

Measuring Housework Participation: The Gap between “Stylised” Questionnaire Estimates and Diary-based Estimates*

**Man Yee Kan
Institute for Social and Economic Research
University of Essex**

**GeNet Working Paper No.20
April 2006**

*This paper is also released as Working Paper of Institute for Social and Economic Research, Paper 2006-11.

Direct correspondence to:

Man Yee Kan
Institute for Social and Economic Research
University of Essex
Wivenhoe Park
Colchester
CO4 3SQ
U.K.
Email: mkan@essex.ac.uk

Please do not cite without the author's permission

Acknowledgements

This paper forms part of the project “Gender, Time Allocation, and the ‘Wage Gap’” funded by the ESRC Gender Equality Network. It has been greatly improved as a result of my discussions with Professor Jonathan Gershuny. It has also benefited from comments received from my Time Use Group colleagues, participants at the International Association for Time Use Research Conference 2005, and seminar participants at the Institute for Social and Economic Research.

Abstract

This article compares stylised (questionnaire-based) estimates and diary-based estimates of housework time collected from the same respondents. Data come from the Home On-line Study (1999 – 2001), a British national household survey that contains both types of estimates (sample size = 632 men and 666 women). It shows that the gap between the two types of estimate is generally smaller in the case of women. But the gap between the estimates in the case of women is associated with the amount of housework performed as secondary activities and the level of irregularity in housework hours. Presence of dependent children, on the other hand, inflates the gap for both men and women. Men holding traditional gender-role attitudes tend to report more housework time in surveys than in diaries, but the tendency is reversed when they undertake long hours of housework. The overall results suggest that there are systematic errors in stylised housework time estimates.

Key words: housework, measurement error, time-diary method, time use

Measuring Housework Participation: The Gap between “Stylised” Questionnaire Estimates and Diary-based Estimates

The domestic division of labour has been a popular topic of social research in recent years. Research on this topic has mainly used estimates of housework time collected by time diary methods and direct questioning in survey interviews. This raises the question of comparability of the results derived from these two types of estimates, and of whether they need any qualifications because of the potential existence of systematic errors. It is widely recognized that questionnaire-based estimates are less accurate than time-diary ones because they contain more random errors (e.g., errors due to recalling problems). Nevertheless, relatively few research efforts have been devoted to finding out whether the direct questioning approach in surveys may also produce systematic errors in the estimates of housework time.

This article therefore aims to assess systematic biases in estimates of housework hours collected in survey interviews by comparing them with time diary information provided by the same respondents of a British survey (Home On-line Study, 1999 - 2001). In particular, it will investigate whether the mechanisms producing systematic errors, if any, are different for men and women. The implications of using survey questionnaire-based housework time estimates for the study of the domestic division of labour will be discussed in the light of the findings.

BACKGROUND

Methods of Collecting Housework Time Estimates

Direct questioning in surveys and time-diary methods are common instruments for collecting housework time estimates. These methods have their respective advantages and disadvantages, although it is established in past research

that diary-based estimates of time use are more reliable and accurate than estimates derived from direct questions (Plewis, Creeser & Mooney, 1990; Robinson, 1985)¹.

In survey interviews, respondents are asked how much time they have spent, for example, in the previous week or normally spend each week, on a particular activity. These estimates of time use are referred to as stylised estimates (Juster & Stafford, 1985). Stylised estimates are cheaper to collect than time diary estimates because they can conveniently be obtained in surveys where time use is only one of the several variables of research interest. The survey approach is also less demanding on respondents (e.g., in terms of time, initiative and energy) than diaries and therefore usually has a higher response rate than diary methods. But stylised estimates contain certain recall biases because it is difficult for a respondent to remember and report exactly the amount of time used. They also tend to produce total time greater than 168 hours for a week (Gershuny & Robinson, 1994), and are generally higher than time diary estimates (See for example, Bianchi, Milkie, Sayer, and Robinson, 2000, Table 4).

The time diary approach asks respondents to fill in activities and time used of the present/past day in slots of a diary. Compared with the survey interview approach, it is less dependent on the respondents' calculation and augmentation of time on different activities, and hence produces more accurate measures of time use. Another advantage of the time diary approach is that it can provide detailed information of the respondent's use of time on various activities during the day, the sequence of these activities and the contexts of performing such activities (e.g., with whom and where).

¹ Juster, Ono, and Stafford however held a different view. In the case of paid work time, they found that the two types of estimates correspond with each other closely for individuals who work regularly. When assessing historical trends, they reported that the two estimates correspond closely, but some time-diary estimates deviate from the trend even when the sample and the codes are standardized. See Juster et al., (2003, Figures 1a – 1d).

But completing a diary for many respondents is an onerous task and therefore the response rate is often low.

Indeed, to collect time use estimates, there is often a trade off between minimizing the burden placed on respondents and achieving a high response rate. At one extreme, the experience sampling method (ESM) places little burden on respondents and often achieves a high response rate. This method requires respondents to record what they are doing at randomly selected moments during the day, which are determined and alerted by an electronic beeper (Csikszentmihalyi & Larson, 1987). Because this method can obtain a sample of high representation of different times of the day and can create virtually no recalling problems to respondents, it usually provides very accurate measures of time use of a particular population group. But because of the relatively high cost of the electronic instruments, the samples in these studies are typically very small. Another major drawback is that it fails to provide an adequate basis to establish an individual-level time budget, i.e., the time use estimates obtained contain little information about intra-personal variation (Gershuny, 2004).

The Gap between Stylised Estimates and Diary Based Estimates

To examine how the gap between stylised estimates and diary-based estimates of housework time might be associated with different individual characteristics, it would be best if these estimates are obtained from the same respondents and are with comparable time frames. Earlier studies were limited by the lack of data sources that contain both types of estimates. It is particularly difficult to find comparable data of these estimates from the same respondents because most surveys have collected stylised estimates of *weekly* housework time, while many time-diary studies have collected only *one or two days* of diary records (usually one weekday diary and one

weekend diary) from each respondent. But a one-day diary provides less representative information about an individual's time budget than a one-week diary. Past studies hence tended to compare the two types of estimates from different data sources. This is less ideal than comparing data from the same source because the gap between the estimates may be due to a certain extent to the difference in samples and designs of the studies. Moreover, comparing estimates from different data sources provides little information about how the gap might vary with individual characteristics.

Comparing questionnaire estimates from the U.S. National Survey of Families and Households, 1992 – 1994, and diary-based estimates from a time diary sample in 1995, Bianchi et al. (2000) found that the questionnaire-based estimates of housework time are generally higher than the diary-based ones. This finding was confirmed in a recent study conducted by Juster, Ono, and Stafford (2003), which provided detailed comparisons between stylised estimates and diary estimates of housework hours and paid work hours from various years of U.S. data (1965 – 2001). Similar results were also reported by Baxter and Bittman (1995) using Australian data, Niemi (1993) using Finnish data, and Marini and Shelton, (1993) and Robinson (1985) using U.S. data. Some of these studies also suggested that the gap between stylised estimates and diary-based estimates is substantially larger in the case of women than in the case of men (Baxter & Bittman, 1995; Niemi, 1993; Robinson, 1985).

Despite the differences between the two types of estimates, these past studies suggested that the two methods often reveal roughly similar patterns of variation between different groups. They therefore concluded that, despite being less accurate and reliable than diary-based estimates, stylised estimates provide a useful ordinal scaling of individuals' time spent on housework and thus are useful for multivariate

analyses of the domestic division of labour (Baxter & Bittman, 1985; Marina & Shelton, 1993; Robinson, 1985).

Recent research however has shown that the magnitude of the gap between diary-based and stylised estimates varies among different social groups. Press and Townsley (1998), compared self-reported housework hours in a U.S. survey (National Survey of Families and Households, 1988) with values of diary-based information imputed from multivariate regression of time-diary data from another study (Americans' Use of Time, 1985), and found that the size of the gap depends on the gender of respondents, total housework hours, education and other socio-economic variables.

