I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10404

Labor Supply Elasticities: Overcoming Nonclassical Measurement Error Using More Accurate Hours Data

Garry F. Barrett
Daniel S. Hamermesh

DECEMBER 2016

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10404

Labor Supply Elasticities: Overcoming Nonclassical Measurement Error Using More Accurate Hours Data

Garry F. Barrett
University of Sydney and IZA

Daniel S. Hamermesh
Royal Holloway University of London, IZA, NBER and University of Texas at Austin

DECEMBER 2016

IZA – Institute of Labor Economics

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

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

www.iza.org

ABSTRACT

Labor Supply Elasticities: Overcoming Nonclassical Measurement Error Using More Accurate Hours Data*

We measure the impact of measurement error in labor-supply elasticities estimated over recalled usual work hours, as is ubiquitous in the literature. Employing hours of work in diaries collected by the American Time Use Survey, 2003-12, along with the same respondents' recalled usual hours, we show that the latter yield elasticities that are positively biased. We argue that this bias arises from the salience on recalled hours of differences in wage rates.

JEL Classification: time use, measurement error, labor supply
Keywords: J22, C21

Nontechnical summary:

Using data for a large set of individuals who *both* say how much they recall working last week and also record a diary of time working the previous day, we examine how these vary among workers in relation to their wage rates. This is important, because all of the immense number of studies that examine labor supply rely on the recall measure. The results show that recalled hours worked respond more to differences in wage rates than does the actual time spent working. This result suggests that previous estimates of responses of effort to income-tax changes, unemployment insurance and other programs overstate their likely effects on effort – the disincentive effects are not so large as is commonly believed.

Corresponding author:
Daniel S. Hamermesh
Royal Holloway University of London
Department of Economics
214 Horton Building
Egham, Surrey TW20 0EX
United Kingdom
E-mail: Daniel.Hamermesh@rhul.ac.uk

* We thank Colin Cameron, Jay Stewart and Felix Weinhardt for helpful comments.

I. Hours on the Job and Hours Working

The elasticity of labor supply is one of the most studied parameters describing individual behavior in the sub-field of labor economics and, because of its importance for explaining macroeconomic fluctuations, perhaps in the entire discipline of economics. Large numbers of estimates have been summarized over the past three decades (Pencavel, 1986; Killingsworth and Heckman, 1986;, Blundell and MaCurdy, 1999; and Keane, 2011). In all studies based on representative samples of employees the measure of the intensive margin of labor supply has been the survey respondent's recalled hours worked in the previous week (typically in monthly labor-force surveys), month or year (typically in longitudinal annual household surveys). The use of such longer-recall data to measure labor-supply elasticities remains ubiquitous (e.g., Bargain *et al*, 2014; Blundell *et al*, 2016; Cherchye *et al*, 2012; Goux *et al*, 2014).[1]

In this note we examine estimates of labor supply elasticities using a data set that, in addition to requiring only very short recall, imposes adding-up restrictions upon respondents' perceptions of their time allocated to different activities. Estimation of labor supply elasticities based on this alternative, and arguably more accurate, survey methodology is novel and of independent interest. Our main purpose, however, is to use the alternative hours of work measure to assess the reliability of elasticity estimates based on the traditional hours measures. While distinctions between an individual's work hours on a particular job reported by the employer have been noted to differ from what the worker recalls (Duncan and Hill, 1985; Bound *et al*, 1994), the

[1]The main exception to this pattern has been a spate of studies on the responses of independent contractors' effort to variations in the price of the product they sell (e.g., most recently Stafford, 2015)

differences between these measures were not and, indeed, could not be used to analyze the impact of measurement errors on estimated labor-supply responses.[2]

II. Time-Use Data in the Estimation of Labor Supply

The American Time Use Survey (ATUS) provides a roughly one-eighth sub-sample of recent respondents to the monthly household Current Population Survey (CPS) in the United States. On some morning two to five months after their final appearance in the CPS, ATUS respondents write up a diary detailing their activities at each minute of the previous day, which upon diurnal aggregation yields total hours worked (defined in various ways), H^{A}. In addition, they answer the standard CPS question requiring them to recollect their usual weekly work hours, H^{CPS} (Hamermesh *et al*, 2005).

We use diarists in the ATUS from 2003-12. Given the relatively few hours worked on weekend days, we use only weekdays. Perhaps most important, because of the division bias induced in calculating the hourly earnings of salaried workers (see Borjas, 1980), we limit the sample to hourly-paid workers, who unlike other workers directly report an hourly wage rate. These restrictions generate samples of 3,925 married and 4,262 unmarried women. Although the literature on women's labor supply elasticities is larger, that on men's is also huge. We thus create samples of male ATUS respondents using the same restrictions, yielding 3,840 married and 3,086 unmarried men.

