The effect of earnings on housework: Pros and cons of HILDA's time use data items

Peter Siminski, PhD Candidate, University of New South Wales

Paper prepared for the ACSPRI Social Science Methodology Conference, December 10-13, 2006

Abstract

The HILDA survey includes several 'stylised' time use data items. In time use research, such data are generally regarded to be of lower quality than diary data (for which respondents record each episode of time use over a given period, typically over one or two days). However, this limitation needs to be considered alongside the advantages of the survey, particularly its panel dimension. These issues are discussed in the context of a replication exercise, which is in progress.

Using cross-sectional data from the 1992 Australian Time Use Survey, Bittman et al. (2003) examined the relationship between people's share of couple earnings and time spent in housework. They argued that a negative relationship would support 'exchange-bargaining theory', which predicts that greater relative earnings result in smaller contributions of housework due to economic power. On the other hand, a positive relationship would suggest that people may compensate for gender deviance in earnings by stringently conforming to gender norms in housework. They found a quadratic relationship for women (holding a number of characteristics constant including time spent in paid work). It was negative for the range where women earned between 0 and 50% of income, and positive if they earned between 50% and 100%. They found no significant relationship for men.

Bittman et al.'s analysis is reproduced as closely as possible using HILDA. The results are quite similar, with a quadratic relationship found for women and no statistically significant relationship found for men. This suggests that the limitations of 'stylised' time use data are not critical for the present investigation.

A second, preferred model is estimated, which differs from the replication in a number of ways, fully utilising the strengths of HILDA. This model includes fixed effects and current income (rather than annual income) and additional control variables. The measure of housework includes household errands. For this model, an approximately linear relationship is apparent for both sexes, thus supporting exchange-bargaining theory across the full range of relative earnings. As a woman earns a greater share of income, not only does she seem to do less housework, but her husband does more. This is quite a different result to published investigations into this issue.

The main factor leading to the differing results appears to be the inclusion of fixed effects in the model. This is interpreted to suggest that the cross-sectional findings (for both sexes) might at least partially result from unobserved heterogeneity. Specifically, there may be unobserved characteristics that are positively correlated

both with (high) earnings and with (high) housework. Such an omitted variable is perhaps related to vitality, energy, work ethic or other characteristics.

The results concerning the substantive issue of housework should be treated as preliminary. The main conclusion of the paper is that researchers should consider HILDA and similar panel surveys as an alternative data source for analysis of time use.

Introduction

The HILDA survey (and numerous similar panel surveys) includes a simple time-use component, with a small number of 'stylised' time use questions. Other surveys rely on time-diaries, which are more complex and yield richer data. Time-diary data are generally regarded to be of higher quality than stylised time data. However, the limitations of the HILDA data need to be considered alongside the advantages of the survey, particularly its panel dimension. These issues are discussed in the context of a replication exercise, which is in progress.

The relationship between share of income earned and housework by members of couple families has been the topic of numerous studies (Bittman et al., 2003; Brines, 1994; Greenstein, 2000). Following Bittman et al., the theoretical literature can perhaps be summarised as presenting two competing hypotheses of the relationship between income and housework within couple families.

'Exchange-bargaining' theories predict that when an individual's earnings increase as a proportion of total couple earnings, they will make smaller contributions of housework due to the economic power they hold. In other words, as the share of a couple's income earned by a wife increases, she will do less housework, while her husband will do more.

The second is termed 'gender display' or 'deviance neutralisation'[1], which predicts that individuals may *compensate* for gender deviance in earnings by stringently conforming to gender-norms in housework. In other words, as the share of couple income earned by a wife increases, she will do more housework, while her husband will do less.

The empirical literature, however, has produced results that are perhaps intuitively unsatisfying. The paper by Bittman et al. (2003) is perhaps the seminal paper on this topic. They addressed the issue using the 1992 Time Use Survey conducted by the Australian Bureau of Statistics. Their modelling approach was OLS multiple regression of cross sectional micro-data. They found that, for wives, the relationship was strong and quadratic (holding a number of characteristics constant including time spent in paid work). The relationship was negative for the range where wives earned between 0 and 50% of income, thereby conforming to exchange-bargaining theory. And it was positive in the range between 50% and 100%, thus conforming to the

[1] The two terms actually describe different theories, but they predict the same relationship between the key variables of interest, and are hence treated as homogenous for the present purposes.

gender-display hypothesis. They found no statistically significant relationship between these variables amongst husbands.

Bittman et al. (2003) repeated their own analysis as closely as possible for the U.S, using data from the National Survey of Families and Households. This analysis suggested a very different pattern. Strong evidence was found for a linear relationship amongst women, supporting the exchange-bargaining hypothesis across the whole range of relative earnings. They also found some evidence ($p<0.10$) for a quadratic relationship for men. These results for the U.S. are consistent to those found in previous studies (Brines, 1994; Greenstein, 2000). Bittman et al. conclude that the results reflect genuine cross-national differences.

The primary aim of this paper is to discuss the strengths and weaknesses of HILDA, whilst replicating the Bittman et al. study using the first four waves of HILDA. The next section describes the HILDA time use data, and the following section summarises their strengths and weaknesses. The subsequent two sections describe the methods used for the HILDA analysis and the results produced. The main conclusions are presented in the last section.

