1998-05-15

Microeconomic analysis of time-use data. Did we reach the promised land?

N. Anders Klevmarken*

ABSTRACT

This paper surveys the last decade of micro-economic research using time-use data. Focusing on the household production model, time-use as an investment activity, and the distribution of extended income, issues of data collection, measurement errors, model specification and estimation as well as substantive results are reviewed and discussed. Although time-use data have specific characteristics which need be considered in analysis there is more to learn from these data and in particular if present short-comings are dealt with in future data collection.

Keywords: time-use, household production, labor supply, auxiliary information
JEL Classification: J22 D13 C81

* Department of Economics, Uppsala University, P.O. Box 513, S-751 20 UPPSALA, Sweden.
Fax +46 18 471 1478, email anders.klevmarken@nek.uu.se

Introduction

The measurement of time-use had an early tradition in the former planned economies. The survey of time-use analysis by Juster and Stafford (1991) mentions an early study from 1924 in the former USSR. There were also a few contemporary studies in the United States and in Japan. The international comparative time-use surveys inspired by the work in the planned economies and organized by Alexander Szalai (Szalai 1966, 1972) are usually mentioned as the starting point of modern time-use measurement in Western market economies. Time-use studies were primarily justified as giving data for an analysis and valuation of household work, but also for their ability to give information about leisure (the ultimate utility yielding activities?), commuting and travel behavior, etc.

The Juster and Stafford(1991) survey distinguishes between micro-economic time-use analysis and work with national accounting. Although the research community seem to have developed a division of work, the two are obviously related, as evidenced, for instance in Quah(1993) and Wagman & Folbre(1996). Theories and empirical results about household behavior in the nonmarket sector will influence the design of national accounting, and when satelite accounts for the household sector are developed they can be used to analyse the productive contribution from the household sector to the economy, see for instance, the discussion about household services and economic growth in the United states in Wagman & Folbre(1996).

Micro-economic analysis of household time-use has usually used the paradigm of the household as a producing and consuming unit which seeks to find an optimal combination of market work, household work and leisure as to maximize its well-being. Although the work by Becker on the economics of the family now forms the basis for most current work in this area there are both theoretical and empirical contributions at least as early as those of Margaret Reid(1934, 1947).

Much of micro-economic analysis has focused on the effects of taxes, benefits and social policy on household behavior and on household differences in well-being. The overwhelming majority of empirical studies in this area did not use time-use data but more conventional measures of market work, and measures of income and consumption expenditures. What expectations did we have about time-use surveys moving research forward in this area? Almost twenty years ago I was visiting the Survey Research Center of ISR, University of Michigan, and preparing for a new household panel survey in my own country. Inspired by the 1975-76 Michigan time-use survey we decided to implement a full scale time-use survey as part of the first wave for the new panel, later called the Swedish HUS panel study. (See Klevmarken & Olovsson 1993.) Almost ten years later, in 1993 still another time-use study was added to the panel (Flood et.al. 1997). We expected that time-use measurements would

- give better measures of market work,
- improve the analysis of labor supply by explicitly including competing activities in the home,
- make feasible studies of gender differences in market and nonmarket work, and thus also improve our understanding of female labor supply,
- improve our understanding of the demand for consumer goods by permitting the estimation of joint demand and time-use models,
- give better measures of economic well-being and further our knowledge of its distribution.

In this paper we will investigate to what extent these expectations or goals have been fulfilled and discuss the circumstances which have led research to its present state of art. The survey article by Juster and Stafford (1991) in an excellent way took stock of economic time-use research until the beginning of the 1990s. This paper will thus primarily focus on the last ten years of research. I shall also primarily discuss research about household behavior and have relatively less to say about time-use accounting.

1. Measures of market time.

Conventional measures of market time based on survey questions about normal weekly hours tend to give empirical frequency distributions which have pronounced peaks at full-time hours for men and at half-time and full-time hours for women. The observed high concentration to peak hours is probably exaggerated. There are good reasons to believe that many respondents report their contracted number of hours disregarding or forgetting any nonwork episodes at work and any irregular overtime work. Even if asked explicitly about secondary work it might also be difficult to report hours retrospectively, in particular if the respondent only works intermittently in this job. In general, those who have irregular work hours will find it difficult to respond to questions about normal hours. Time-use diaries are, however, normally collected such that meals, coffee breaks and other work breaks, over-time and time on secondary jobs are carefully recorded. In particular if a time diary is given in a "yesterday interview" and not in a leave behind diary, its sequential nature makes it difficult to falsify. [1] Time-use surveys also have the advantage of giving data on travel to and from work. Sometimes it is desirable to add commuting time to pure market work time.