The first three columns of Table 1 present descriptive statistics on H^{CPS} and H^{A}. Throughout we use two measures of H^{A}, one reporting total work time on the diary day, H^{A1}, the other calculated as total work time minus commuting time, H^{A2}, in both cases converting reported minutes to daily hours. The correlations between these latter two measures are nearly one. Thus

[2] French (2004) considered the effect of nonclassical measurement error in recalled annual hours and wages for the estimation of the intertemporal elasticity of substitution based on the Panel Study of Income Dynamics.

while we present labor-supply elasticities estimated using both measures, we expect very small differences between the results.

Assuming a five-day workweek (and regrettably there is no information on days worked in these data), the mean reported hours of work less commuting time (third column of Table 1) differ only very slightly from recalled usual weekly hours.[3] This result on means repeats similar findings by Juster and Stafford (1991, Table 5) for the U.S. and Gershuny and Robinson (2013) for Europe. The correlations between the diary measures and usual weekly hours, however, are not high. The final column of Table 1 describes hourly wages among these samples of hourly-paid workers. Not surprisingly, given selectivity into hourly-paid employment, the means are well below economy-wide average hourly earnings.

Since our main purpose here is to compare estimates of labor supply elasticities using H^{CPS} and H^{A}, we use simple stylized labor-supply models that only adjust for obviously exogenous covariates. Thus we estimate:

(1) $h^i = \alpha^i w + \beta X + e$, i = CPS, A1 or A2,

where the lower-case symbols indicate logarithms, the vector X contains a quadratic in age, and fixed effects for state of residence, day of week, and year of the survey, the α^i and β are parameters to be estimated, and e is an error term. The estimated α^i are shown for married and unmarried women and men separately in Table 2. The estimated responses are fairly low, but not far from the weighted middle of the vast range of estimates produced in the literature.

The crucial thing to note from these estimates is that $\hat{\alpha}^{CPS} > \hat{\alpha}^{A}$ for both time-diary measures of work time, for both married and unmarried people, and for both women and men. The

[3]Very occasionally Supplements to the May CPS have collected information on days worked. In the May 1991 Supplement the mean days worked was 4.75 (5.07) among female (male) workers (Hamermesh, 1998).

absolute differences are not huge, but they approach or exceed conventional levels of statistical significance.[4] The central question thrown up by these results is why this difference arises? If, as we believe, the diary records of yesterday's work time are less error-prone than the one-week recalls of usual hours of work, the existence of classical measurement errors in the latter would suggest that the estimated α^{CPS} would be less than the estimated α^{A}, and the R^2 would be lower, in both cases the opposite of what we find.

III. Nonclassical Measurement Error

H^{CPS} is collected by interview shortly before former CPS respondents complete their one-day time diaries. The hourly wage rates of hourly-paid diarists are transcribed from their responses in their eighth month in the CPS if they were working and paid hourly then. In our sub-samples about 85 (87) percent of the female (male) diarists were employed and paid hourly in their eighth CPS month. 14 (13) percent of female (male) diarists were in the ATUS two months after their final CPS interview; 71 (69) percent three months later; 14 (17) percent four months later, and 1 (1) percent five months later. Thus the measure of hourly wages precedes responses about H^{CPS} and H^{A} on average by three months.[5]

The clear implication from the estimation of the labor supply elasticities using ATUS data is that the CPS question eliciting information on usual weekly work hours introduces nonclassical measurement error. There are three sources of measurement error in H^{CPS} relative to H^{A}. They may be categorized as relating to recall error, the specificity of the observation period and the constraint implied by the application of a 24-hour time budget. Recall error is likely greater in

---

[4]We test this by calculating the statistics generated after using STATA's *sureg* procedure to estimate (1) for h^{CPS} and h^{A1} (h^{A2}) jointly.

[5]Re-estimating the models in Section II restricting the sample to those who reported hourly pay in CPS month 8 produced only very small changes in the results. Similarly small changes were produced when we deleted the roughly 10 percent of each sample who report holding more than one job.

H^{CPS} than in H^A, as the report is based on the recall of usual hours over a prior week, in contrast to recorded hours worked on the previous day. A further source of measurement error arises from the specificity of the ATUS diary (the actual hours at work in a specific 24-hour time period) relative to the more general usual week reference period for H^{CPS}. Finally, H^A from the time diary includes an explicit aggregate 24-hour time constraint on all activities over the previous day, which imposes additional discipline on the diary measure that is absent from H^{CPS}.

