(NOT FOR PUBLICATION)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 (2021). 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 (2021).

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 and the IVs Z_i are generated from the $n \times 5$ matrix with elements H_{it}^{DTUS} ($t = 2, \dots, 6$), denoted by $H^{DTUS,5}$, using

[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.

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 (2021), 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.

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

[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.

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]

[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.

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 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 β 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 β is just the usual argument for 2SLS (i.e., generalized method of moments) estimators. Formally, β 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 β, 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 β 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 $E(Z_i H_i^w)$ in eq. (B.3). Since the expression of β in eq. (B.3) is the *unique solution* of eq. (B.2), the identification of β 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

$$\beta = \left(E(X_i Z_i')[E(Z_i Z_i')]^{-1} E(Z_i X_i')\right)^{-1} E(X_i Z_i')[E(Z_i Z_i')]^{-1} \sum_{t=1}^{7} E(Z_i H_{it}) \tag{B.4}$$

$$= \left(E(X_i Z_i')[E(Z_i Z_i')]^{-1} E(Z_i X_i')\right)^{-1} E(X_i Z_i') \sum_{t=1}^{7} [E(Z_i Z_i'|d_{it}=1)]^{-1} E(Z_i H_{it}|d_{it}=1) \tag{B.5}$$

$$= \left(E(X_i Z_i')[E(Z_i Z_i')]^{-1} E(Z_i X_i')\right)^{-1} E(X_i Z_i')[E(Z_i Z_i')]^{-1} \sum_{t=1}^{7} E(r_{nt} d_{it}) E(Z_i H_{it})$$

$$= \left(E(X_i Z_i')[E(Z_i Z_i')]^{-1} E(Z_i X_i')\right)^{-1} E(X_i Z_i')[E(Z_i Z_i')]^{-1} \sum_{t=1}^{7} E(r_{nt} d_{it} Z_i H_{it})$$

$$= \left(E(X_i Z_i')[E(Z_i Z_i')]^{-1} E(Z_i X_i')\right)^{-1} E(X_i Z_i')[E(Z_i Z_i')]^{-1} \sum_{t=1}^{7} E(r_{nt} Z_i H_{it}|d_{it}=1) \tag{B.6}$$

$$= \sum_{t=1}^{7} \left(E(X_i Z_i'|d_{it}=1)[E(Z_i Z_i'|d_{it}=1)]^{-1} E(Z_i X_i'|d_{it}=1)\right)^{-1}$$
$$\times E(X_i Z_i'|d_{it}=1)[E(Z_i Z_i'|d_{it}=1)]^{-1} E(Z_i H_{it}|d_{it}=1), \tag{B.7}$$

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 $\hat{\beta}_{wk}$:

$$\hat{\beta}_{wk} - \beta = A_n^{-1} X' P_z U = A_n^{-1} B_n C_n^{-1} (Z'U/n) \xrightarrow{p.} A^{-1} B C^{-1} E(Z_i U_i) = 0.$$

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:

$$\hat{\beta}_{im} - \hat{\beta}_{wk} = (X'P_z X)^{-1} X' P_z \left[\sum_{t=1}^{7} Z(Z'D_t Z)^{-1} Z' D_t H_t - H^w\right]$$
$$= (X'P_z X)^{-1} X' P_z \left[\sum_{t=1}^{7} Z(Z'D_t Z)^{-1} Z' D_t H_t - P_z \sum_{t=1}^{7} H_t\right]$$
$$= \sum_{t=1}^{7} (X'P_z X)^{-1} X' P_z Z[(Z'D_t Z)^{-1} Z' D_t H_t - (Z'Z)^{-1} Z' H_t]$$

$$= \sum_{t=1}^{7} (X'P_z X)^{-1} X'Z[(Z'D_t Z)^{-1} Z'D_t H_t - (Z'Z)^{-1} Z'H_t].$$

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

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

Define

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

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

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

because $E(Z_i V_{it}) = 0$. 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

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

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

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

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

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

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

$$E(r_t D_t - I_n) = 0$$

implies

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

As a result,

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

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

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

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

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

and

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

As a result,

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

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 $1 \le t \le \tau \le 7$,

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

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 Ω_{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 $\mathrm{Var}(\sqrt{n}(\hat\beta_{im}-\hat\beta_{wk}))$ as the following sum,

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

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

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

We hence conclude that $\mathrm{Var}(\sqrt{n}(\hat\beta_{im}-\hat\beta_{wk})) = \mathrm{Var}(\sqrt{n}(\hat\beta_{im}-\beta)) - \mathrm{Var}(\sqrt{n}(\hat\beta_{wk}-\beta))$. 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

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

We can then replace $E(U_i^2 Z_i Z_i')$ in the middle of Ω_{wk} in eq. (11) by eq. (B.19). Part (i) follows by adding Ω_{wk} and Ω_{im-wk} together, which are given in eq. (11) and eq. (12), respectively. Since Ω_{im-wk} involves terms like $E(Z_i V_{it} V_{i\tau} Z_i')$, it may seem at a glance that Ω_{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 Ω_{wk} and Ω_{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 $E\left[\left(Z_i'\sum_{t=1}^{7}\alpha_t - X_i'\beta\right)^2 Z_i Z_i'\right]$ in Ω_{im} and adding it together with $\Omega_{pool-im}$ in eq. (13). □

Proof of Theorem 6. Part (i). For every $t = 1, \ldots, 7$, 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 Ω_{day}. Similar to eq. (B.18), we can decompose U_{it} in a similar manner:

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

which implies that

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

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

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

Subtracting Ω_{im} in eq. (14) from Ω_{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

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

We will show that Ω^a_{day-im} is a variance-covariance matrix, Ω^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 $(U_{i1}, \dots, U_{i7})'$ and $(V_{i1}, \dots, V_{i7})'$.