Time use data in HILDA

The time use component of HILDA is contained in the Self-Completion Questionnaire. In Waves 2-4, each respondent was asked to indicate how much time they spend on each of nine activities in a typical week (there were only seven items in Wave 1). The relevant section from the Wave 4 instrument is reproduced in Figure 1, below. The corresponding sections from Waves 2 and 3 are identical. There are several differences, however, between these questions and those asked in Wave 1:

- The question on paid employment was not asked in Wave 1. But a similar question was asked in the Person Questionnaire: "Including any paid or unpaid overtime, how many hours per week do you usually work in all your jobs?"[2] That question was also asked in Waves 2-4. The correlation between these two items is high (0.89 amongst the sample of relevance to this study). Since the question asked in the Person Questionnaire is more detailed than the one in the Self-Completion questionnaire, it might be of a higher quality.
- In Wave 1, the question on travelling to and from employment was presented last in the sequence.
- The two questions on playing with children (labelled f and g in Figure 1) were asked as one combined question in Wave 1.

[2] Respondents who indicated that their hours vary were prompted with an alternate question: "Including any paid or unpaid overtime, how many hours per week do you usually work *on average* over a usual 4-week period in *all* your jobs?" [*italics* in original]

Figure 1 Time Use component of Wave 4 HILDA questionnaire

B17 How much time would you spend on each of the following activities in a typical week?

	IMPORTANT: • *Please do not count any activity twice* • *If you do not do an activity, write "0" in the hours box*	Hours per week	Minutes (if applicable)
a	Paid employment		
b	Travelling to and from a place of paid employment		
c	Household errands, such as shopping, banking, paying bills, and keeping financial records (but do not include driving children to school and to other activities)		
d	Housework, such as preparing meals, washing dishes, cleaning house, washing clothes, ironing and sewing		
e	Outdoor tasks, including home maintenance (repairs, improvements, painting etc.), car maintenance or repairs and gardening		
f	Playing with your children, helping them with personal care, teaching, coaching or actively supervising them, or getting them to child care, school and other activities		
g	Looking after other people's children (aged under 12 years) on a regular, unpaid basis.		
h	Volunteer or charity work (for example, canteen work at the local school, unpaid work for a community club or organisation)		
i	Caring for a disabled spouse or disabled adult relative, or caring for elderly parents or parents-in-law		

Strengths and limitations of HILDA time use data

Limitations

1. *Few categories of time*

An obvious limitation is that time spent was only measured for nine separate categories of activity (eight in Wave 1 including the paid work item). This is far fewer than in the ABS Time Use Surveys. This imposes a limit to the types of enquiries that the data will support.

2. *'Stylised' data*

There are two main ways in which time use data is collected. These are time diaries (such as the ABS Time Use Surveys) and survey questions (such as HILDA). The time diary method requires respondents to record every episode of time use in a given day or series of days, including the nature of the activity, it's start and end time. The survey method usually asks people to record the length of time spent on various activities in a 'typical' day or week, and is hence sometimes dubbed as 'stylised' time use data. The time diary method is generally regarded as being of higher quality than

survey data, but it is more complicated and expensive to conduct. It provides a much richer source of data, which can be analysed in a variety of ways.

Several studies have found that in survey questions, women report their housework time more accurately than men. On average, men tend to overestimate their time spent in housework more so than do women (Baxter and Bittman, 1995; Kan, 2006; Press and Townsley, 1998). This may reflect social desirability bias, to the extent that men desire to demonstrate a more egalitarian contribution to domestic labour. However, such bias is correlated with attitudes towards gender for both men and women. Women with traditional attitudes towards gender are less likely to overestimate their time spent in housework, while the opposite is true for men (Press and Townsley, 1998).

Whilst it is clear that survey data may produce biased estimates of summary statistics, the effect of such bias on regression analyses has been given less attention. Schulz & Grunow (2006) addressed this issue using a data source that had both diary data and survey data on time use. They ran a series of regressions, alternating between the two measures of housework as the dependent variable. They found that the regression results were very different between the two measures, and concluded that regression analysis is more susceptible to bias than summary statistics if one relies on survey data. However, the sample size for the regressions was probably too small for the conclusions to be reliable (a maximum of 139 observations). Regardless, it is important to be mindful of the strong likelihood of biased reports of housework in the HILDA data, and to consider the possibility that this might affect regression results.

Strengths

1. *Panel Data*

The greatest and most obvious strength of HILDA data is it's longitudinal dimension. By including fixed or random effects, one can hold constant all time-invariant individual characteristics. Two examples of the types of concerns that such an approach can navigate are given here.

As mentioned previously, Bittman et al. (2003) found a quadratic relationship between share of earnings and housework time for women. It is possible that there are unobserved characteristics of women which are positively correlated with (high) earnings and with (high) housework. Such an omitted variable is perhaps related to vitality, energy, work ethic or other characteristics. If so, this could explain the apparently quadratic relationship as a result of omitted variable bias. Similarly, such an omitted variable might also explain the absence of a relationship amongst men. More generally, omitted variable bias may be an important issue for cross-sectional time use analyses, especially since these are often characterised by low explanatory power (typically R-squared is around 10%). To the extent that such an omitted variable is time-invariant for each individual, it can be successfully controlled for using a fixed or random effects model.

Also as mentioned above, there is evidence that the bias inherent in survey estimates of housework time is correlated with attitudes towards gender. It is likely that gender attitudes are correlated with socioeconomic variables (including the proportion of

total couple income earned by the individual). This is another example of potential omitted variable bias, which can be navigated by a fixed or random effects model to the extent that the bias is time-invariant.

