat{\beta}_{wk} &= \sum_{t=1}^{7} A_n^{-1}B_n\left[C_{nt}^{-1}\frac{1}{n_t}Z'D_t\tilde{V}_t - C_n^{-1}\frac{1}{n}Z'V_t\right] \\
&\xrightarrow{p.} \sum_{t=1}^{7} A^{-1}BC^{-1}[E(Z_id_{it}\tilde{V}_{it}) - E(Z_iV_{it})] \\
&= \sum_{t=1}^{7} A^{-1}BC^{-1}[E(Z_iV_{it})E(d_{it}) - E(Z_iV_{it})] \\
&= 0,
\end{aligned}$$

where the second equality holds by $E(Z_i\tilde{V}_{it}) = E(Z_iV_{it})$ and $d_{it} \perp\!\!\!\perp (Z_i, V_{it}, e_{it}^{ATUS})$. Since $\hat{\beta}_{wk}$ is consistent, so is $\hat{\beta}_{im}$. Let $e_t^{ATUS} = (e_{1t}^{ATUS}, \ldots, e_{nt}^{ATUS})'$, then eq. (B.10) becomes

$$\begin{aligned}
\hat{\beta}_{pool} - \hat{\beta}_{wk} &= \sum_{t=1}^{7} A_n^{-1}B_nC_n^{-1}\frac{Z'(r_{nt}D_t - I)H_t}{n} + \sum_{t=1}^{7} A_n^{-1}B_nC_n^{-1}\frac{Z'r_{nt}D_te_t^{ATUS}}{n} \\
&\xrightarrow{p.} 0 + A^{-1}BC^{-1}\sum_{t=1}^{7}\frac{Z'r_tD_te_t^{ATUS}}{n} \qquad \text{(by eq. (B.10))} \\
&\xrightarrow{p.} 0 + A^{-1}BC^{-1}\sum_{t=1}^{7} E(r_td_{it}Z_ie_{it}^{ATUS}) \\
&= 0,
\end{aligned}$$

where the last equality holds by Assumption C1. With measurement error, eq. (B.12) becomes

$$\begin{aligned}
\hat{\beta}_{day} - \beta &= \sum_{t=1}^{7} (X' P_{zt} X)^{-1} X' P_{zt} \tilde{U}_t \\
&\xrightarrow{p.} \sum_{t=1}^{7} A^{-1} B C^{-1} [E(Z_i U_{it}) + E(Z_i e_{it}^{ATUS})] \qquad \text{(by eq. (B.12))} \\
&= \sum_{t=1}^{7} A^{-1} B C^{-1} E(Z_i U_{it}) \\
&= 0,
\end{aligned}$$

where the second equality holds also by Assumption C1.

## C.3 Asymptotic Variances and Efficiency

First, the asymptotic variance of $\hat{\beta}_{wk}$ is not affected by the measurement error in the ATUS. To derive the asymptotic variance of the feasible estimators when the ATUS contains classical measurement error, we modify eq. (B.13), eq. (B.16) and eq. (16), which were the key steps in deriving the asymptotic variance without measurement error.

For the asymptotic variance of $\hat{\beta}_{im}$, eq. (B.13) becomes,

$$\begin{aligned}
\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk}) &= \sum_{t=1}^{7} A_n^{-1} B_n \left[ C_{n_t}^{-1} r_{nt} \frac{1}{\sqrt{n}} Z' D_t \tilde{V}_t - C_n^{-1} \frac{1}{\sqrt{n}} Z' V_t \right] \\
&= \sum_{t=1}^{7} A_n^{-1} B_n \left[ C_{n_t}^{-1} r_{nt} \frac{1}{\sqrt{n}} Z' D_t (V_t + e_t^{ATUS}) - C_n^{-1} \frac{1}{\sqrt{n}} Z' V_t \right].
\end{aligned}$$

By Assumption C1 and $n^{-1/2} Z' D_t e_t^{ATUS} = O_p(1)$, we see that

$$\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk}) = \underbrace{A^{-1} B C^{-1} \sum_{t=1}^{7} \frac{1}{\sqrt{n}} Z' (r_t D_t - I_n) V_t}_{\equiv \text{ part 1}} + \underbrace{A^{-1} B C^{-1} \sum_{t=1}^{7} \frac{1}{\sqrt{n}} Z' r_t D_t e_t^{ATUS}}_{\equiv \text{ part 2}} + o_p(1).$$

By Assumption C1, we get: (i) the asymptotic variance of part 2 is $\sum_{t=1}^{7} r_t \operatorname{Var}(e_{it}^{ATUS}) A^{-1}$; (ii) part 1 and part 2 are asymptotically independent; and (iii) part 1 is the same as the leading term in eq. (B.14). Taking account of these, we get

$$\widetilde{\Omega}_{im-wk} \equiv \mathrm{Var}\left(\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk})\right) = \Omega_{im-wk} + \sum_{t=1}^{7} r_t \,\mathrm{Var}(e_{it}^{ATUS})A^{-1},$$

where $\Omega_{im-wk}$ is defined in eq. (12). By Assumption C1, we have $e_{it}^{ATUS} \perp\!\!\!\perp U_i$, so we still have

$$\mathrm{Cov}\left(\sqrt{n}(\hat{\beta}_{im} - \hat{\beta}_{wk}), \sqrt{n}(\hat{\beta}_{wk} - \beta)\right) = o_p(1).$$