More recently, two studies compared the two types of estimates of housework hours from a single data source. But we should note that these studies were based on data collected from one-day diary records, which, as mentioned before, provide less ideal information for comparison with stylised housework time estimates than weekly diary records. Bonke (2005) compared time-diary estimates and stylised estimates of paid work time and unpaid household work time from the Danish Time Use Survey, 2001. Contrary to previous studies that used data from two different sources, he found that both men and women report less household work time in surveys than in their diaries, and in absolute terms, women “under-report” their housework hours to a greater extent than men. He also found that parents “under-report” their housework time more than non-parents, and older people “under-report” it more than younger people (where age was thought to be related to gender attitudes)².

Similarly, Kitterød and Lyngstad (2005) analyzed time-diary estimates and stylised estimates of housework from the Norwegian Time Use Survey, 2000 – 01 but

² The terms “over-report” and “under-report” are adopted from Bonke (2005). To be exact, we should note these are based on the assumption that the diary estimates are more accurate than the stylised ones.

found only modest differences between the estimates. Contrary again to previous studies, they discovered that the gap is associated significantly only with age groups, but not with the gender of respondents. They accounted for the discrepancy between their results with those of previous studies by the nearly disappearance of housewife role in Norway, “over-reporting of housework in direct questions because of social desirability and perceived pressures to do much housework is probably rather modest in Norway” (Kitterød & Lyngstad, 2005, p. 30). This view matches well with Bonke’s (2005) findings from Danish data, which indicate that men and women tend to report less housework time in surveys than in diaries. It might also be because in Denmark gender relations are generally more equal than many non-Scandinavian countries, and hence men and women experience relatively little pressure to over-report their housework time.

It will be interesting to find out whether men and women differ in their accuracy in reporting housework participation when the gender division of domestic labour in the country/region surveyed is highly unequal. Lee and Waite (2005) conducted a relevant study using U.S. data. They examined housework time derived from ESM estimates and that from direct questions of 265 married men and women with children aged 5 – 18, and found that both men and women inflate their housework time substantially in direct questions, with wives over-reporting more in absolute values. For example, the ESM estimate for wives equals 26 hours per week, whereas the ESM estimate is only 15 (the ESM estimates is 21.2 when housework as secondary activities are also counted). The sample used in Lee and Waite’s (2005) study is however not nationally representative and the ESM measures, as mentioned before, provide little information on intra-personal variation.

To examine systematic errors in stylised estimates of housework time, it will be best to compare them with diary-based estimates derived from information collected from the same respondents. It will also be important to estimate diary-based housework time from one-week diary-records rather than one-day records in order to capture variations at the individual level. Furthermore, it is expected that systematic biases in stylised estimates of housework may differ according to the gender relations and values of the country where the survey was conducted.

Mechanisms to Explain the Gap

First, the gap between stylised questionnaire estimates and diary-based estimates of housework hours can be due to differing conceptions of housework. In many cases, housework is not a clearly and consistently defined concept. Unlike gainful employment work, which can readily be defined as the work done for the exchange of labour market income, housework is unpaid and may be defined as different sets of household tasks by the respondent and by the researcher. In diary-based methods, the researcher plays a stronger role in defining housework than in surveys, typically by setting pre-coding activities for respondents to fill in and then by summing relevant activities to calculate the total time spent on housework. In survey interviews, on the other hand, the researcher describes briefly the activities that are defined as housework (in many surveys, it includes routine and labour costing household tasks, such as cleaning, cooking, washing and doing the laundry, and so on), and the respondent then comprehend the question wording and estimate the time spent on the defined housework accordingly in a short time. Some common household tasks that are not normally defined as routine housework in surveys, such as care of family members, DIY and gardening, but might be counted by some respondents when they report their housework time in surveys. The discrepancy

between the respondent's definition and the researcher's depend on the clarity of the survey questions as well as the respondent's attentiveness and comprehension.

Moreover, it might well be the case that the discrepancy varies across different population groups, such as women and men, depending on the levels of their usual participation in household tasks. For example, research has shown that women on average spend more time on childcare than men (Gershuny, 2000; Robinson, 1997) and therefore the gap caused by mixing up childcare and housework may affect women more than men.

The second factor accounting for the gap is inaccuracy in the respondent's estimation. Owing to the limitations of human memory, a difference between the actual time spent and that recalled by the respondent is almost unavoidable. But the gap should also be considerably larger when the respondent has irregular patterns of housework participation, which makes accurate estimation over a week/month more difficult. In the case of paid work hours, Robinson (1997) analyzed U.S. data and reported that the gap between the two types of estimates is larger when people have long work hours and irregular work patterns. More recently, Gershuny (2005) found from British data that the size of discrepancy between the two kinds of estimates is associated with the degree of variation in one's paid work hours across different days of a week. Because the time diary method requires fewer efforts from the respondent in calculating the time than the survey approach, it should measure the time spent on housework more accurately.

Regarding the problem of inaccuracy in estimation, it has also been suggested that women should report their housework hours more accurately than men, since they usually undertake the bulk of housework and should therefore be more familiar with how much time each of the chores has cost them; men, on the other hand, tend to be

oriented towards labour market work and thus pay less attention to housework (Fenstermaker Berk & Shih, 1980; Press & Townsley, 1998; Warner, 1986).

Furthermore, it is proposed that respondents are likely to include the time when housework is undertaken simultaneously with one or more other activities at home, and therefore stylised estimates of housework tend to be higher than diary-based estimates (Juster & Stafford, 1991; Marini & Shelton, 1993). Following this argument, we would expect to find that the bias in estimates due to double counting of concurrent tasks should affect women more than men because women perform more than one housework task simultaneously more often and are more likely to report housework as secondary activities than men (Lee & Waite, 2005).

Finally, Press and Townsley (1998) pointed out that the discrepancy between the two types of estimates of housework hours could not be fully explained by the double counting of multiple tasks and inaccuracy of memory. They suggested that reporting housework hours in survey interviews is a gendered process and the role of social desirability plays a key role in explaining the gap between the estimates. The domestic division of labour in the U.K. and many other countries are still highly gendered, with women undertaking the major share of it regardless of their employment status (e.g., Bianchi et al., 2000; Laurie & Gershuny, 2000). Some respondents might feel the pressure to report a level of housework participation that agrees with the normative gender roles, should their actual participation depart from what their gender role attitudes enjoin. This view parallels a burgeoning area of research that focuses on the function of housework in fulfilling one's gender identity. It is suggested that couples that violate gendered expectations in labour market earnings (e.g., cases where the husband is dependent on the wife financially) tend to compensate their "loss" in gender identities by adopting a traditional gendered

division of household labour (Brines, 1994; Greenstein, 2000; Bittman, England, Folbre, Sayer, & Matheson, 2003). It will be interesting to examine whether the process of reporting housework participation in surveys itself is gendered, and if so, whether this has any implications for research on the relationship between gender traditionalism and the domestic division of labour.

It is expected that the influence of social desirability will depend on one's attitudes towards gender roles. The data in Press and Townsley's (1998) study came from two separate sources (one provided the diary-based estimates and the other provided the stylised estimates) and did not contain any direct measures of gender role attitudes. The authors therefore could only use imputed data of reported housework hours and tap gender role attitudes by the respondent's age and educational level. Based on these indirect measures, they found that young and highly educated men, whom they argued were socially expected to exercise gender egalitarianism in the domestic division of labour, were more likely to over-report their housework participation than other men.

Hypotheses

The present study will test the following hypotheses:

H1: Concerning the mechanism producing inaccuracies in estimation, it is hypothesized that women will report their housework hours more accurately than men (i.e., the gap between the estimates will be smaller in the case of women). Because women on average spend far more time on housework and are more specialized in housework than men, they should have better knowledge of how much time household chores have cost them.

H2: Also, it is hypothesized that when variation in housework hours during a week increases, the gap between the stylised housework hours and the diary-based

hours will become larger. It is because irregularity in housework patterns should have made it more difficult for the respondent to recall and estimate their housework hours.

H3: Considering the mechanism of double counting of simultaneous activities, it is expected to find that the gap between the two types of estimates of housework hours will be larger the more often the respondent undertakes one or more other activities together with housework. That is, the reported housework hours will increase with increases in the time spent on housework as a secondary activity. It is also expected that the effect is greater on women than on men, since women are more likely to undertake housework as secondary activities.