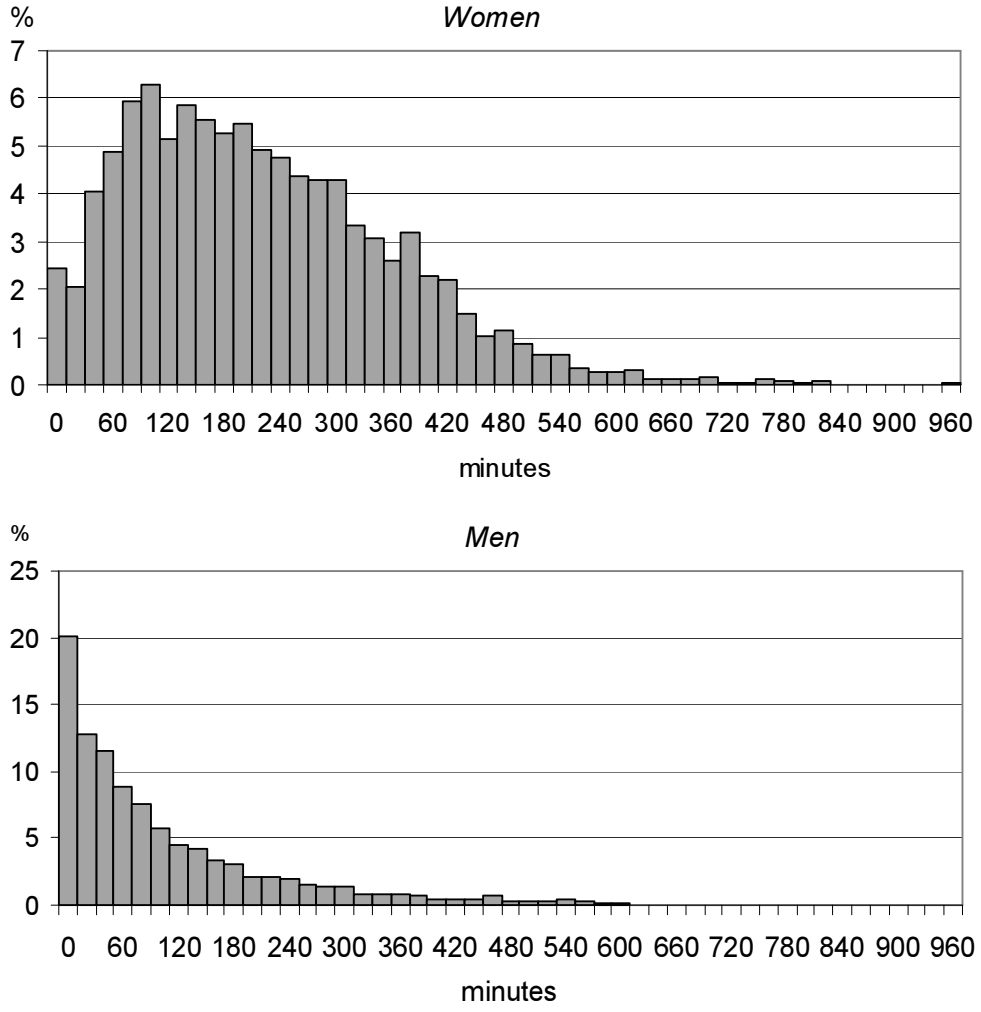
2. Few zero values

Episodic time-diary data is often aggregated to the diary-day level. Thus the object of analysis is often the total time spent by people on a given activity per day. The distribution of such data items typically has the following characteristics. It usually contains a large number of zero values, since it is rare for a given activity to be performed by every person on every day. As an example, Figure 2 shows the distribution of time spent on housework by women and men aged 55 or younger who are members of couple families, from the 1992 ABS Time Use Survey.[3] About 2.4% of women's records and 20% of men's records have zero housework time.[4] The corresponding distributions from the first four waves of HILDA are shown in Figure 3. These have very few zeros (0.1% of women and 0.7% of men have zeros recorded). There has been some debate as to whether zeros in time use data should be treated as censored. An emerging consensus is that they should not and that methods such as Tobit regression might not be appropriate (Brown and Dunn, 2006; Gershuny and Egerton, 2006; Stewart, 2006). If, however, one takes the view that such zeros should be treated as censored, then the relatively few zeros in the HILDA data can perhaps be seen as a negligible proportion of the sample.

[3] The distribution is very similar for the 1997 Time Use Survey.

[4] Note that the left-most bar in both charts represents genuine zeros, while the second bar actually represents 1-20 minutes, the third bar represents 21-40 minutes, and so on.