Therefore, the asymptotic variance of $\hat{\beta}_{im}$, when the ATUS contains classical measurement error, is $\widetilde{\Omega}_{im} \equiv \Omega_{wk} + \widetilde{\Omega}_{im-wk} = \Omega_{im} + \sum_{t=1}^{7} r_t \,\mathrm{Var}(e_{it}^{ATUS})A^{-1}$, where $\Omega_{wk}$ is defined in eq. (11) and $\Omega_{im}$ is defined in eq. (14). The new term $\sum_{t=1}^{7} r_t \,\mathrm{Var}(e_{it}^{ATUS})A^{-1}$ arises due to the measurement error.

For the asymptotic variance of $\hat{\beta}_{pool}$, eq. (B.16) remains valid even when we substitute $V_t$ with $\tilde{V}_t$, because $n^{-1/2} r_{nt} Z' D_t e_t^{ATUS} = O_p(1)$. So the asymptotic efficiency gap $\Omega_{pool-im}$ between $\hat{\beta}_{pool}$ and $\hat{\beta}_{im}$ remains unchanged even with classical measurement error in the ATUS hours. This further implies that the asymptotic variance of $\hat{\beta}_{pool}$ becomes $\widetilde{\Omega}_{pool} \equiv \Omega_{pool} + \sum_{t=1}^{7} r_t \,\mathrm{Var}(e_{it}^{ATUS})A^{-1}$, where $\Omega_{pool}$ is defined in eq. (15).

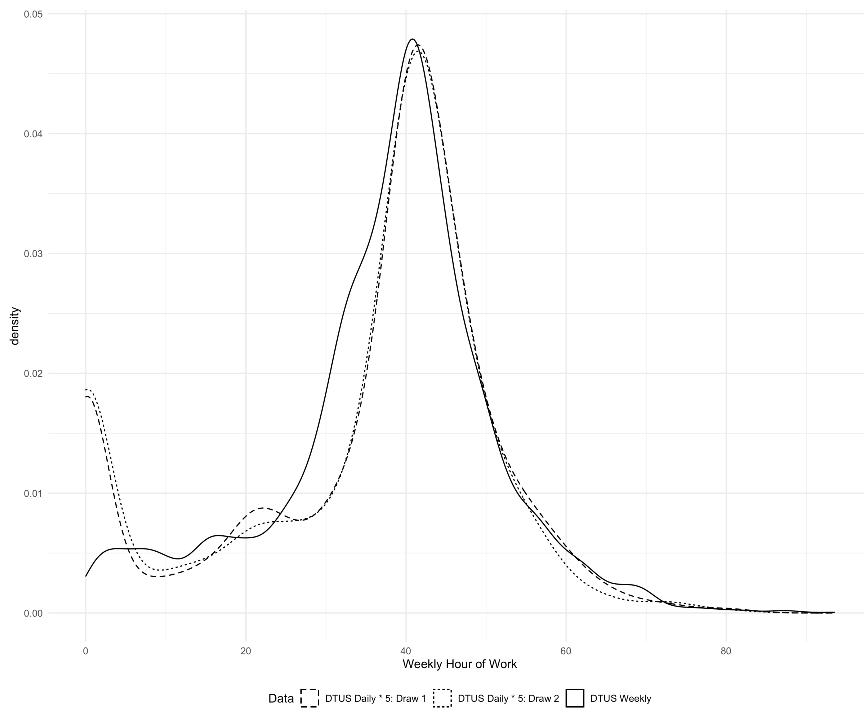
For the asymptotic variance of $\hat{\beta}_{day}$, we replace $U_{it}$ with $\tilde{U}_{it}$ in eq. (16). By Assumption C1 and the same argument as for $\hat{\beta}_{im}$, the asymptotic variance of $\hat{\beta}_{day}$, when the ATUS contains classical measurement error, is $\widetilde{\Omega}_{day} \equiv \Omega_{day} + \sum_{t=1}^{7} r_t \,\mathrm{Var}(e_{it}^{ATUS})A^{-1}$, where $\Omega_{day}$ is defined in eq. (16).

## References

**Angrist, Joshua D.**, "Grouped-Data Estimation and Testing in Simple Labor-Supply Models," *Journal of Econometrics*, 1991, *47* (2-3), 243-266.

**Chou, Cheng and Ruoyao Shi**, "What Time Use Surveys Can (And Cannot) Tell Us About Labor Supply," *unpublished manuscript*, 2020.

Figure A.1: DTUS Weekly Hours vs. Randomly Drawn Weekday Daily Hours ×5




*Note*: The DTUS sample used here is pooled across the years 1985, 1990, 1995, 2000, and 2005. The sample includes only full-time workers aged between 25 and 54 at the time of interview. We used the default sample weight of the DTUS, which makes the weighted frequencies of the diaries within each age and sex group are evenly distributed in a week.