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 Ω^a_{day-im} as follows,

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

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

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

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 η_{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

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

$$= \mathrm{E}\Big[\Big(\sum_{t=1}^{7}(r_t d_{it} - 1)(U_{it} + V_{it})Z_i\Big)\Big(\sum_{t=1}^{7}(r_t d_{it} - 1)(U_{it} - V_{it})Z_i'\Big)\Big]$$

$$= \mathrm{Cov}\Big(\Big(\sum_{t=1}^{7}(r_t d_{it} - 1)(U_{it} + V_{it})Z_i\Big), \Big(\sum_{t=1}^{7}(r_t d_{it} - 1)(U_{it} - V_{it})Z_i\Big)\Big).$$

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. □

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 β_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 ξ_{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

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

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 Σ. 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

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

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:

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

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 $\hat{\beta}_{im}$. $\square$

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 β 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 β still results from eq. (B.3) if the ATUS contains measurement errors.

For the feasible estimators based on the ATUS data, the identification of β 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

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

$$\begin{aligned} E(r_{nt} Z_i H_{it}^{AUTS}|d_{it}=1) &= E(r_{nt} Z_i H_{it}|d_{it}=1) + E(r_{nt} Z_i e_{it}^{AUTS}|d_{it}=1) \\ &= E(r_{nt} Z_i H_{it}|d_{it}=1) + E(r_{nt} Z_i e_{it}^{AUTS}) \\ &= E(r_{nt} Z_i H_{it}|d_{it}=1) + E(r_{nt} Z_i)E(e_{it}^{AUTS}) \\ &= E(r_{nt} Z_i H_{it}|d_{it}=1). \end{aligned} \tag{C.2}$$

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 β 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

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

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

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

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

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

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

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

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

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

Therefore, the asymptotic variance of $\hat{\beta}_{im}$, when the ATUS contains classical measurement error, is $\widetilde{\Omega}_{im} \equiv \Omega_{wk} + \widetilde{\Omega}_{im-wk} = \Omega_{im} + \sum_{t=1}^{7} r_t \operatorname{Var}(e_{it}^{ATUS}) A^{-1}$, where Ω_{wk} is defined in eq. (11) and Ω_{im} is defined in eq. (14). The new term $\sum_{t=1}^{7} r_t \operatorname{Var}(e_{it}^{ATUS}) A^{-1}$ 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 $\widetilde{\Omega}_{pool} \equiv \Omega_{pool} + \sum_{t=1}^{7} r_t \operatorname{Var}(e_{it}^{ATUS}) A^{-1}$, where Ω_{pool} is defined in eq. (15).

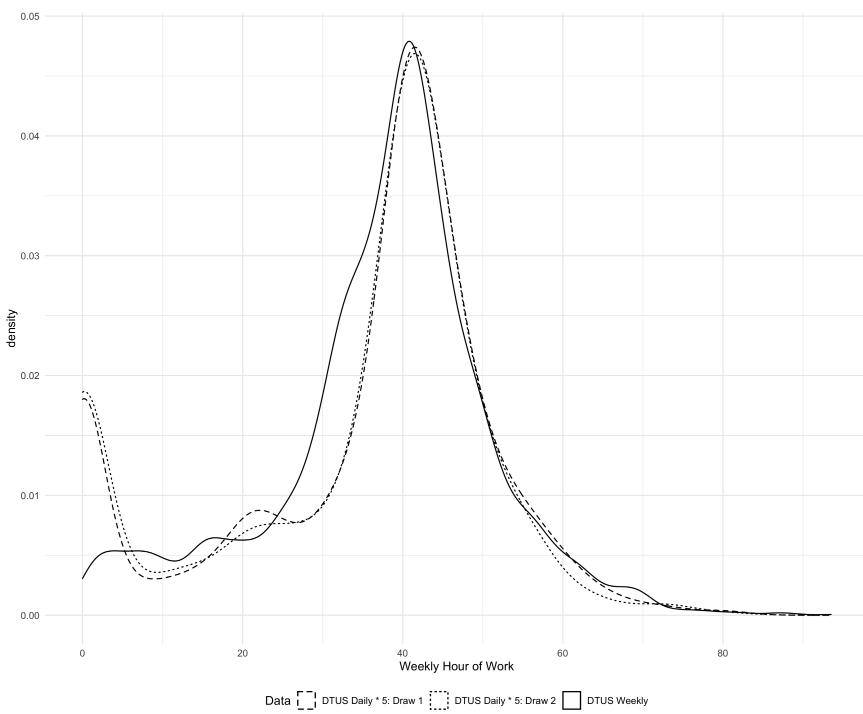
For the asymptotic variance of $\hat{\beta}_{day}$, we replace U_{it} with $\tilde{U}_{it}$ in eq. (16). By Assumption C1 and the same argument as for $\hat{\beta}_{im}$, the asymptotic variance of $\hat{\beta}_{day}$, when the ATUS contains classical measurement error, is $\widetilde{\Omega}_{day} \equiv \Omega_{day} + \sum_{t=1}^{7} r_t \operatorname{Var}(e_{it}^{ATUS}) A^{-1}$, where Ω_{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," *Journal of Applied Econometrics*, forthcoming, 2021.

Figure A.1: DTUS Weekly Hours vs. Randomly Drawn Weekday Daily Hours $\times 5$



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)

Corr($\tilde{X}_i, U_i$) / 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
Corr($\tilde{X}_i, U_i$) / 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 ρ 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