H4: The gap in the estimates is expected to be greater when dependent children are present in the household. It is because the respondent might confuse childcare activities with housework (and hence causing greater inaccuracy in the estimation), and childcare is frequently carried out simultaneously with housework (and thus housework is likely to be double-counted). Again, it is expected that the influence of the presence of children is greater on women than on men, since women are mainly responsible for childcare.

H5: With regard to the mechanism of social desirability, it is hypothesized that when gendered expectations are violated, e.g., when a traditional man spends long hours on housework or a traditional woman undertakes little housework, the respondent tends to report their housework time inaccurately in order to be consistent with their gender role attitudes. That is, there will be an interaction between gender role attitudes and housework hours.

ANALYTICAL STRATEGIES

To test the above hypotheses, I will compare stylised estimates of housework hours with diary-based estimates first by simple cross-tabulations and then by multivariate analysis.

Data

The data used in this study come from the Home On-line Study, which consists of three annual waves of household panel data (1999 – 2001). A distinctive advantage of this study is that it contains both stylised estimates and diary-based estimates of time spent on housework. It also surpasses other time budget studies because it collected seven-day diaries from respondents, while other studies usually collected only one- or two-day diaries. It interviewed about 1,000 households drawn from a national random sample in Great Britain. It was originally intended for the estimation of time-use patterns as a result of the everyday use of information-and-computer technology, and therefore has an over-sample to make sure that 50% of the households have a personal computer. Individuals aged 16 or over in the selected households were interviewed in all waves. A one-week self-completion diary designed to record what respondents' were doing each day of that week every quarter hour of a day was given to the respondents after the interview. They were asked to fill in the diary with 35 pre-coded activities at least once each day and then return it at the end of the designated week. The Home On-line Study collected around 2,300 diaries (i.e., 16,100 diary days) from respondents in all the three waves.

The sample selected for the present study includes married and cohabiting men and women, since past studies mainly used the housework estimates to study of the domestic division of labour of marital couples. It is also restricted to cases where both questionnaire and diary records are present and where the missing time of the

total weekly diary records is less than 3.5 hours (i.e., only completed or fairly completed diaries are included). All cases are weighted to adjust for diary non-response. The final sample for analysis contains 1,298 cases (632 men and 666 women).

Measures

1. Stylised questionnaire estimates of housework hours

In the survey part of the study, respondents were asked the following question: “About how many hours do you spend on housework in an average week, such as time spent on cooking, cleaning and doing the laundry?” Just next to this question, respondents were asked how much time they spent on DIY and gardening in an average week. From the survey question, housework is referred as routine and labour consuming household work, but excludes occasional chores and care. The respondent’s reported housework hours is taken as the dependent variable in the multivariate analysis.

2. Diary-based estimates of housework hours

To be consistent with the questionnaire question, housework time derived from diary records includes time spent on routine housework, such as cleaning and washing and cooking but excludes DIY, gardening, and care. The diary part of the study asked respondents to fill in their main activity and, if any, their secondary activity in a given slot of the diary. The total time spent on housework when housework is a primary activity and that when it is a secondary activity on all seven days of a week are summed up and are introduced as two separate independent variables in the regression analysis.

It should be noted that the time referents of the diary records and the survey question are not exactly the same. The notion of “average week” in the survey

question might have confused respondents, and might have given them scope to portray their activities in a socially acceptable way. On the other hand, the diary days were randomly sampled, and therefore might contain some random differences from normal days. Despite these, I will show in a separate paper that the diary estimates contain less measurement error variances than the stylised estimates, i.e., the diary estimates are more accurate than the stylised estimates³. It would be important, however, to bear in mind that some of the gap might be due to difference in the time referents when we interpret the findings.

3. Irregularity in housework hours

The irregularity of housework time over the week is measured by the coefficient of variation (α):

$$\alpha = SD \text{ of Duration} / M, \text{ if } M > 0;$$

$$\alpha = 0, \text{ if } M = 0$$

where *Duration* = *Length of housework time during a day* = *Starting time* - *End time*;

M = *Sum(Duration)/Number of days doing housework*.

Here only housework coded as primary activities is taken into account.

4. Traditionalism in gender role attitudes

Gender role attitudes are measured by four items about women's and men's roles in the family. Respondent were asked about their opinions on each item in an agree/disagree format. Each of the four items is measured by a five-point scale and responses to them are recoded and then added up to create a score indicating the respondent's degree of traditionalism in gender role attitudes. The score ranges from

³ In a forthcoming paper, the stylised and the diary-based estimates of housework time will be treated as the dependent variables in two separate equations, which will be modeled by the method of "seemingly unrelated regressions" (Zellner, 1962). The results show that the diary estimates contain less measurement error variances than the stylised estimates. For more details, see (Kan & Pudney, forthcoming).

0 to 16, where higher values indicate more traditional attitudes and 8 is neutral. Responses were then classified into two groups: *Traditional*, where the score is higher than 8, and *Non-traditional*, otherwise. Details of the questions and of the coding of this variable are given in Appendix I. Here the focus is placed on the potential influence of gender traditionalism on the reporting of housework hours in surveys. This variable is included because research on the domestic division of labour has suggested traditional men are prone to “do gender” by undertaking less housework when their orthodox breadwinner role is not fulfilled (Bittman et al., 2003; Brines, 1994; Greenstein, 2000). It is possible that the process of “doing gender” operates later in the reporting of housework hours.

Methods

One issue is whether to include in the analysis cases where the respondent’s reported housework hours is zero. Research focusing on the discrepancy between stylised and diary-based estimates of paid work hours usually excludes cases where the respondent has reported zero work hours, on the ground that there should have been no recall error when the respondent has not participated in labour market work (see Gershuny, 2005; Robinson, 1997). The case of housework is however more complicated than that of paid work. Housework is not as clearly and consistently defined as paid work. Some people may have performed and recorded some housework chores, but reported none in the survey. Moreover, unlike paid work, housework can readily be performed as a secondary activity (e.g., one may read a novel and monitor the operation of a washing machine at the same time). The line between little and absolutely no housework participation is often blurred. Furthermore, given the difference in time referents of the diary and the survey, as mentioned before, it might well be the case that some respondents usually undertake

some housework in a normal week but happened not to have done any during the week of completing the diary. Hence it is theoretically possible for a respondent to report some participation in the survey but recorded zero housework hours in the diary. Empirically, 92 men and 3 women in the sample (14.6% of men and 0.5% of women) recorded zero housework hours in their dairies. Of the 92 cases of male respondents, 40 (43.8%) recorded non-zero housework hours in the secondary activities of their dairies and 72 (78.5%) reported non-zero housework hours in the survey interviews. In fact, only in a handful of cases ($n = 10$, 1.6% of the male respondents sample) did the respondents show zero housework hours both in the survey and in the primary as well as secondary activities of their dairies. Therefore cases where respondents recorded zero housework hours in the diary will be included in the analysis.

OLS regression will be employed as the main instrument to test the association between reported housework hours and housework hours recorded in dairies. Men's and women's hours of housework reported in the survey will be modeled separately (since the mechanisms for producing systematic biases in them can be different) by the following equation:

$$\begin{aligned}
 y_i = & \beta_0 + \beta_1 \text{Primary Housework}_i + \beta_2 \text{Secondary Housework}_i + \beta_3 \text{Variation}_i \\
 & + \beta_4 \text{Child}_i + \beta_5 \text{Traditional Attitudes}_i + \beta_6 \text{Primary Housework}_i * \text{Variation}_i \\
 & + \beta_7 \text{Primary Housework}_i * \text{Child}_i \\
 & + \beta_8 \text{Primary Housework}_i * \text{Traditional Attitudes}_i + \varepsilon_i
 \end{aligned}$$

where for a respondent i , y_i is a dependent variable indicating the weekly housework hours reported by the respondent in the survey. *Primary Housework* is a continuous variable denoting the respondent's hours on housework recorded in the diary, and housework was reported as a primary activity. *Secondary Housework* denotes the

number of hours on housework when housework was recorded as a secondary activity. *Variation* is the coefficient of variation of housework hours in the week of completion of the diary. *Child* and *Traditional Attitudes* are binary variables, indicating respectively whether one or more children under the age of 16 were present in the household (*Yes* = 1; *No* = 0) and whether the respondent held traditional attitudes towards gender roles (*Yes* = 1; *No* = 0). β_0 to β_8 are parameters to be estimated. ε_i is an error term.