Table A.1: Simulations Based Only on Weekdays in the Dutch Time Use Survey (DTUS)

| $\mathrm{Corr}(\tilde{X}_i, U_i)$ / $\mathrm{Corr}(\tilde{X}_i, \tilde{Z}_i)$ | | Panel A: $n = 250$ | | | | Panel B: $n = 500$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{pool}$ | $\hat{\beta}_{day}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{pool}$ | $\hat{\beta}_{day}$ |
| 0 / 1 | MSE | 0.002 | 0.019 | 0.019 | 0.019 | 0.001 | 0.009 | 0.009 | 0.009 |
| | Bias$^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.002 | 0.019 | 0.019 | 0.019 | 0.001 | 0.009 | 0.009 | 0.009 |
| 0.25 / 0.95 | MSE | 0.000 | 0.017 | 0.017 | 0.017 | 0.000 | 0.008 | 0.008 | 0.008 |
| | Bias$^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.000 | 0.017 | 0.017 | 0.017 | 0.000 | 0.008 | 0.008 | 0.008 |
| 0.5 / 0.80 | MSE | 0.002 | 0.019 | 0.019 | 0.020 | 0.001 | 0.009 | 0.009 | 0.009 |
| | Bias$^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.002 | 0.019 | 0.019 | 0.020 | 0.001 | 0.009 | 0.009 | 0.009 |
| 0.75/ 0.43 | MSE | 0.047 | 0.064 | 0.064 | 124.978 | 0.022 | 0.031 | 0.031 | 0.043 |
| | Bias$^2$ | 0.000 | 0.000 | 0.000 | 0.008 | 0.000 | 0.000 | 0.000 | 0.004 |
| | Var | 0.047 | 0.064 | 0.064 | 124.970 | 0.022 | 0.031 | 0.031 | 0.039 |
| $\mathrm{Corr}(\tilde{X}_i, U_i)$ / $\mathrm{Corr}(\tilde{X}_i, \tilde{Z}_i)$ | | Panel C: $n = 1000$ | | | | Panel D: $n = 2500$ | | | |
| | | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{pool}$ | $\hat{\beta}_{day}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{pool}$ | $\hat{\beta}_{day}$ |
| 0 / 1 | MSE | 0.001 | 0.004 | 0.005 | 0.004 | 0.000 | 0.002 | 0.002 | 0.002 |
| | Bias$^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.001 | 0.004 | 0.005 | 0.004 | 0.000 | 0.002 | 0.002 | 0.002 |
| 0.25 / 0.95 | MSE | 0.000 | 0.004 | 0.004 | 0.004 | 0.000 | 0.002 | 0.002 | 0.002 |
| | Bias$^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.000 | 0.004 | 0.004 | 0.004 | 0.000 | 0.002 | 0.002 | 0.002 |
| 0.5 / 0.80 | MSE | 0.001 | 0.004 | 0.005 | 0.005 | 0.000 | 0.002 | 0.002 | 0.002 |
| | Bias$^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.001 | 0.004 | 0.005 | 0.005 | 0.000 | 0.002 | 0.002 | 0.002 |
| 0.75/ 0.43 | MSE | 0.011 | 0.015 | 0.015 | 0.017 | 0.004 | 0.006 | 0.006 | 0.006 |
| | Bias$^2$ | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Var | 0.011 | 0.015 | 0.015 | 0.016 | 0.004 | 0.006 | 0.006 | 0.006 |

[1] This table compares finite sample performance of various estimators using the DTUS data. $10,000$ random samples of different sizes are drawn from the original DTUS sample of $6,567$ individual-year records.

[2] The two numbers in the first column represent: (i) correlation coefficient between regressor $\tilde{X}_i$ and error term $U_i$ (degree of endogeneity); (ii) correlation coefficient between regressor $\tilde{X}_i$ and IV $\tilde{Z}_i$ (strength of IV). Both are adjusted by changing the parameter $\rho$ in the simulation setup.

[3] $\hat{\beta}_{wk}$ is the 2SLS estimator given in eq. (5), which uses the accurate hours worked from Mondays to Fridays in the DTUS and serves as an infeasible benchmark for the three estimators based on the ATUS. $\hat{\beta}_{wk}$ has virtually no bias and the smallest variance.

[4] For each individual in the DTUS, we randomly draw one from the five weekdays using the (equal) diary day sampling probabilities of the ATUS, thus obtained samples that imitate the ATUS, and we apply $\hat{\beta}_{im}$, $\hat{\beta}_{pool}$ and $\hat{\beta}_{day}$ to them in order to evaluate their performance.

[5] $\hat{\beta}_{im}$ has virtually no bias and the smallest variance among the three, followed closely by $\hat{\beta}_{pool}$.

[6] $\hat{\beta}_{day}$ is numerically equivalent to $\hat{\beta}_{im}$ when $\tilde{X}_i$ is exogenous. When $\tilde{X}_i$ is endogenous, however, $\hat{\beta}_{day}$ could display notable bias and considerable variance, especially when the sample size is smaller (and hence each day subsample is even smaller).

[7] $\tilde{\beta}_{day}$ introduced in Remark 11 performs almost identically to $\hat{\beta}_{im}$, but we do not report it here to avoid repetition.