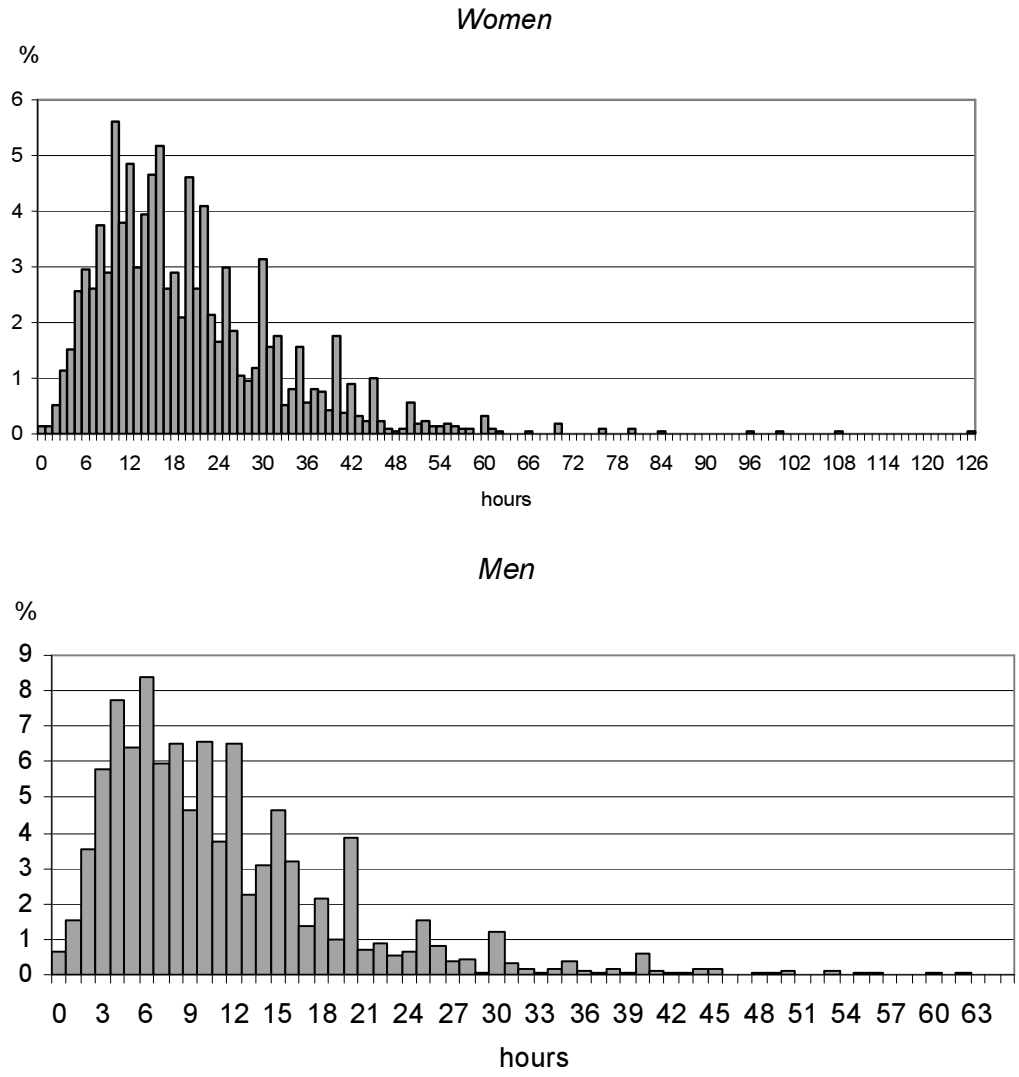
Figure 2 – Distribution of housework time – TUS 1992 (% of diary days by sex)[a]



(a) Restricted to members of a couple, aged 55 or under

Source: Author's calculations from ABS TUS 92 CURF

Figure 3 – Distribution of usual weekly housework time – HILDA Waves 1-4 (% of persons by sex)[a]



(a) Restricted to members of a couple, aged 55 or under

Source: Author's calculations from HILDA

3. *Clustering*

HILDA and the ABS Time Use Surveys are complex surveys, in which observations are clustered. The Time Use Surveys have three levels of clustering. Diary days are clustered within persons, which are clustered within households, which are clustered with geographic sampling units. In HILDA, observations are clustered within years (waves) and within persons. Persons are clustered within households, which are clustered within (initial) geographic sampling units.[5] Failure to take account of clustering will usually lead to underestimated standard errors, and hence the possibility of incorrect conclusions. In cross-sectional analysis, robust standard errors can be calculated if one has access to an identifier for the clustering variables, or through alternate means if replicate weights are provided on the file. Since neither replicate weights nor geographic sampling unit identifiers are provided on the publicly available ABS Time Use Survey CURFs for 1992 and 1997, it is not possible to accurately account for geographical clustering in those data. It is, however, possible to account for the other levels of clustering in those data, particularly through methods such as Generalised Estimating Equations (GEE), which takes account of correlations within clusters to estimate standard errors (Liang and Zeger, 1986).

For HILDA, the story is simpler. A two-way fixed effects model will 'automatically' produce standard errors that are robust to clustering.

4. *Income*

HILDA also has the advantage of more detailed income data items, which are useful for the present investigation. Most importantly, it includes current income data, while the ABS TUS only has previous year's financial year income. It seems reasonable to assume that one's current time spent on housework will be more closely related to current income than to the previous year's income. Many people will have changed their income earning activities in the interim. Some will have also changed their household composition.

5. *Sample size*

As the number of waves increases with time, the potential sample size for time use analysis in HILDA increases. In the present investigation, the sample size is between two and three times larger than the corresponding data in the ABS Time Use Survey.

Methods[6]

The empirical analysis is conducted in three stages using regression analysis of HILDA (release 4.1) data. In the first stage, the Bittman et al. (2003) study is replicated as closely as possible, utilising HILDA as a source of pooled cross-sectional data. In the second stage, an alternate (preferred) model is estimated, which

[5] In the present analysis, the person within household level is not relevant, since each regression includes only one member of a household.

[6] All analysis conducted using PROC GENMOD in SAS V9.1

differs from the initial method in a number of ways, as described below. Finally, an attempt is made to identify the methodological difference(s) that accounts for the discrepancy in the results.

Replication Stage

The main purpose of replicating Bittman et al. (2003) is to test the HILDA time use data items. Consistent results between the two studies would give some support to the suggestion that the HILDA time use items are not greatly affected by the weaknesses described above for the present analysis.