RESULTS

Descriptive Figures

Let us first focus on some descriptive figures that depict the difference between stylised and diary-based housework hours. Figure 1 compares the stylised and the diary-based weekly housework hours of men and women. The distribution of housework hours by gender is consistent with findings of past research on the domestic division of labour: Women as a group undertake far more housework than men, regardless of whether diary-based or questionnaire-based estimates are used. Concerning the gap between the two kinds of estimates, the results are nevertheless different from those of past studies. Women do not appear to over-report their housework hours more than men. In contrast, the gap between the two types of estimates is smaller in the case of women than in the case of men. When only primary activities are taken account of, the difference is 0.79 hour for women and 1.96 hours for men. As has been expected, the gap becomes narrower when secondary activities are taken account of, suggesting that some respondents tend to count the housework that is undertaken as a secondary activity. The figures are 1.02 hours for men and -.49 hour for women (the negative value shows that some women

actually report less housework time in surveys than in diaries when secondary activities are taken into account). In other words, assuming that diary-estimates of housework hours are more accurate than stylised questionnaire estimates, the results show that men, as a group, over-report their housework hours more than women. The gap is even more obvious when compared with the total housework hours: It constitutes 30.3% of men's reported hours (only primary activities are used to estimate the diary-based hours) but only 4.9% of women's reported hours.

Figure 1. Participation in housework per week by Gender, Stylised Vs Diary-based Estimates

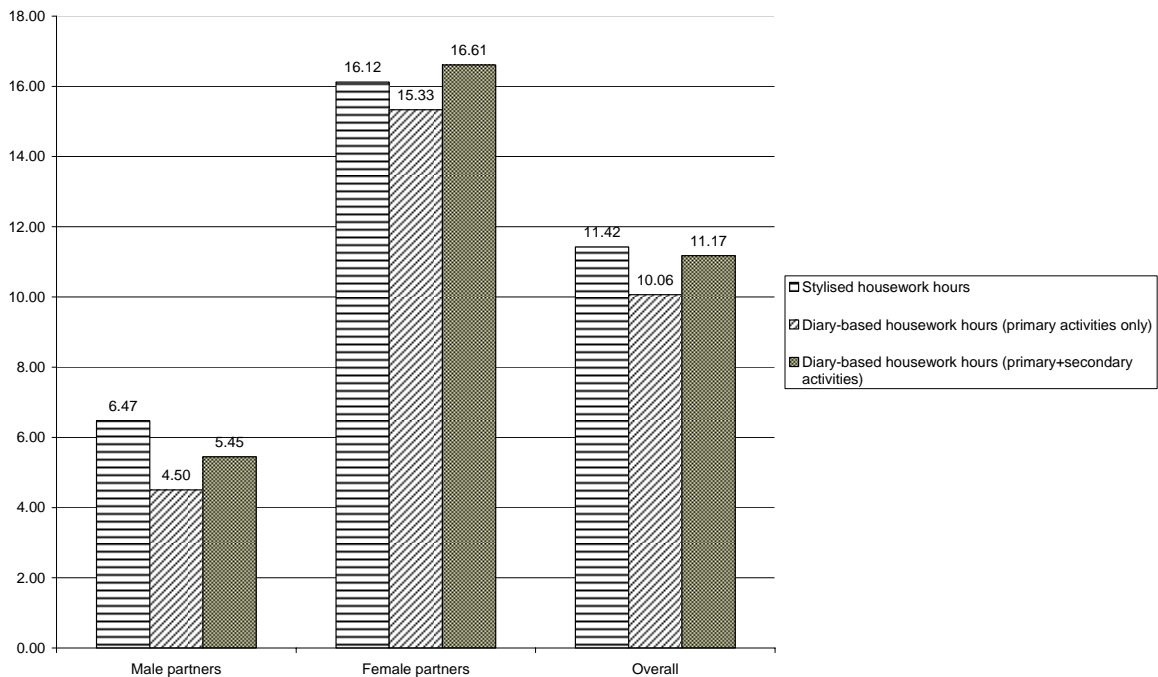
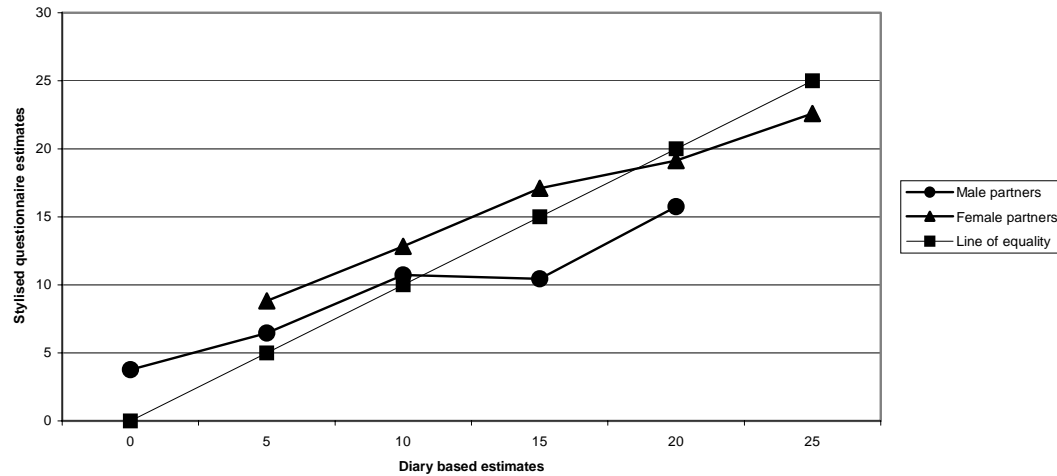


Figure 2 shows whether the gap between the two kinds of estimates varies with the length of housework hours. It plots the mean housework hours reported in the survey against the mid-point of 5-hour intervals of the diary-based housework hours. For example, the “10 hours per week point” represents the housework time of all those recording between 7.5 and 12.5 hours of housework per week in the diaries (where only primary activities are included). A 45-degree “line of equality” is included for reference. (The point against 0 in the x -axis is not shown in the case of women’s hours and that against 25 is not shown for men’s hours, because the numbers of cases are small). As we can see, for women, the stylised estimates are generally higher than the diary-based estimates when the mid-points of diary-based hours are below 20; from 20 onwards, they are slightly lower, indicating that women tend to report less housework hours in surveys than in diaries when their housework hours are long. In the case of men, the curves of estimates cut at around 10 of the mid-points of diary-based hours, which is significantly greater the mean value. In other words, most men report longer housework hours in surveys than in diaries; but their stylised hours are substantially lower when the mid-point of the diary-based hours is over 10.

Figure 2. Stylised Vs Diary Estimates of Weekly Housework Hours, Married/Cohabiting Men and Women



To explore the effect of irregularity in housework hours, men and women in Figures 3.1 and 3.2 respectively are divided into two groups, with half of them being defined as having high variation in housework hours and the rest have low variation (high variation, if coefficient of variation > 0.48 ; low variation, otherwise). We see that in the case of men little difference is observed between the two curves until the 5-hour mid-point of diary-based hours; but from this point to 10, the high variation group is above the low variation one, showing that they tend to report longer housework hours in surveys than in diaries. In the case of women, no substantial difference between the curves is seen until the mid-point 10 of diary-based hours; from this point onwards, the high variation group remains under the low variation one until the point 20, where we see the trend starts to reverse.