Table A.2: Weekly Labor Supply Regression Coefficient Estimates: the DTUS

| | Married Men | | | Married Women | | |
|---|---|---|---|---|---|---|
| | $\hat{\beta}_{re}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ | $\hat{\beta}_{re}$ | $\hat{\beta}_{wk}$ | $\hat{\beta}_{im}$ |
| $n$ of kids aged $< 18$ | 0.42 | 0.16 | 0.09 | 0.01 | −4.17 | −5.24 |
| | (0.18) | (0.24) | (0.48) | (0.36) | (0.43) | (0.83) |
| Educ: completed 2ndry | 0.95 | −0.48 | −3.10 | −0.96 | 2.95 | 2.44 |
| | (0.50) | (0.66) | (1.25) | (0.94) | (1.11) | (2.19) |
| Educ: above 2ndry | 1.84 | −0.85 | −2.33 | −0.39 | 5.63 | 5.37 |
| | (0.53) | (0.70) | (1.34) | (1.12) | (1.32) | (2.62) |
| P value of joint Hausman test | 0.00 | 0.11 | | 0.00 | 0.53 | |
| $n$ of Obs. | 1746 | 1746 | 1746 | 835 | 835 | 835 |
| $R$ squared[5] | 0.06 | 0.03 | 0.07 | 0.18 | 0.39 | 0.26 |

[1] The other control variables are age, age-squared, a dummy of working in private sector (with public sector as base group), an urban area dummy (with rural being base group), and year dummies.

[2] $\hat{\beta}_{re}$ uses the recalled weekly hours; $\hat{\beta}_{wk}$ uses the true diary weekly hours; $\hat{\beta}_{im}$ uses the fictitious sample where only one day is randomly chosen for each individual using the ATUS diary day sampling weights.

[3] Standard errors are in parentheses.

[4] We conduct the joint Hausman tests (i.e., the coefficients associated with the three regressors in the table) regarding whether there are significant differences between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$, and between $\hat{\beta}_{wk}$ and $\hat{\beta}_{im}$, respectively.

[5] The $R$ squared for impute estimator is the average $R$ squared of the seven linear regression of daily hours worked $H_{it} = X_i'\beta_t + U_{it}$ for $t = 1, \ldots, 7$.

Table A.3: Comparison between the Respondents in the ATUS and the CPS

| | ATUS | CPS (in ATUS or not, Table A.4) | Entire CPS |
|---|---|---|---|
| Male | 40.5% | 48.3% | 48.6% |
| College graduates | 21.3% | 18.1% | 18.5% |
| Age | 39.4 | 39.3 | 39.3 |
| s.d. | (8.4) | (8.6) | (8.7) |
| Hours usually worked per week | 36.1 | 38 | 38 |
| s.d. | (9.0) | (8.5) | (8.5) |
| Hourly wage (2017 US dollars) | 18.7 | 18.4 | 18.4 |
| s.d. | (9.0) | (8.8) | (8.8) |
| Num. of children aged $< 5$ | 0.23 | 0.21 | 0.20 |
| s.d. | (0.52) | (0.50) | (0.50) |
| Num. of children aged 5–18 | 0.79 | 0.92 | 0.90 |
| s.d. | (1.00) | (1.11) | (1.11) |
| Num. of obs. | 19,038 | 73,429 | 991,116 |

[1] “ATUS” column refers to the sample that was used in our empirical studies. “CPS (in ATUS or not, Table A.4)” column refers to the CPS 2003-2017 sample after the same sample selection criterion (hourly paid workers aged between of 25 and 54, whose wage rate is positive, and spouse earnings and total usual weekly hours worked at all jobs reported in the CPS are observed) is applied, whether they participate in the ATUS or not. “Entire CPS” differs from “CPS (in ATUS or not, Table A.4)” only in that “Entire CPS” keeps the respondents whose hourly wage or spouse weekly earnings is missing.

Table A.4: Weekly Labor Supply Elasticity Estimates: the CPS (in the ATUS or not)

| Panel A: Mean and std dev of hours and wage | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| CPS Usual Weekly Hours Worked | 41.02 | 39.21 | 34.90 | 36.65 |
| s.d. | (7.01) | (7.99) | (9.16) | (8.29) |
| Hourly Wage (2017 US dollars) | 21.22 | 17.92 | 17.79 | 16.23 |

| Panel B: Elasticities (hundredths)[2] | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| Wage | 7.66 | 11.15 | 10.02 | 12.41 |
| | (0.36) | (0.48) | (0.55) | (0.58) |
| Spouse weekly earnings | −0.29 | | −2.52 | |
| | (0.12) | | (0.24) | |
| Num. of kids age < 5 | 0.34 | | −6.10 | |
| | (0.21) | | (0.42) | |
| Num. of kids ages 5–18 | 0.30 | | −2.18 | |
| | (0.11) | | (0.17) | |
| $R$ squared | 0.16 | 0.18 | 0.18 | 0.17 |
| $n$ of obs. | 20,307 | 15,134 | 21,165 | 16,823 |

[1] The sample here contains the CPS 2003-2017 sample after the same sample selection criterion (hourly paid workers aged between of 25 and 54, whose wage rate is positive, and spouse earnings and total usual weekly hours worked at all jobs reported in the CPS are observed) is applied, whether they participate in the ATUS or not.