Following Bittman et al., separate regression models are estimated by OLS for men and for women. Only members of couples where both are under 55 years of age are included. Couples were excluded if either member's main source of income was anything but earnings, or if either member had a disability that impaired employment. All variables are specified in levels. The dependent variable is weekly hours of housework (including outdoor tasks, but not including household errands). The explanatory variables are:

- Wife's share of annual income (0-1);
- Wife's share of annual income squared;
- Combined annual income;
- Number of children aged:
  - 0-4 years
  - 5-9 years
  - 10-14 years
- Usual weekly hours of paid work (including commuting time)
- Spouse's usual weekly hours of paid work (including commuting time)
- Highest level of qualification:
  - Tertiary (1 = yes)
  - Trade (1 = yes)
  - High School (1 = yes)
- Age
- Husband unemployed (1 = yes)
- Husband out of labour force (1 = yes)
- Wife out of labour force (1 = yes)
- Evenly split income from self-employment (1 = yes)

Data from the four available waves were pooled, which necessitated some minor adjustments. Dummy variables (fixed effects) were added for each Wave (except Wave 1) to account for any general changes between years. Clustering of responses within persons was accounted for by GEE.

Apart from the diary-day versus stylised data issue of interest, the differences between Bittman et al. and this replication are:

- Approximately 10 years timing difference.

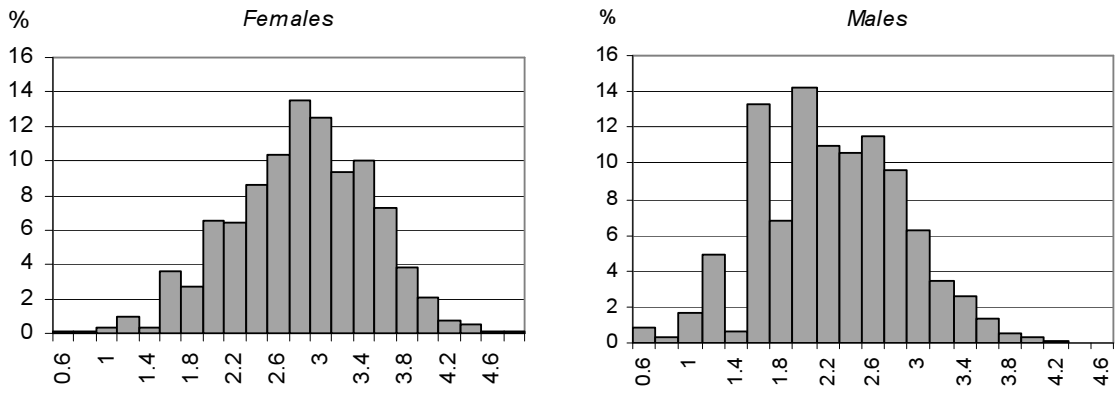
- Bittman et al. had variables for number of children aged 0-2 and 2-4, while in HILDA these are one variable (children aged 0-4).
- In HILDA, four cross-sections of data are pooled. Any correlation of responses within persons is accounted for through GEE, while overall differences between waves are accounted for through year-level fixed effects.

Preferred Model (interim)

The correct functional specification of the relationship between housework and the explanatory variables has not yet been given sufficient attention. Future versions of this investigation will address this in more detail, in a more rigorous theoretical context. Some observations are made here in the interim and the chosen functional specification is preliminary. But it is shown below that the results are not sensitive to a number of alternate specifications.

Bittman et al. assume that the relationship between housework and the explanatory variables is linear (in levels). This implies that a one-unit increase in an explanatory variable is associated with a fixed increase in the dependent variable. An alternate model might have a log-linear specification. In such a model, a one-unit increase in an explanatory variable is associated with a fixed percentage increase in the dependent variable. It is not clear whether either specification is an adequate estimation of the data generating process. The distribution of the dependent variable closely resembles a log-linear distribution, suggesting that a log-linear specification might be more appropriate. But such a distribution might also be consistent with a linear data generating process, depending on the distributions of the explanatory variables and their co-efficients in the population model. Without performing any formal specification tests, it is assumed for now that the log-linear model is a better approximation of the data generating process. However, a small number of observations take-on zero values for the housework variable (0.1% of records for women; and 0.4% for men), which complicates this assumption. Given their paucity, it might be justified to censor (set to a small positive number) or truncate (exclude) these observations, rather than to complicate the process. Both options are problematic and so a slightly modified specification was chosen. Rather than taking the log of the dependent variable, the left hand side of the model is specified as $\ln(y+2)$. Such a transformation is mentioned as a possibility by Gujarati (1995: 387). The distribution of the transformed variable is shown in Figure 4. The specification is thus similar to a log-linear specification, but the resulting coefficients will not have the same intuitive interpretation.

Figure 4 – Distribution of ln(housework + 2) – HILDA Waves 1-4 (% of persons by sex)[a]



(a) Restricted to members of a couple, aged 55 or under

Source: Author's calculations from HILDA

The preferred model differs from the replication in the following additional ways:

- The model includes person-level fixed effects
- The housework variable includes household errands (presumably mainly shopping). Bittman et al. excluded shopping from their measure in order to maximise comparability with previous studies, but their results were not sensitive to its inclusion
- Current income is used instead of annual income
- Couples were *not* excluded if their main source of income was anything but earnings.
- Couples were excluded if either member was self-employed in their main job. It is well known that the recorded income of the self-employed is an underestimation of associated living standards. See for example Bradbury (1996).
- Additional Control variables added:
  - Number of persons in household aged 15 + (minus two to account for the couple themselves);
  - Age squared;
  - Combined current income squared;
  - Remoteness (Major City; Inner Regional; Outer Regional; Remote/ Very Remote);
  - Dwelling Structure (house; semi; unit; other)
- The age variable (in levels) was not necessary as a control variable. This is because age is almost perfectly co-linear with the person and year fixed effects.[7]

[7] If each given person were interviewed on exactly the same date across years, these variables would be perfectly co-linear. Thus the age variable adds virtually nothing to the regression over and above the two-way fixed effects.