Figure 3.1. Stylised Vs Diary Estimates of Weekly Housework Hours, Married/Cohabiting Men, by Variation in Housework Time

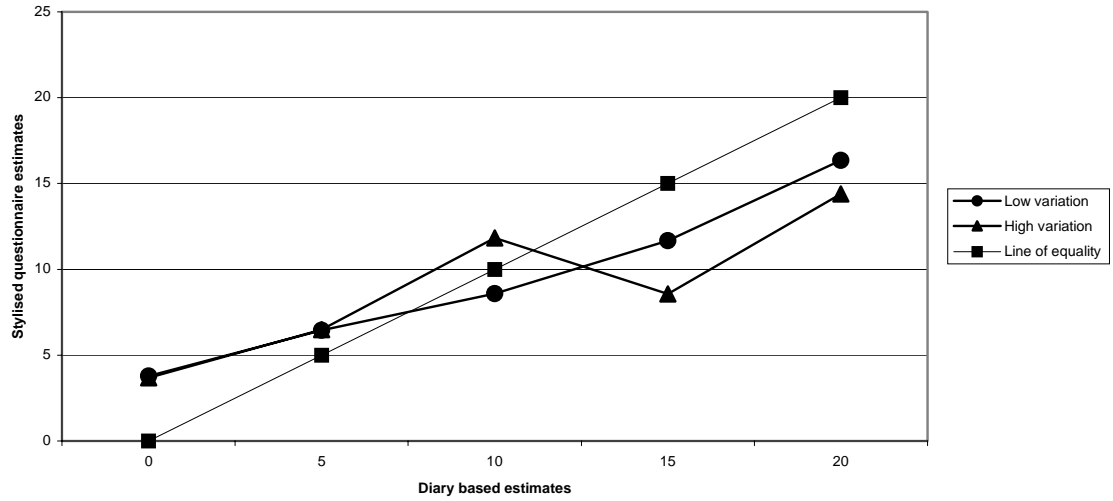


Figure 3.2. Stylised Vs Diary Estimates of Weekly Housework Hours, Married/Cohabiting Women, by Variation in Housework Time

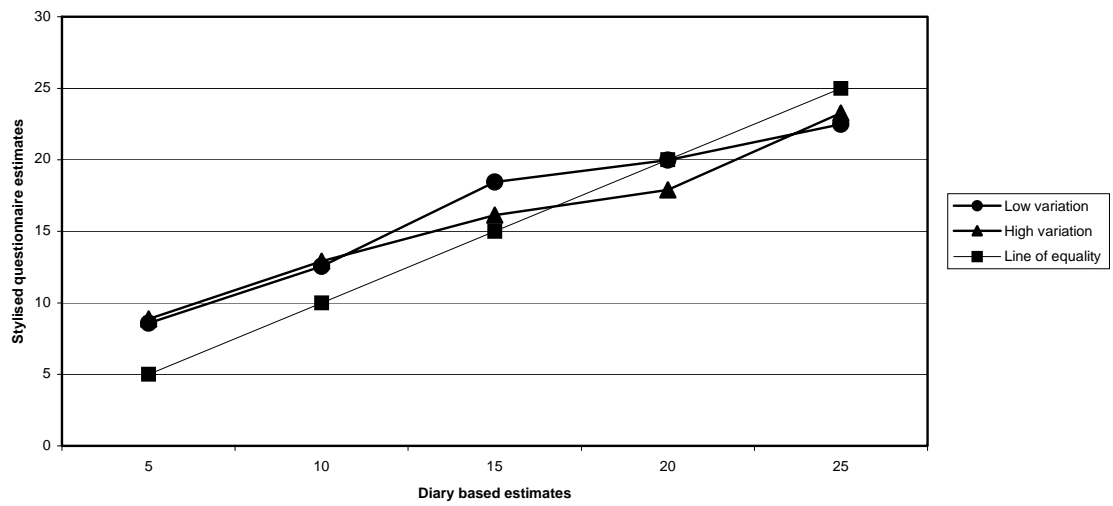


Figure 4.1. Stylised Vs Diary Estimates of Weekly Housework Hours, Married/Cohabiting Men, by the Presence of Dependent Child

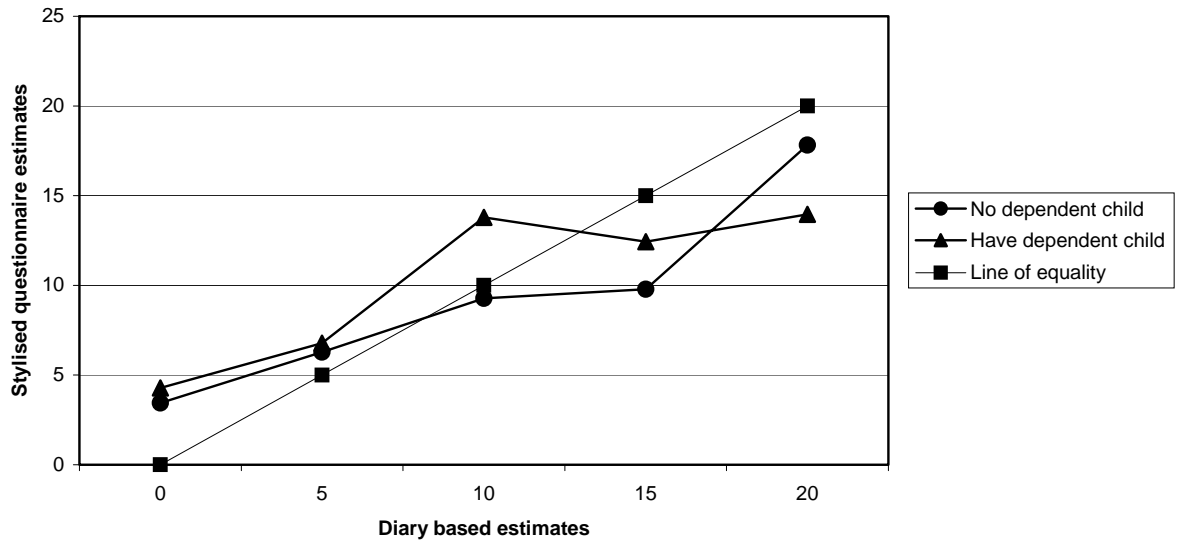
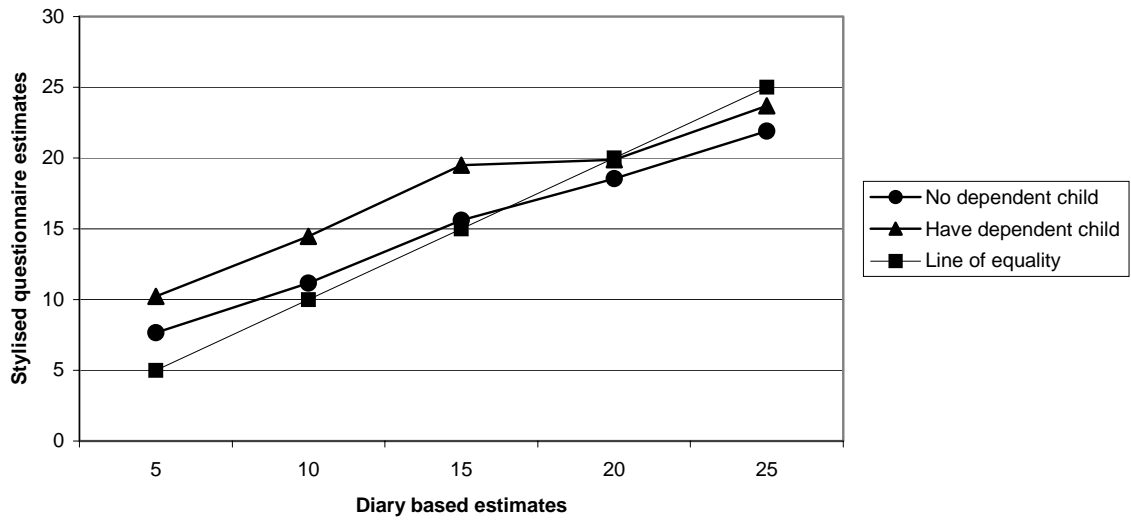


Figure 4.2. Stylised Vs Diary Estimates of Weekly Housework Hours, Married/Cohabiting Women, by the Presence of Dependent Child



Figures 4.1 and 4.2 depict the effect of dependent children on the reported housework hours. As we can see, for both men and women, the presence of one or more dependent child (aged under 16) in the household generally moves the curve upwards (i.e., it tends to make respondents report longer housework hours in the survey). For men, the gap between the two curves is rather narrow until after the 5 hour mid-point of diary-based hours, from which point the gap increases up to about 4 hours. For women, the gap remains steadily at about 3 hours until the 20 mid-point of diary-based hours, from which it drops to about 1 hour.