[2] The elasticities are evaluated at the respective mean hours worked in each data source.

[3] The other control variables are including age, age-squared, two education dummies, eight Census division dummies, a metropolitan area dummy, race dummies, year dummies, occupation dummies and industry dummies.

Table A.5: Weekly Labor Supply Regression Coefficient Estimates: the CPS and the ATUS

| Panel A: Mean and std dev of hours and wage | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| CPS Usual Weekly Hours Worked | 39.625 | 38.421 | 32.499 | 35.524 |
| s.d. | (6.130) | (7.260) | (10.430) | (8.630) |
| ATUS Hours Worked on Diary Day | 4.698 | 4.741 | 3.557 | 4.182 |
| s.d. | (4.550) | (4.440) | (4.000) | (4.210) |
| ATUS Imputed Weekly Hours Worked | 41.270 | 40.380 | 31.960 | 36.180 |
| s.d. (lower bound)[1] | (9.569) | (9.792) | (9.255) | (9.677) |
| Hourly Wage (2017 US dollars) | 21.877 | 18.649 | 18.699 | 16.564 |

| Panel B: Elasticities (hundredths)[2] | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| Wage (CPS) | 2.136 | 4.371 | 5.163 | 4.165 |
| | (0.353) | (0.406) | (0.410) | (0.380) |
| Wage (ATUS) | 0.607 | 1.902 | 3.349 | 2.945 |
| | (1.387) | (1.315) | (1.061) | (1.194) |
| Spouse weekly earnings ($100) (CPS) | −0.000 | | −0.003 | |
| | (0) | | (0) | |
| Spouse weekly earnings ($100) (ATUS) | −0.002 | | −0.002 | |
| | (0.001) | | (0.001) | |
| Num. of kids age $<$ 5 (CPS) | −0.316 | | −2.788 | |
| | (0.192) | | (0.266) | |
| Num. of kids age $<$ 5 (ATUS) | −0.445 | | −2.868 | |
| | (0.792) | | (0.673) | |
| Num. of kids ages 5–18 (CPS) | −0.002 | | −0.932 | |
| | (0.101) | | (0.138) | |
| Num. of kids ages 5–18 (ATUS) | −0.183 | | −0.383 | |
| | (0.464) | | (0.379) | |
| $R$ squared (CPS) | 0.083 | 0.149 | 0.219 | 0.147 |
| $R$ squared (ATUS) | 0.155 | 0.242 | 0.174 | 0.169 |
| $p$ value of joint Hausman test | 0.254 | 0.048 | 0.064 | 0.281 |
| $n$ of obs. | 3889 | 3816 | 5602 | 5731 |

[1] See footnote 47 in the paper for more details.
[2] The estimates based on the CPS recalled weekly hours are $\hat{\beta}_{re}$; the estimates based on the ATUS diary day hours are $\hat{\beta}_{im}$.
[3] The standard errors are in parentheses.
[4] The $R$ squared for impute estimator is the average $R$ squared of the seven linear regression of daily hours worked $H_{it} = X_i'\beta_t + U_{it}$ for $t = 1, \ldots, 7$.
[5] For each sample group, we conduct joint Hausman tests regarding whether there are significant differences between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$.
[6] The other control variables are including age, age-squared, two education dummies, eight Census division dummies, a metropolitan area dummy, race dummies, year dummies, occupation dummies and industry dummies.

Table A.6: The ATUS Sample Sizes of All Occupations and Percentages by Month

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Total $n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Management occupations | 10.9 | 9.2 | 10.8 | 7.4 | 5.8 | 7.8 | 9.3 | 7.3 | 8.6 | 8.2 | 6.5 | 8.3 | 1262 |
| Computer and mathematical science occupations | 10.0 | 8.2 | 9.2 | 8.5 | 8.7 | 8.0 | 6.7 | 7.6 | 8.1 | 8.8 | 8.4 | 7.8 | 3575 |
| Healthcare support occupations | 9.8 | 8.3 | 9.6 | 8.2 | 8.6 | 7.4 | 7.9 | 8.1 | 7.7 | 8.8 | 8.0 | 7.6 | 3777 |
| Sales and related occupations | 11.3 | 9.2 | 9.2 | 7.8 | 7.2 | 8.0 | 9.3 | 7.5 | 7.2 | 7.8 | 7.4 | 8.2 | 1443 |
| Office and administrative support occupations | 10.9 | 7.9 | 8.5 | 8.5 | 7.2 | 8.6 | 7.3 | 8.0 | 8.1 | 8.3 | 8.3 | 8.5 | 3669 |
| Construction and extraction occupations | 10.4 | 8.1 | 9.0 | 9.6 | 6.9 | 7.6 | 8.6 | 8.9 | 7.9 | 8.0 | 8.0 | 7.0 | 1032 |
| Installation, maintenance, and repair occupations | 9.8 | 8.1 | 9.9 | 8.5 | 8.4 | 7.6 | 7.2 | 7.3 | 8.5 | 8.3 | 8.7 | 7.7 | 885 |
| Production occupations | 9.6 | 7.8 | 9.2 | 8.6 | 7.9 | 8.2 | 7.9 | 8.3 | 7.6 | 9.0 | 8.9 | 7.1 | 2066 |
| Transportation and material moving occupations | 11.1 | 6.9 | 10.8 | 8.4 | 7.2 | 6.1 | 8.4 | 7.8 | 7.8 | 9.3 | 9.0 | 7.2 | 1329 |
| Monthly num. of obs. | 10.4 | 8.2 | 9.4 | 8.4 | 7.8 | 7.8 | 7.8 | 7.9 | 7.9 | 8.6 | 8.2 | 7.8 | 19 038 |

[1] The numbers are the percentage of sample size in the total sample size per occupation.