The preferred model thus takes the following form:

$$\ln(y+2) = b_0 + b_1 S + \beta X \qquad (1)$$

Where:
y is hours of housework performed per week including household errands;
S is the share of the couple's current income received by the wife (S squared was also added in some versions of the model to account for a non-linear relationship);
X is a vector of control variables including combined real income, combined real income squared, hours of paid work time per week of the person; hours of paid work time per week of the person's spouse; age squared; number of children in the household aged 0-4 years; 5-9 years; 10-14 years; number of persons in the household aged 15+ (minus two to account for the couple themselves); highest educational qualifications (tertiary, trade, high school, other); dummy variables for husband unemployed, husband NILF, wife NILF; remoteness (Major City; Inner Regional; Outer Regional; Remote/ Very Remote); Dwelling Structure (house; semi; unit; other); fixed effects for person and year.
β is a vector of coefficients of the control variables.

Interpreting the coefficients is not straightforward. Consider the following arrangement of equation (1) at y_0 and S_0:

$$y_0 + 2 = e^{(b_0 + b_1 S_0 + \beta X)} \qquad (2)$$

If S increases by c to the new value $S_1 = S_0 + c$, then the new value of y (y_1) is a function of c, the old value of y (y_0) and the co-efficient of S (b_1):

$$y_1 + 2 = e^{(b_0 + b_1 S_1 + \beta X)} = e^{[b_0 + b_1 (S_0 + c) + \beta X]} = e^{b_1 c} e^{[b_0 + b_1 S_0 + \beta X]} = e^{b_1 c}(y_0 + 2) \qquad (3)$$

A convenient way to simplify this is to focus on the case where y_0 is at its mean value (26.1 hours for women and 14.1 hours for men). Thus (3) becomes:

$$y_1 = 28.1 e^{(b_1 + c)} - 2 \qquad \text{(3a), for women; and}$$

$$y_1 = 16.1 e^{(b_1 + c)} - 2 \qquad \text{(3b), for men,}$$

where b_1 is estimated separately for men and women.[8] Consider the following example for women, assuming y_0 is at its mean. If S increases from 0 (its minimum value) to 1 (its maximum value), then y_1 increases from 26.1 to $28.1 e^{b_1} - 2$.

[8] The algebra is slightly more complicated similar where S-squared is also a regressor, but not prohibitively so.

Results

Replication

The regression results for the replication exercise are shown in Table 1, with and without the 'wife's share of income squared' term. The relationship between the variables of interest from Model 1 is plotted in Figure 5, (with all other explanatory variables held at their mean values if they are quantitative; or at their mode values if they are qualitative). Bittman et al.'s results are shown for comparison in Figure 6.

Table 1 Results of Replication exercise

	Women		Men	
	Model 1	Model 2	Model 1	Model 2
Intercept	12.162***	10.885***	8.877***	9.093***
Wife's share of income	-8.927***	-1.314	1.175	-0.074
Wife's share of income squared	8.985**		-1.469	
Combined real weekly income	-0.001***	-0.001***	0.000	0.000
Weekly hours of paid work	-0.183***	-0.188***	-0.068***	-0.068***
Weekly hours of paid work of spouse	0.077***	0.077***	0.052***	0.053***
Age	0.284***	0.289***	0.064***	0.064***
Number of children in household aged				
0-4 years	3.941***	3.915***	1.557***	1.562***
5-9 years	2.544***	2.566***	1.144***	1.140***
10-14 years	1.686***	1.680***	0.462**	0.463**
Highest Educational qualification (0 = none)				
Tertiary	-2.502***	-2.543***	-1.379***	-1.382***
Trade	-1.591***	-1.633***	0.015	0.013
High School	-1.606**	-1.600**	-1.503	-1.506***
Husband unemployed (1 = Yes)	1.551	1.774	5.233***	5.192***
Husband NILF (1 = Yes)	-0.547	0.037	8.670***	8.580***
Wife NILF (1 = Yes)	1.547*	1.918**	-0.185	-0.246
Evenly split self-employment income (1 = Yes)	2.459	2.141	-1.551	-1.491
R-squared	0.265	0.232	0.070	0.070
n	5985		6004	

*** Denotes statistically significant at 1% level ** 5% level * 10% level

Figure 5 Predicted hours of housework by share of couple income from replication exercise - HILDA