Lastly, Figures 5.1 and 5.2 compare respondents who hold traditional gender role attitudes with those who do not. In male partners' case, the gap between the stylised hours and the diary-based ones is higher for those who are traditional than those who do not. While all men tend to report longer housework hours in surveys than in diaries when their diary-based hours are below 10, the extent of "over-reporting" is higher for traditional men. When the diary-based hours are over 10, men's stylised hours are lower than their diary-based hours, but again the figure suggests that traditional men tend to "under-report" their housework time to a greater extent than other men. The curves in the figure of women's housework hours do not form a clear pattern. The two curves intertwined together and not much difference is observed between them. The gap between the curves grows only after the 20 mid-point diary-based hours, with traditional women "under-reporting" about 1 hour more than other women.

Figure 5.1. Stylised Vs Diary Estimates of Weekly Housework Hours, Married/Cohabiting Men, by Gender Role Attitudes

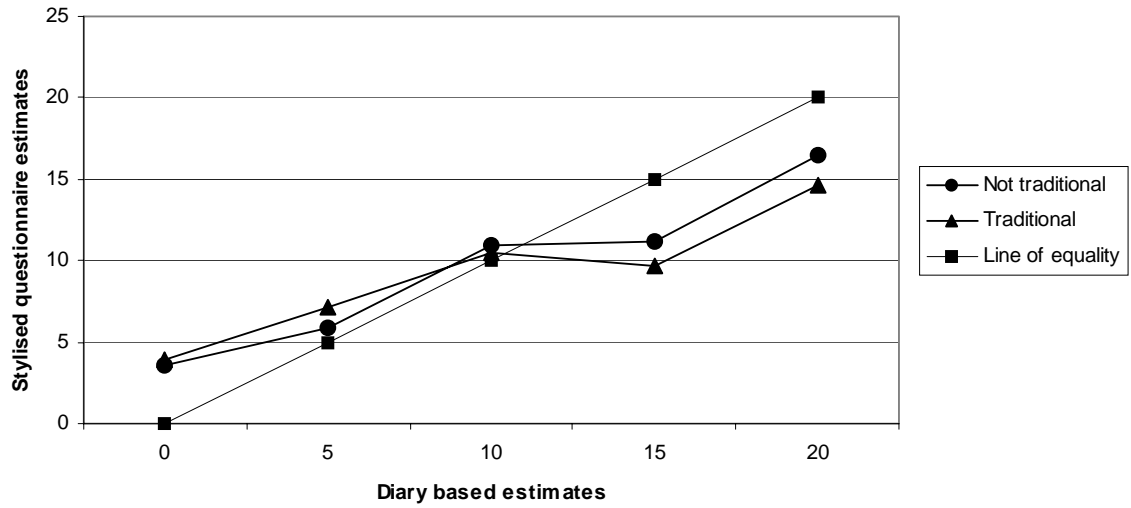
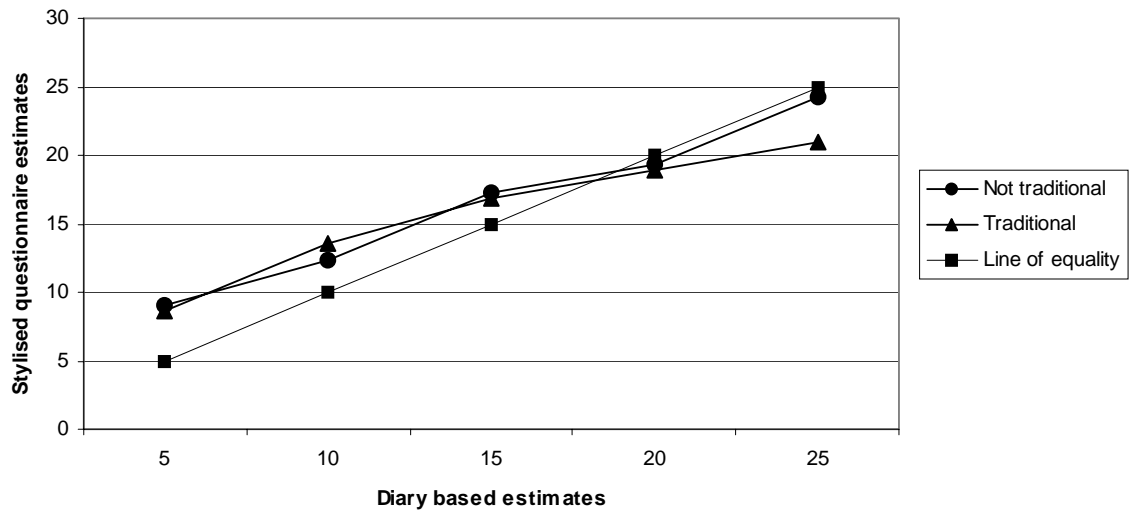


Figure 5.2. Stylised Vs Diary Estimates of Weekly Housework Hours, Married/Cohabiting Women, by Gender Role Attitudes



To sum up results of the figures, there is supportive evidence for **H1**:

Although women on average have longer housework hours than men, the gap between their stylised and diary-based housework hours is not larger than that of men; rather, they report their housework hours more accurately than men. Moreover, the gap between the two kinds of estimates becomes narrower when hours of housework as a secondary activity are taken into account. This result concurs well with the hypothesis that in answering the survey question some respondents count the time during which they undertake housework simultaneously with other activities (**H3**). The figures have also shown that the gap between stylised and diary-based housework hours may vary according to the length of housework hours, irregularity in of the distribution of housework time over the week, the presence of dependent children and gender role attitudes (**H2**, **H4** and **H5**). However, the mechanisms that produce the gap might be different for men and women.

Table 1. *F Tests of OLS Models of Stylised Weekly Housework Hours (Baseline Plus Interactions)*

Effect	<i>R</i> ² of reduced model	<i>df</i>	<i>R</i> ² change	<i>F</i> for <i>R</i> ² change
<u>Male partners</u>				
Primary housework hours*Variation	0.222	1	0.003	2.327
Primary housework hours*Dependent children	0.224	1	0.000	0.038
Primary housework hours*Traditional attitudes	0.218	1	0.007	5.512*
<u>Female partners</u>				
Primary housework hours*Variation	0.235	1	0.007	5.946*
Primary housework hours*Dependent children	0.242	1	0.000	0.006
Primary housework hours*Traditional attitudes	0.241	1	0.001	1.156

Note: Data from Home On-line Study, 1999 - 2001. *N* = 632 for male partners' models; *N* = 666 for female partners' models. The *R*² change indicates the difference in *R*² between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. $F = [R^2 \text{ change}/1]/[(1-R^2)/N-k-1]$, where *N* = sample size, and *k* = the number of independent variables.

The final model contains the interactions and the following independent variables: Housework hours (primary activities), housework hours (secondary activities), variation in primary housework hours, the presence of young children (*Yes* = 1; *No* = 0), and traditional attitudes (*Yes* = 1; *No* = 0). *R*² of the male partners' final model = 0.224, *df* = 8, *R*² of the female partners' final model = 0.242, *df* = 8.

Multivariate Analyses

Men's and women's reported housework hours in the survey are regressed separately with their diary-based housework hours (primary activities), hours of housework as secondary activities, coefficient of irregularity of housework time, the presence of children, and traditional gender role attitudes⁴. Table 1 shows the significance test results of the interactions of primary housework hours with irregularity of day housework time, the presence of children and traditional gender role attitudes respectively. Here the backward elimination method is used. The removal of the interaction between primary housework hours and traditional gender attitudes brings a significant reduction in fit of the male partners' model (i.e., this interaction term would improve the fit of the model significantly should it be included). But the other two interactions are both insignificant. As to the female partners' model, only the interaction between primary housework hours and the coefficient of irregularity of housework hours would make a significant decrease in fit of the model should it be removed.