Table A.7: Weekly Labor Supply Elasticity Estimates: the CPS and the ATUS (Computer & Mathematical, Healthcare, Office & Administrative Occupations)

| Panel A: Mean and std dev of hours and wage[1] | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| CPS Usual Weekly Hours Worked | 38.87 | 37.22 | 31.97 | 35.20 |
| s.d. | (7.12) | (8.13) | (10.68) | (8.90) |
| ATUS Hours Worked on Interview Day | 4.64 | 4.76 | 3.47 | 4.18 |
| s.d. | (4.57) | (4.46) | (4.01) | (4.21) |
| ATUS Imputed Weekly Hours Worked | 40.69 | 37.85 | 30.72 | 35.89 |
| s.d. (lower bound)[2] | (10.37) | (10.63) | (9.41) | (9.67) |
| Hourly Wage (2017 US dollars) | 21.91 | 17.79 | 19.39 | 17.01 |

| Panel B: Elasticities (hundredths)[2] | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| Wage (CPS) | 6.61 | 13.78 | 13.65 | 9.22 |
| | (1.93) | (1.88) | (1.51) | (1.32) |
| Wage (ATUS) | 10.82 | 8.65 | 6.71 | 3.81 |
| | (6.39) | (6.13) | (4.02) | (3.84) |
| Spouse weekly earnings (CPS) | −1.67 | | −10.58 | |
| | (0.97) | | (0.94) | |
| Spouse weekly earnings (ATUS) | −5.01 | | −7.20 | |
| | (3.19) | | (2.62) | |
| Num. of kids age < 5 (CPS) | 0.77 | | −8.95 | |
| | (1.10) | | (0.97) | |
| Num. of kids age < 5 (ATUS) | 5.15 | | −9.67 | |
| | (3.54) | | (2.64) | |
| Num. of kids ages 5–18 (CPS) | 0.08 | | −3.26 | |
| | (0.59) | | (0.51) | |
| Num. of kids ages 5–18 (ATUS) | −1.84 | | −2.77 | |
| | (2.08) | | (1.43) | |
| $R$ squared (CPS) | 0.13 | 0.19 | 0.22 | 0.12 |
| $R$ squared (ATUS) | 0.42 | 0.40 | 0.18 | 0.18 |
| $p$ value of joint Hausman test | 0.46 | 0.40 | 0.04 | 0.15 |
| $n$ of obs. | 1227 | 1483 | 4224 | 4087 |

[1] This table only contains the three occupations with the most observations in the ATUS (see Table A.6).
[2] See footnote 47 in the paper for more details.
[3] The estimates based on the CPS recalled weekly hours are $\hat{\beta}_{re}$; the estimates based on the ATUS diary day hours are $\hat{\beta}_{im}$.
[4] The standard errors are in parentheses.
[5] The elasticities are evaluated at the respective mean hours worked in each data source.
[6] The $R$ squared for impute estimator is the average $R$ squared of the seven linear regression of daily hours worked $H_{it} = X_i'\beta_t + U_{it}$ for $t = 1, \ldots, 7$.
[7] For each sample group, we conduct joint Hausman tests regarding whether there are significant differences between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$.
[8] The other control variables are including age, age-squared, two education dummies, eight Census division dummies, a metropolitan area dummy, race dummies, year dummies, occupation dummies and industry dummies.

Table A.8: Weekly Labor Supply Elasticity Estimates: the CPS and the ATUS (Work-related Hours)

| Panel A: Mean and std dev of hours and wage[1] | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| CPS Usual Weekly Hours Worked | 39.63 | 38.42 | 32.50 | 35.52 |
| s.d. | (6.13) | (7.27) | (10.44) | (8.63) |
| ATUS Hours Worked on Diary Day | 4.70 | 4.75 | 3.56 | 4.19 |
| s.d. | (4.55) | (4.44) | (4.01) | (4.21) |
| ATUS Imputed Weekly Hours Worked | 41.38 | 40.45 | 31.99 | 36.19 |
| s.d. (lower bound)[2] | (9.57) | (9.80) | (9.26) | (9.69) |
| Hourly Wage (2017 US dollars) | 21.88 | 18.65 | 18.70 | 16.56 |