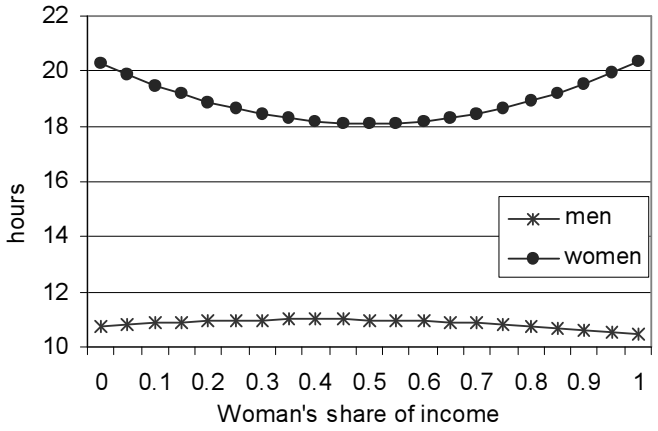
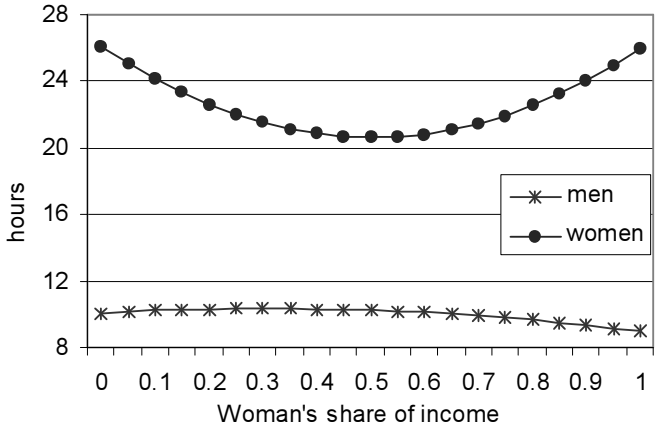


Figure 6 Bittman et al.'s results from 1992 TUS



The replication has produced quite similar results to Bittman et al. The relationship between income and housework appears quadratic for women ($p<0.02$). For men, it is not statistically significant quadratically ($p<0.57$), nor linearly ($p<0.93$). The main difference between these findings and those of Bittman et al. (2003) is that the female relationship is not as strong in HILDA. The parabola is approximately twice as deep in Bittman et al.'s results as in the present replication. Nevertheless, this exercise provides reasonably strong circumstantial evidence that the HILDA data are suitable for the present investigation. HILDA produces very similar results to the Time Use Survey when the same methods are used.[9] It has been argued, however, that these methods are not the best use of HILDA data. Attention is turned to the preferred specification.

[9] In future work it will be useful to seek further further confirmation by repeating the exercise using the 1997 Time Use Survey and/or the 2006 Time Use Survey when the data become available.

Preferred Model

The full regression results for the preferred model are shown in Table 2. The main results are quite different to those from the replication shown above. Share of income is not statistically significant in its quadratic form for either sex, but it is significant linearly ($p<0.03$ for each sex). Many of the variables that were statistically significant in the earlier models are not significant in these models. In some cases this may reflect their limited variation over time rather than the absence of a genuine relationship. For example, educational qualifications only change for a small proportion of people from year to year. A fixed effects model is not appropriate for measuring the effects of time-invariant factors.

Table 2 Regression results for preferred models

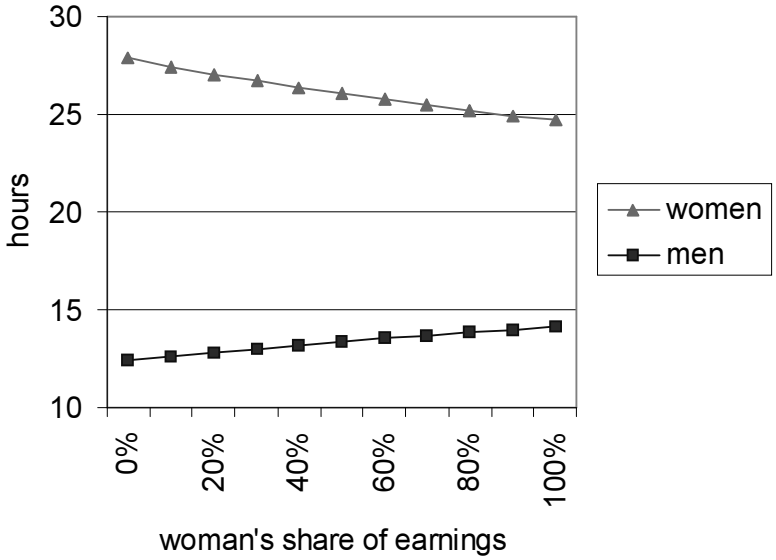
	Women		Men	
	Model 1	Model 2	Model 1	Model 2
Intercept	2.337***	2.339***	3.222***	3.223***
Wife's share of income	-0.136	-0.114**	0.148	0.115**
Wife's share of income squared	0.025		-0.037	
Combined real weekly income ($'00)	-0.002	-0.003	-0.004	-0.004
Combined real weekly income squared ($'00)	0.000	0.000	0.000	0.000
Weekly hours of paid work	-0.006***	-0.006***	-0.004***	-0.004***
Weekly hours of paid work of spouse	0.002***	0.002***	0.003***	0.003***
Age squared	-0.001***	-0.001***	-0.001	-0.001
Number of people in household aged				
under 5 years	0.120***	0.120***	0.023	0.022
5-9 years	0.077***	0.077***	0.005	0.005
10-14 years	0.027	0.027	0.036	0.036
15 years+ (excluding the couple)	-0.021	-0.021	0.033*	0.033*
Highest educational qualification (0 = none)				
Tertiary	0.060	0.060	-0.141	-0.141
Trade	-0.029	-0.029	-0.306**	-0.306**
High School	0.189*	0.189*	-0.145	-0.145
Husband unemployed (1 = Yes)	0.043	0.046	0.277	0.273***
Husband NILF (1 = Yes)	-0.014	-0.011	0.133**	0.128**
Wife NILF (1 = Yes)	0.093***	0.094***	0.004	0.001
Remoteness Area (0 = major city)				
Inner Regional	-0.006	-0.006	0.016	0.016
Outer Regional	0.000	0.000	0.066	0.066
Remote or Very Remote	0.034	0.034	0.039	0.039
Dwelling Structure (0 - house)				
semi-detached	0.045	0.044	-0.028	-0.028
appartment	-0.006	-0.006	-0.117***	-0.116***
other	-0.494***	-0.494***	-0.356*	-0.356*
R-squared	0.736	0.736	0.693	0.693
n	4932		5050	