⁴ Descriptive statistics and correlations of variables for the multivariate regressions are given in Appendix II.

Table 2. *The Preferred OLS Models of Stylised Weekly Housework Hours*

	Male partners			Female partners		
	<i>B</i>	<i>SE B</i>	β	<i>B</i>	<i>SE B</i>	β
Primary housework hours from diary	0.80***	0.07	0.52	0.40***	0.07	0.33
Secondary housework hours from diary	-0.01	0.11	0.00	0.40**	0.13	0.11
Coefficient of variation in housework hours	0.63	0.57	0.04	-5.45**	1.93	-0.17
Presence of dependent children (<i>Yes = 1; No = 0</i>)	1.21*	0.55	0.08	2.14**	0.73	0.10
Traditional attitudes (<i>Yes = 1; No = 0</i>)	1.69*	0.74	0.11	-0.48	0.73	-0.02
Primary housework*Coefficient of variation				0.36*	0.15	0.13
Primary housework*Traditional attitudes	-0.26*	0.11	-0.13			
Constant	1.88**	0.61		9.22***	1.50	
<i>R</i> ²		0.221			0.241	
<i>F</i>		29.62***			34.79***	
<i>Df</i>		6			6	

Note: Data from Home On-line Study, 1999 - 2001. *N* = 632 for the male partners' model; *N* = 666 for the female partners' model.
 p* < .05. *p* < .01. ****p* < .001.

The final preferred models for male partners and female partners with inclusion of the significant interaction terms are presented in Table 2. As has been expected, primary housework hours are a significant predictor in both models. But it should be noted that this variable is associated with men's stylised housework hours to a greater extent (the coefficient is substantially larger), because the coefficient of secondary housework hours is significant in the female partners' model but not in the male partners' one. Moreover, the coefficient for secondary housework hours is of a similar size to that of the primary housework hours in the female partners' model. These results suggest that secondary housework hours influence only women, but not men, significantly on their reported housework hours in surveys. In other words, there is supportive evidence for **H3**: Housework undertaken as a secondary activity affects women's reporting of housework hours in surveys significantly, as much as their primary housework hours do. But their association with men's stylised housework hours is insignificant and small. This can be due to the fact that men undertake less housework, whether primary or secondary, than women and their stylised hours are hence less influenced by activities performed simultaneously with housework.

Turning to the coefficient of variation in housework hours, in the case of men it is not significantly associated with stylised housework hours. On the other hand, it is significantly and substantially associated with women's stylised hours. The negative value of the coefficient indicates that stylised estimates are lower than diary-based estimates, i.e., there is a tendency for women to report less in surveys than they have done (according to their diary records). But the irregularity in housework time also interacts with primary housework hours significantly and the corresponding coefficient is of an opposite sign. That is, the level of under-reporting decreases with

increases in primary housework hours. When housework hours are sufficiently long (i.e., greater than 15.14; $0.36 \times 15.14 - 5.45 = 0$, where holding the coefficient of variation and other variables constant), the under-reporting effect by irregularity in hours of housework will be annulled and instead women tend to report longer hours than the diary-based estimates. These results are consistent with the earlier findings in Figure 3.2. Hence the findings support **H2**: Irregularity in housework hours is associated with the gap between the two types of estimates. Again, we see a difference in the effect by gender. It appears that men's stylised housework hours are not associated significantly with the level of irregularity in housework hours. Again, this is perhaps because men undertake less housework than women and are less likely to have to juggle between paid work and domestic work. We can see from the table in Appendix II that the level of irregularity in housework hours is negatively correlated with housework hours in the case of women, and the correlation is positive in the case men. That is, irregularity of housework hours in the case of men are less likely due to the competition between conflicting demands on time for paid work and domestic work, which may create time pressure and affect accuracy in recalling housework hours.

Moreover, concurring with the earlier observations in Figures 4.1 and 4.2, the presence of dependent children is positively and significantly associated with the stylised housework hours in both models. That is, both men and women tend to report longer housework hours in surveys than in diaries with the presence with one or more dependent children, even when primary and secondary housework hours are controlled for. The coefficient in the female partners' model is greater than that in the male partners' one. This can be due in part to the collinearity between women's primary housework hours and the presence of dependent children (See Appendix II).

The overall findings are in accord with **H4**: The presence of dependent children will inflate the gap between the estimates.

Finally, traditional gender role attitudes are associated significantly only with men's stylised housework hours but not with women's. The positive coefficient (1.69) indicates that traditional men tend to report more housework hours in surveys than other men. But the interaction between traditional gender role attitudes and primary housework hours is of the opposite sign, suggesting that traditional men who have long housework hours tend to report less housework in surveys. For example, when a traditional man undertakes more than 6.5 hours of housework, the over-reporting effect of traditional attitudes will be annulled and further increments in diary-hours will predict decreases in stylised housework hours. This finding constitutes a powerful supportive evidence for the social desirability mechanism described by **H5**: when men holding traditional gender role attitudes undertake long hours of housework, i.e., when the amount of housework done conflicts with their gendered expectations, they tend to report less housework hours in the survey. Interesting, significant results are only found for men but not for women. This is likely because housework is regarded as a "women's sphere" and men are more prone to feel a clash with their male identity when they perform long hours of housework.

This finding also forms an interesting parallel with earlier research findings on the relationship between economic dependency and housework participation, which show that economically dependent men tend to "do gender" by avoiding participation in housework (e.g., Brines, 1994; Greenstein, 2000). These past studies are however based on stylised housework hours. Because the reporting process in surveys is gendered and it affects particularly men with traditional gender role attitudes, some qualifications should be made on the conclusion about the "doing gender" effect on

men's housework participation. It might well be the case that traditional men "do gender" in the reporting process by claiming shorter housework hours in survey interviews.

CONCLUSIONS

Stylised estimates of housework hours are widely used in research on the domestic division of labour as if they are unproblematic. This study has shown that there are systematic errors, which may affect the use of these estimates for the analysis of men's and women's participation in housework.

It is found that women generally report their housework hours more accurately than men; the difference between the two types of estimates constitutes more than 30% of men's stylised housework time, whereas the proportion is less than 5% in the case of women. Therefore, the gender gap in housework participation will be underestimated when we compare directly the means of stylized estimates of men's and women's housework hours.

These results depart from those of Bonke (2005) and Kitterød and Lyngstad (2005), which suggest no tendency of over-reporting in surveys, and no or little gender difference in the accuracy of reporting housework in surveys held in Denmark and Norway respectively. This is perhaps due to the differences in societal contexts between these countries and the U.K., where gender relations and the gender division of labour are comparatively less equal. It is expected that gender differences in housework reporting are related to differences in social desirability of women's and men's participation in housework. This argument can be further verified by comparing the gap between diary-based and stylised estimates of housework using data from other countries in future research.

The findings also indicate that the mechanisms that produce systematic errors in stylised estimates of housework hours are different for men and women. Men holding traditional gender-role attitudes tend to report more housework time in surveys than in diaries, but the tendency is reversed when they undertake long hours of housework. This finding casts doubts on earlier research that suggested economically dependent men tend to undertake less housework in order to avoid further loss in their masculine identity. It is because many traditional men might have claimed shorter housework hours than they actually did in survey interviews, which might in turn have obscured the linear relationship between economic dependency and housework hours.

Other main factors accounting for the gap include the presence of dependent children, which will inflate both men's and women's stylised housework hours; housework undertaken as secondary activities and irregularity in housework hours, which affect particularly women.

Are stylised estimates of housework hours still acceptably reliable? The survey approach, after all, remains a relatively flexible and a low cost way of collecting time use data. Stylised estimates of housework time are available in a number of large-scale national surveys. In many cases, as past studies showed, they provide a fairly useful ordinal scaling of individuals' time on housework (Baxter & Bittman, 1995; Niemi, 1993; Robinson, 1985). Rather than abandoning all these rich data, researchers may make better use of them by being cautious in the interpretation of results derived from them. They should be aware that the systematic errors may reduce or exaggerate the gap in housework participation between different population groups. It should also be noted that systematic errors may affect dependencies between variables and hence the results of multivariate analyses. The trustworthiness

of the results will be undermined when variables known to generate systematic errors (e.g., gender role attitudes, irregularity in housework hours and presence of dependent children) are of theoretical interest in the interpretation.