| Panel B: Elasticities (hundredths)[2] | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| Wage (CPS) | 5.39 | 11.38 | 15.89 | 11.72 |
| | (0.89) | (1.06) | (1.26) | (1.07) |
| Wage (ATUS) | 1.55 | 4.76 | 10.44 | 8.15 |
| | (3.35) | (3.25) | (3.32) | (3.31) |
| Spouse weekly earnings (CPS) | −0.19 | | −9.43 | |
| | (0.41) | | (0.77) | |
| Spouse weekly earnings (ATUS) | −3.47 | | −5.80 | |
| | (1.62) | | (2.12) | |
| Num. of kids age < 5 (CPS) | −0.80 | | −8.58 | |
| | (0.48) | | (0.82) | |
| Num. of kids age < 5 (ATUS) | −1.03 | | −8.95 | |
| | (1.90) | | (2.10) | |
| Num. of kids ages 5–18 (CPS) | −0.00 | | −2.87 | |
| | (0.26) | | (0.42) | |
| Num. of kids ages 5–18 (ATUS) | −0.47 | | −1.19 | |
| | (1.12) | | (1.18) | |
| $R$ squared (CPS) | 0.08 | 0.15 | 0.22 | 0.15 |
| $R$ squared (ATUS) | 0.16 | 0.24 | 0.17 | 0.17 |
| $p$ value of joint Hausman test | 0.26 | 0.05 | 0.06 | 0.28 |
| $n$ of obs. | 3889 | 3816 | 5602 | 5731 |

[1] The ATUS hours worked in this table include all work-related hours.
[2] See footnote 47 in the paper for more details.
[3] The estimates based on the CPS recalled weekly hours are $\hat{\beta}_{re}$; the estimates based on the ATUS diary day hours are $\hat{\beta}_{im}$.
[4] The standard errors are in parentheses.
[5] The elasticities are evaluated at the respective mean hours worked in each data source.
[6] The $R$ squared for impute estimator is the average $R$ squared of the seven linear regression of daily hours worked $H_{it} = X_i'\beta_t + U_{it}$ for $t = 1, \ldots, 7$.
[7] For each sample group, we conduct joint Hausman tests regarding whether there are significant differences between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$.
[8] The other control variables are including age, age-squared, two education dummies, eight Census division dummies, a metropolitan area dummy, race dummies, year dummies, occupation dummies and industry dummies.

Table A.9: Weekly Labor Supply Elasticity Estimates: the CPS and the ATUS (OLS)

| Panel A: Mean and std dev of hours and wage | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| CPS Usual Weekly Hours Worked | 39.63 | 38.42 | 32.50 | 35.52 |
| s.d. | (6.13) | (7.26) | (10.43) | (8.63) |
| ATUS Hours Worked on Diary Day | 4.70 | 4.74 | 3.56 | 4.18 |
| s.d. | (4.55) | (4.44) | (4.00) | (4.21) |
| ATUS Imputed Weekly Hours Worked | 41.39 | 40.30 | 31.95 | 36.18 |
| s.d. (lower bound)[1] | (9.57) | (9.79) | (9.26) | (9.68) |
| Hourly Wage (2017 US dollars) | 21.88 | 18.65 | 18.70 | 16.56 |

| Panel B: Elasticities (hundredths)[2] | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| Wage (CPS) | 5.24 | 10.99 | 15.31 | 11.47 |
| | (0.89) | (1.06) | (1.25) | (1.07) |
| Wage (ATUS) | 2.18 | 5.78 | 11.19 | 8.56 |
| | (3.21) | (3.14) | (3.21) | (3.17) |
| Spouse weekly earnings (CPS) | −0.26 | | −9.53 | |
| | (0.40) | | (0.75) | |
| Spouse weekly earnings (ATUS) | −2.94 | | −6.75 | |
| | (1.56) | | (2.02) | |
| Num. of kids age < 5 (CPS) | −0.80 | | −8.56 | |
| | (0.49) | | (0.82) | |
| Num. of kids age < 5 (ATUS) | −1.07 | | −8.19 | |
| | (1.92) | | (2.08) | |
| Num. of kids ages 5–18 (CPS) | −0.01 | | −2.87 | |
| | (0.26) | | (0.42) | |
| Num. of kids ages 5–18 (ATUS) | −1.03 | | −1.26 | |
| | (1.11) | | (1.17) | |
| $R$ squared (CPS) | 0.08 | 0.15 | 0.22 | 0.15 |
| $R$ squared (ATUS) | 0.16 | 0.24 | 0.17 | 0.17 |
| $p$ value of Hausman test | 0.36 | 0.11 | 0.14 | 0.37 |
| $n$ of obs. | 3889 | 3816 | 5602 | 5731 |

[1] See footnote 47 in the paper for more details.
[2] The estimates based on the CPS recalled weekly hours are $\hat{\beta}_{re}$; the estimates based on the ATUS diary day hours are $\hat{\beta}_{im}$.
[3] The standard errors are in parentheses.
[4] The elasticities are evaluated at the respective mean hours worked in each data source.
[5] The $R$ squared for impute estimator is the average $R$ squared of the seven linear regression of daily hours worked $H_{it} = X_i'\beta_t + U_{it}$ for $t = 1, \ldots, 7$.
[6] For each sample group, we conduct joint Hausman tests regarding whether there are significant differences between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$.
[7] The other control variables are including age, age-squared, two education dummies, eight Census division dummies, a metropolitan area dummy, race dummies, year dummies, occupation dummies and industry dummies.