Of these sources of measurement error, the lesser specificity and clarity of H^{CPS} is likely to introduce only random noise, which should make the R^2 using this measure lower and should attenuate the estimates of α, compared to those based on h^A. In contrast, recall error and the absence of a requirement that the respondent must list all of his/her activities so that they add to 24 hours are likely to introduce nonclassical measurement error in H^{CPS}. Reported work hours, H^{CPS}, may be related to salience, as suggested by Akerlof and Yellen (1985) in their analysis of measurement error in CPS responses about labor-force status in the previous year compared to contemporaneous measures, and as discussed more generally in the survey by Bound *et al* (2001, Section 5).

These different sources of measurement error have also been considered in the literature examining the reliability of information from family expenditure surveys, such as the U.S. Consumer Expenditure Surveys. Like the ATUS used here, the two main approaches to measurement are based on either longer-term recall or short-recall diary methods. Analogous to our discussion, this literature considers how reliability generated by short-term diaries, precisely specified reporting periods and expenditure items, and an aggregate budget constraint (based on reported cash flow or total spending) relative to longer-term recall yield higher quality data (see the essays in Carroll *et al*., 2015).

To develop intuition on the nature of the measurement error, we make the strong assumption that the time-use diary reliably measures work hours.[6] The measurement error in h^{CPS}, denoted by $e^{CPS} = h^{CPS} - h^{A}$, leads to:

(2) $\text{Bias}(\hat{\delta}^{CPS}) = (Z'Z)^{-1}E[Z'e^{CPS}]$,

where $Z = [w, X]$ and $\delta = [\alpha, \beta]'$. The effect of nonclassical measurement error in h^{CPS} is analogous to an omitted variable, leading to bias in the estimation of α. To simplify further, assume a bivariate relationship between h^{CPS} and w (equivalently, no linear relationship between w and X), then the bias in $\hat{\alpha}^{CPS}$ can be expressed as:

(3) $\text{Bias}(\hat{\alpha}^{CPS}) = \rho_{we}\frac{\sigma_{we}}{\sigma_{w}}$

Using this simple formula, the relative biases $\left(\frac{\hat{\alpha}^{CPS} - \hat{\alpha}^{A}}{\hat{\alpha}^{A}}\right)$ in estimating these labor supply elasticities in the samples of married workers are 0.316 (0.115) and 0.339 (0.920) for A1 and A2 respectively for married women (men), while in the samples of unmarried workers they are 0.247 (0.497) and 0.360 (0.768) respectively for women (men).

IV. Conclusion

In this note we have provided alternative estimates of elasticities describing labor supply using recalled usual weekly hours, and time worked from a one-day diary kept by the same respondents. Estimates using the latter measure are lower, which we argue results from measurement errors induced by the greater salience of recalled work time among higher-earning workers.

Measures of usual hours, recalled either for the previous week, or even for the past year (as in the major household longitudinal surveys), underlie nearly all studies of labor supply. Our

[6]We do not claim that this is strictly the case. Rather, we are demonstrating how the errors induced in H^{CPS} by the issues we have raised with the way this measure is obtained yield biased estimates compared to a measure that removes these errors.

results suggest that the nature of surveys producing data on usual work hours creates measurement errors that are positively correlated with the respondent's wage rate. Using these measures to estimate elasticities of labor supply then generates upward biases. The differences we have shown between these biased estimates and others that avoid this difficulty are not huge, but they do suggest some care in interpreting standard estimates of labor-supply elasticities.

REFERENCES

George Akerlof and Janet Yellen, "Unemployment through the Filter of Memory," *Quarterly Journal of Economics*, 100 (Aug. 1985): 747-73.

Olivier Bargain, Kristian Orsini and Andreas Peichl, "Comparing Labor Supply Elasticities in Europe and the United States: New Results," *Journal of Human Resources*, 49 (Summer 2014): 723-838.

Richard Blundell, Luigi Pistaferri and Itay Saporta-Eksten, "Consumption Inequality and Family Labor Supply," *American Economic Review*, 106 (Feb. 2016): 387-435.

-------------------- and Thomas MaCurdy, "Labor Supply: A Review of Alternative Approaches," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, vol. 3, 1559-1695. Amsterdam: North-Holland, 1999.

John Bound, Charles Brown, Greg Duncan and Willard Rodgers, "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics*, 12 (July 1994): 345-68.

John Bound, Charles Brown and Nancy Mathiowetz, "Measurement Error in Survey Data," in James Heckman and Edward Leamer, eds., *Handbook of Econometrics*, vol. 5. Amsterdam: North-Holland, 2001, 3705-3843.

George Borjas, "The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias," *Journal of Human Resources*, 15 (Summer 1980): 409-23.

Christopher Carroll, Thomas Crossley and John Sabelhaus, eds., *Improving the Measurement of Consumer Expenditures*. Chicago: University of Chicago Press, 2015.

Laurens Cherchye, Bram de Rock and Frederic Vermeulen, "Married with Children: A Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure Information," *American Economic Review*, 102 (Dec. 2012): 3377-405.