Finally, to improve the accuracy of stylised estimates of housework, it will be worthwhile to adopt a combination of the diary and the survey approaches to collect time use data. Despite being able to produce more accurate time use estimates, the diary method is expensive and often impractical in large-scale household panel studies. A sub-sample of respondents in a large-scale survey can be requested to record time use information in diaries. The diary-based estimates of time use can be regressed with the stylised estimates collected in the survey, and the coefficients of regressors can then be used to enhance the stylised estimates of the entire sample (See Gershuny, 2003; Kan & Gershuny, forthcoming).

REFERENCES

- Baxter, J., & Bittman, M. (1995). Measuring time spent on housework: A comparison of two approaches. *Australian Journal of Social Research*, *1*, 21-46.
- Bianchi, S. M., Milkie, M. A., Sayer, L. C., & Robinson, J. P. (2000). Is anyone doing housework? Trends in the gender division of household labor. *Social Forces*, *79*, 191-222.
- Bittman, M., England, P., Folbre, N., Sayer, L., & Matheson, G. (2003). When Does Gender Trump Money? Bargaining and Time in Household Work. *American Journal of Sociology*, *109*, 186-214.
- Bonke, J. (2005). Paid work and unpaid work: Diary information versus questionnaire information. *Social Indicators Research*, *70*, 349-368.
- Brines, J. (1994). Economic Dependency, Gender and the Division of Domestic Labour at Home. *American Journal of Sociology*, *100*, 652-688.
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disorders*, *175*, 526-536.
- Fenstermaker Berk, S., & Shih, A. (1980). Contributions to household labor: Comparing wives' and husbands' reports. In S. Fenstermaker Berk (Ed.) *Women and household labor*. Beverly Hills, CA: Sage.
- Gershuny, J. (2000). A concise atlas of time use: Twenty countries, thirty-three years' change. In J. Gershuny (Ed.) *Changing times: Work and leisure in postindustrial society* (pp.160-219). New York: Oxford University Press.
- Gershuny, J. (2003). Time, through the lifecourse, in the family. *Working Paper of Institute for Social and Economic Research, Paper 2003-3*. Colchester, UK: The University of Essex.

- Gershuny, J. (2005). *Stylised estimates, activity logs and diaries: estimating paid and unpaid work time*. Paper presented at the XXVII International Association for Time-Use Research Conference (2 November-4 November). Halifax, Canada.
- Gershuny, J., & Robinson, J. P. (1994). Measuring hours of paid work: time-diary vs. estimate questions. *Bulletin of Labour Statistics*. Geneva: International Labour Office.
- Greenstein, T. N. (2000). Economic Dependence, Gender, and the Division of Labor in the Home: A Replication and Extension. *Journal of Marriage and the Family*, 62, 322-335.
- Juster, F. T., Ono, H., & Stafford, F. P. (2003). An assessment of alternative measures of time use. *Sociological Methodology*, 33, 19-54.
- Juster, F. T., & Stafford, F. P. (1985). *Time, goods, and well-being*. Ann Arbor, MI: The University of Michigan.
- Kan, M. Y., & Gershuny, J. (forthcoming). Calibrating time use estimates for the BHPS. *Working Paper of the Institute for Social and Economic Research*. Colchester, UK: University of Essex.
- Kan, M. Y., & Pudney, S. (forthcoming). Measurement errors in stylised and diary-based data on time use. *Working Paper of the Institute for Social and Economic Research*. Colchester, UK: University of Essex.
- Kitterød, R. H., & Lyngstad, T. H. (2005). Diary versus questionnaire information on time spent on housework - The case of Norway. *Electronic International Journal of Time Use Research*, 2, 13-32.
- Laurie, H., & Gershuny, J. (2000). Couples, work and money. In R. Berthoud & J. Gershuny (Eds.) *Seven Years in the Lives of British Families: Evidence on the*

dynamics of social change from the British Household Panel Survey. Bristol: Policy Press.

- Lee, Y.-S., & Waite, L. J. (2005). Husbands' and wives' time spent on housework: A comparison of measures. *Journal of Marriage and Family*, *67*, 328-336.
- Marini, M. M., & Shelton, B. A. (1993). Measuring household work: Recent experience in the United States. *Social Science Research*, *22*, 361-382.
- Niemi, I. (1993). Systematic error in behavioural measurement: Comparing results from interview and time budget studies. *Social Indicators Research*, *30*, 229-244.
- Plewis, I., Creeser, R., & Mooney, A. (1990). Reliability and validity of time budget data: Children's activities outside school. *Journal of Official Statistics*, *6*, 411-419.
- Press, J. E., & Townsley, E. (1998). Wives' and husbands' housework reporting: Gender, class, and social desirability. *Gender and Society*, *12*, 188-218.
- Robinson, J. P. (1985). The validity and reliability of diaries versus alternative time use measures. In F. T. Juster & F. P. Stafford (Eds.) *Time, goods, and well-being* (pp.33-62). Ann Arbor, MI: The University of Michigan.
- Robinson, J. P. (1997). The overestimated workweek and trends in hours at work. In J. P. Robinson & G. Godbey (Eds.) *Time for life: The surprising ways Americans use their time* (pp.81-96). Pennsylvania: The Pennsylvania States University Press.
- Warner, R. L. (1986). Alternative strategies for measuring household division of labor: A comparison. *Journal of Family Issues*, *7*, 179-195.
- Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, *57*, 348-368.

APPENDIX I

The following four items are used to measure gender role attitudes:

- a) Both the husband and wife should contribute to the household income;
- b) Having a full-time job is the best way for a woman to be an independent person;
- c) A husband's job is to earn money; a wife's job is to look after the home and family.
- d) A pre-school child is likely to suffer if his or her mother works.

In the Home On-line Study, respondents were asked to indicate if they *strongly agree, agree, neither agree nor disagree, disagree* or *strongly disagree* with each of the statements. Responses to the above statements are recoded and then added up to create a score ranging from 0 to 16, where higher values indicate more traditional attitudes and 8 is neutral. There are about 20% of cases with missing values. For these cases, the values are imputed using education, age and gender.

APPENDIX II

*Correlations and Descriptive Statistics for All Variables*1. Male partners

Variables	1	2	3	4	5	6
1. Stylised housework hours	–					
2. Primary housework hours from diary	0.46***	–				
3. Secondary housework hours from diary	0.00	0.01	–			
4. Coefficient of variation	0.14***	0.22***	-0.05	–		
5. Having a dependent child (<i>Yes = 1; No = 0</i>)	0.05	-0.05	-0.05	0.06	–	
6. Traditional gender attitudes (<i>Yes = 1; No = 0</i>)	0.01	-0.02	-0.03	-0.02	-0.17***	–
M	6.47	4.50	0.94	0.50	0.38	0.45
SD	7.43	4.82	2.43	0.48	–	–

2. Female partners

Variables	1	2	3	4	5	6
1. Stylised housework hours	–					
2. Primary housework hours from diary	0.45***	–				
3. Secondary housework hours from diary	0.12**	0.01	–			
4. Coefficient of variation	-0.29***	-0.52***	-0.06	–		
5. Having a dependent child (<i>Yes = 1; No = 0</i>)	0.09*	-0.05	0.04	-0.02	–	
6. Traditional gender attitudes (<i>Yes = 1; No = 0</i>)	-0.02	0.03	0.04	-0.01	-0.15***	–
M	16.12	15.33	1.28	0.55	0.44	0.45
SD	10.57	8.74	2.82	0.33	–	–

Note: Data from the Home On Line Study, 1999 – 2001. $N = 632$ for the male partners' model; $N = 666$ for the female partners' model.

* $p < .05$. ** $p < .01$. *** $p < .001$.