*** Denotes statistically significant at 1% level ** 5% level * 10% level

The modelled relationships between housework and share of income are plotted in Figure 7. These plots are from Model 1 (those that include share of income squared as

an explanatory variable). The results for women and men almost mirror one another. They suggest that as a woman's share of couple income increases, her hours of housework decrease, while her husband's hours of housework increase. These relationships appear to hold across the whole range of the explanatory variable. A woman who earns all of the income is predicted to perform about 3.1 hours of housework fewer than if she were to earn none of the income (11% fewer). The corresponding effect for men is 1.7 hours per week, or 14% greater. The combined effect across genders is thus substantial.

Figure 7 Predicted hours of housework by share of couple income – preferred model (Model 2)



Sensitivity analysis and explanations for the discrepancy

The sensitivity of the main result was tested to a number of modifications. The results were not sensitive to:

- The use of a linear specification. The size of the modelled effect was similar for the two methods. The level of significance for share of income was weaker in the linear specification for males ($p<0.09$), but similar for females ($p<0.02$).
- The use of a log-linear specification with censoring: observations with less than two hours of housework were set to equal two (this applied to 0.3% of female records and 1.3% of males records). The size of the modelled effects and p-values were very similar.
- The use of a log-linear specification with truncation: observations with less than two hours of housework per week were excluded. The size of the modelled effects and p-values were very similar.
- The exclusion of household errands (including shopping) from the measure of housework. The size of the effects were very similar. The level of significance was weaker for men ($p<0.08$).
- To the use of cross-sectional weights in the regression. This strengthened the level of significance to $p<0.01$ for both sexes, and increased the size of the relationships.

The results were sensitive to:

- The use of person-level fixed effects. When these were excluded from the model, the share of couple income variable was not significant for either sex, either in quadratic or linear form (the p-values for the linear form were 0.12 for females and 0.56 for males). This suggests that omitted variable bias is likely to influence results that do not utilise fixed effects.
- For men only, the use of annual instead of current income for the share of income variable resulted in an insignificant linear relationship ($p<0.16$), though the estimated co-efficient was in the same direction. The result for women was not sensitive to this change.

Conclusion

It has been argued that there are a number of advantages to HILDA data over cross-sectional diary data, particularly for regression analysis. Judging by the empirical results, the greatest advantage appears to be its panel dimension, allowing for the use of fixed effects models. There are also disadvantages to 'stylised' time use data. For the present investigation, however, the replication exercise suggests that these disadvantages are not crucial. Of course this does not prove that longitudinal 'stylised' time use data are superior to cross-sectional diary data for every purpose. Further research would be required to test whether there are indeed any generalisations that can be made. It appears reasonable to suggest that researchers should consider HILDA (and similar surveys) as an alternative data source for time use analysis.

The specification of the preferred model is preliminary and the results should be treated with caution. These results suggest a neat and simple relationship between share of couple income and housework. They provide evidence to support exchange-bargaining theory for both sexes across the full range of the share of income. As a woman earns a greater share of income, not only does she seem to do less housework as a result, her husband does more. This is quite a different result to other investigations into this issue.

References

Baxter, J. and M. Bittman (1995) 'Measuring Time Spent on Housework: A comparison of two approaches' *Australian Journal of Social Research* 1(1): 24-46.

Bittman, M., P. England, N. Folbre, L. Sayer and G. Matheson (2003) 'When Does Gender Trump Money? Bargaining and Time in Household Work' *The American Journal of Sociology* 109(1): 186-214.

Bradbury, B. (1996) *Are the Low Income Self-employed Poor?*, SPRC Discussion Paper No. 73, Social Policy Research Centre, UNSW, Sydney.

Brines, J. (1994) 'Economic Dependency, Gender and the Division of Labour at Home' *American Journal of Sociology* 100(3): 652-88.

Brown, J. and P. Dunn (2006) *Too little time? A fresh look at right skewed data and how Poisson-gamma distributions may provide a viable solution*, International Association of Time Use Research Conference 2006, Copenhagen.

Gershuny, J. and M. Egerton (2006) *Evidence on participation and participants' time use from day- and week-long diaries: implications for modelling time use*, International Association of Time Use Research Conference 2006, Copenhagen.

Greenstein, T. N. (2000) 'Economic Dependence, Gender, and the Division of Labour in the Home: A Replication and Extension' *Journal of Marriage and the Family* 62: 322-35.

Gujarati, D. N. (1995) *Basic Econometrics*, New York, McGraw-Hill.

Kan, M. Y. (2006) *Measuring Housework Participation: The Gap Between "Stylised" Questionnaire and Diary-Based Estimates*, GeNet Working Paper No. 20, Institute for Social and Economic Research, University of Essex, Essex.

Liang, K. Y. and S. L. Zeger (1986) 'Longitudinal analysis using generalized linear models' *Biometrika* 73: 13-32.

Press, J. E. and E. Townsley (1998) 'Wives' and Husbands' Housework Reporting: Gender, class, and social desirability' *Gender & Society* 12: 188-218.

Schulz, F. and D. Grunow (2006) *Time Diaries and Survey Questions on Time Use. Different Approaches - Different Results?!*, International Association of Time Use Research Conference 2006, Copenhagen.

Stewart, J. (2006) *Tobit or not Tobit?*, International Association of Time Use Research Conference 2006, Copenhagen.