Greg Duncan, and Daniel Hill, "An Investigation of the Extent and Consequences of Measurement Error in Labor-economic Survey Data," *Journal of Labor Economics*, 3 (Oct. 1985): 508-32.

Eric French, "The Labor Supply Response to (Mismeasured but) Predictable Wage Changes," *Review of Economics and Statistics*, 86 (May 2004): 602-13.

Jonathan Gershuny and John Robinson, "Workweek Estimate-Diary Differences and Regression to the Mean," Oxford University, Sociology Working Paper No. 2013-01, 2013

Dominique Goux, Eric Maurin and Barbara Petrongolo, "Worktime Regulations and Spousal Labor Supply," *American Economic Review*, 104 (Jan. 2014): 252-76.

Daniel Hamermesh, "When We Work," *American Economic Association: Papers & Proceedings*, 88 (May 1998): 321-5.

--------------------, Harley Frazis and Jay Stewart, "Data Watch: The American Time Use Survey," *Journal of Economic Perspectives*, 19 (Winter 2005): 221-32.

F. Thomas Juster and Frank Stafford, "The Allocation of Time: Empirical Findings, Behavioral Models and Problems of Measurement," *Journal of Economic Literature*, 29 (June 1991): 471-522.

Michael Keane, "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49 (Dec. 2011): 961-1075.

Mark Killingsworth and James Heckman, "Female Labor Supply: A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of Labor Economics*, vol. 1, pp. 103-204. Amsterdam: North-Holland, 1986.

John Pencavel, "Labor Supply of Men: A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of Labor Economics*, vol. 1, pp. 3-102. Amsterdam: North-Holland, 1986.

Tess Stafford, "What Do Fishermen Tell Us That Taxi Drivers Do Not? An Empirical Investigation of Labor Supply," *Journal of Labor Economics*, 33 (July 2015): 683-710.

Table 1. Descriptive Statistics, Working Time and Wages, Hourly Paid Women, and Men, CPS/ATUS 2003-12*

Variable:	CPS Usual Weekly Hours (H^{CPS})	ATUS Total Work Time (H^{A1})	ATUS Work Time Working (H^{A2})	Hourly Pay
WOMEN				
Married (N=3925):				
Mean	36.21	7.97	7.39	$15.49
S.E.	(0.16)	(0.04)	(0.04)	(0.15)
Correlation:				
ATUS Total	0.51			
ATUS Time Working	0.51	0.98		
Unmarried (N=4262):				
Mean	36.89	7.99	7.39	$12.79
S.E.	(0.17)	(0.04)	(0.04)	(0.11)
Correlation:				
ATUS Total	0.42			
ATUS Time Working	0.44	0.94		
MEN				
Married (N=3840):				
Mean	43.30	9.28	8.49	$18.18
S.E.	(0.16)	(0.04)	(0.04)	(0.16)
Correlation:				
ATUS Total	0.52	0.96		
ATUS Time Working	0.53			
Unmarried (N=3086):				
Mean	40.08	8.72	8.01	$14.05
S.E.	(0.20)	(0.05)	(0.04)	(0.14)
Correlation:				
ATUS Total	0.52			
ATUS Time Working	0.53	0.96		

*Here and Table 2, includes people ages 18-69, who reported positive usual weekly hours and minutes working on a weekday, working on a weekday, and whose hourly wage was greater than $5.

Table 2. Labor Supply Elasticities, Hourly Paid Women, and Men, CPS/ATUS 2003-12*

Dep. Var.:	Usual hours (h^{CPS})	ATUS Total (h^{A1})	ATUS Working (h^{A2})
WOMEN			
Married (N = 3925)			
	0.1158	0.0880	0.0865
	(0.0137)	(0.0212)	(0.0221)
R^2	0.0471	0.0278	0.0285
t-statistic: $\alpha^{CPS} = \alpha^{A}$		1.40	1.42
Unmarried (N = 4262)			
	0.1814	0.1455	0.1334
	(0.0156)	(0.0221)	(0.0228)
R^2	0.1312	0.0619	0.0581
t-statistic: $\alpha^{CPS} = \alpha^{A}$		1.61	2.11
MEN			
Married (N = 3840)			
	0.0651	0.0584	0.0339
	(0.0094)	(0.0188)	(0.0199)
R^2	0.0813	0.0366	0.0292
t-statistic: $\alpha^{CPS} = \alpha^{A}$		0.35	1.57
Unmarried (N = 3086)			
	0.1837	0.1227	0.1039
	(0.0156)	(0.0216)	(0.0232)
Adj. R^2	0.1556	0.0751	0.0678
t-statistic: $\alpha^{CPS} = \alpha^{A}$		2.99	3.66

*The equations also include a quadratic in age, indicators of race and ethnicity, and indicators for day of week, state and year.