Table A.10: Weekly Labor Supply Elasticity Estimates: the CPS and the ATUS (Year-Month Grouped IV)

| Panel A: Mean and std dev of hours and wage | | | | |
|---|---|---|---|---|
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| CPS Usual Weekly Hours Worked | 39.63 | 38.42 | 32.50 | 35.52 |
| s.d. | (6.13) | (7.26) | (10.43) | (8.63) |
| ATUS Hours Worked on Diary Day | 4.70 | 4.74 | 3.56 | 4.18 |
| s.d. | (4.55) | (4.44) | (4.00) | (4.21) |
| ATUS Imputed Weekly Hours Worked | 41.56 | 40.51 | 31.85 | 35.79 |
| s.d. (lower bound)[1] | (9.57) | (9.79) | (9.26) | (9.68) |
| Hourly Pay (2017 US dollars) | 21.88 | 18.65 | 18.70 | 16.56 |
| Panel B: Elasticities (hundredths)[2] | | | | |
| | Married Men | Unmarried Men | Married Women | Unmarried Women |
| Wage (CPS) | 6.04 | 10.15 | 21.78 | 18.81 |
| | (2.68) | (2.93) | (3.97) | (3.51) |
| Wage (ATUS) | 0.00 | 1.59 | −2.10 | 1.72 |
| | (11.17) | (9.80) | (12.23) | (10.47) |
| Spouse weekly earnings (CPS) | −0.18 | | −11.45 | |
| | (1.27) | | (2.59) | |
| Spouse weekly earnings (ATUS) | 0.00 | | 0.49 | |
| | (5.84) | | (7.77) | |
| Num. of kids age $<5$ (CPS) | −0.91 | | −8.86 | |
| | (0.49) | | (0.82) | |
| Num. of kids age $<5$ (ATUS) | −0.16 | | −8.52 | |
| | (1.98) | | (2.11) | |
| Num. of kids ages 5–18 (CPS) | 0.02 | | −2.77 | |
| | (0.26) | | (0.43) | |
| Num. of kids ages 5–18 (ATUS) | −0.87 | | −1.87 | |
| | (1.14) | | (1.19) | |
| $R$ squared (CPS) | 0.08 | 0.14 | 0.21 | 0.13 |
| $R$ squared (ATUS) | 0.12 | 0.20 | 0.15 | 0.14 |
| $p$ value of Hausman test | 0.60 | 0.39 | 0.04 | 0.09 |
| $n$ of obs. | 3889 | 3816 | 5602 | 5731 |

[1] See footnote 47 in the paper for more details.
[2] The estimates based on the CPS recalled weekly hours are $\hat{\beta}_{re}$; the estimates based on the ATUS diary day hours are $\hat{\beta}_{im}$.
[3] The standard errors are in parentheses.
[4] The elasticities are evaluated at the respective mean hours worked in each data source.
[5] The $R$ squared for impute estimator is the average $R$ squared of the seven linear regression of daily hours worked $H_{it} = X_i'\beta_t + U_{it}$ for $t = 1, \ldots, 7$.
[6] For each sample group, we conduct joint Hausman tests regarding whether there are significant differences between $\hat{\beta}_{re}$ and $\hat{\beta}_{im}$.
[7] The other control variables are including age, age-squared, two education dummies, eight Census division dummies, a metropolitan area dummy, race dummies, year dummies, occupation dummies and industry dummies.

Table A.11: Pearson's Chi-squared Test for Independence Between Diary Day and Other Variables

| Variables | P-Values[1] |
|---|---|
| Wage decile | 0.63 |
| Spouse wage decile | 0.87 |
| CPS usual weekly hours worked[2] | 0.58 |
| Education | 0.91 |
| Num. of kids age $< 5$ | 0.61 |
| Num. of kids ages 5–18 | 0.07 |
| Age | 0.46 |
| Marriage status | 0.68 |
| Occupation | 0.69 |
| Industry | 0.82 |
| Metropolitan area dummy | 0.83 |
| Region | 0.35 |
| Year | 0.55 |
| Race | 0.01[3] |

[1] The null hypothesis is that the diary day is independent of the corresponding variable.

[2] The CPS recalled hours in our sample have only 76 different values, which is likely due to "bagging" issue in recalled hours. We treat the recalled hours as discrete variable in implementing the chi-squared test.

[3] Though the P-value associated with race is small, Table A.12 below shows that there is in fact no substantial variation of racial composition across the seven days of a week.

Table A.12: Proportion of Races Across Seven Days

| Day | White Non-Hispanic | Black Non-Hispanic | Other Race Non-Hispanic | Hispanic |
|---|---|---|---|---|
| 1 | 0.64 | 0.15 | 0.05 | 0.16 |
| 2 | 0.63 | 0.17 | 0.05 | 0.15 |
| 3 | 0.63 | 0.15 | 0.05 | 0.18 |
| 4 | 0.67 | 0.15 | 0.05 | 0.13 |
| 5 | 0.61 | 0.17 | 0.05 | 0.17 |
| 6 | 0.64 | 0.15 | 0.05 | 0.17 |
| 7 | 0.64 | 0.15 | 0.04 | 0.17 |