Figure 1 illustrates the differences between data based on questions about current hours per week including overtime and secondary jobs ("survey data") and time-use data from the same samples of people. Data were obtained from the HUS surveys (Klevmarken & Olovsson 1993, Flood et.al 1997). The time-use data distributions are much smoother and have a larger variance. The explanation is partly that just given but also the noisiness of time-use data due to the fact that only a few days are observed for each respondent. (One might also note that independently of data source the distribution for women have become more alike that of men.)

Carlin & Flood(1997) compared estimates of male labor supply from time-use data with those from conventional survey data using a so called double-hurdle model. Referring to previous studies they noted that the presence of young children normally decreases work hours for women while the effect for males has typically become nonsignificant or weakly positive. Sweden's active policy to bring women into the labor market and involve fathers more actively in the care of children, and independent evidence that this policy to some extent has been successful, suggested that one might expect a negative effect on labor supply also for Swedish males. They found no significant effect of the presence of young children when the estimates were based on the responses to the survey question about normal weekly hours, but they found a negative effect when they used time-use data. The double-hurdle model suggested that the largest share of this effect came from fathers loosing entire days rather than reducing hours of work when working. The explanation to this difference in results is thus that time-use data recorded temporary, unusual and unexpected episodes of absence from work (in this case

[1] For this reason one might also argue that work time from a time-use diary will not only include time in the regular "white" market but also in the "black" market.

probably because kids became sick), while this was not the case with data based on responses to the question about normal hours of work. In their study the number of children in different age groups were treated as exogenously given. Using a Hausman test they could not reject this assumption, but there is still a question as to what would happen if children were endogenous to the labor supply decisions. For the other explanatory variables the parameter estimates did not depend much on whether time-use or survey data were used. The invariance of wage rate and wealth elasticities may not be a general result. Graversen & Smith (1997) using Danish data demonstrated that the average wage rate and income elasticities were reduced by more than a half when overtime and time in secondary jobs were deleted from their workhours measure. Since individual adjustments to incentive changes are most easily done by a change in overtime work and in secondary jobs this result would seem more plausible.

Depending on the importance of breaks, nonwork at work, overtime and secondary jobs and irregular jobs there are reasons to expect systematic differences in estimates of average work hours from time-use diaries compared to conventional labor force surveys. Juster and Stafford(1991) report that "conventional respondent reports of labor supply seriously overstate the amount of hours actually supplied to the market" (p. 486). They also claimed that conventional Current Population Survey estimates underestimated the 1965-1981 decline in work hours (compared to time-use estimates). This conclusion was, however, questioned by Leete & Schor(1994) who suggested that the Michigan Study, only measuring weekly hours, did not adequately reflect the substantial rise in weeks worked per year found in the CPS.[2] Leete & Schor(1994) found support for the "time-squeeze" hypothesis. According to their results Americans worked longer hours and enjoyed less leisure at the end of the period.

Table 1 compares a number of different measures of weekly hours worked for Sweden. There are three groups of estimates: survey-type estimates from the HUS surveys, estimates from the official labor force surveys and time-use estimates. They all differ in level as well as to rate of change. The official estimates give an average of about 35 hours per week for all men in work including those who are temporarily absent, and about 26 hours for women. These estimates show a small decrease in hours for men and a small increase for women. Conceptually closest comparison is the HUS data estimates for normal hours including overtime and secondary jobs. For men they show a small increase from 41.8 to 42.5 hours, and for females there is an increase from 35 to 37.5 hours. The difference in change between the two types of estimates is probably on the boarder line of being significant for males but it is significant for females. The time-use estimates show a completely different picture. The rows "All" include everybody, not only those in the labor force, and thus give much lower mean estimates. To obtain something which is closer to the "survey estimates" the estimates on the last and third last rows in the table were restricted to those who had responded that they worked on a job or were temporarily absent at the time of the time-use interviews. In 1984 these estimates were lower than the corresponding Labor Force Survey (LFS) estimates for males whether break time was included or not, while for women the estimate excluding break time was about the same as the corresponding LFS estimate. In 1993 all time-use estimates exceeded the LFS estimates. Time-use data thus show a strong increase in hours worked, while LFS data only show a modest increase for women and virtually no change for men.

[2] They also argued that Juster & Stafford(1991) had "not corrected for the fact that in the 1965 sample all household heads were employed. This is especially important because 1981 was a recession year" (p. 41).

One might have thought that the increase in unemployment from 2-3 per cent to about 8 per cent would have decreased the average hours of work for the entire group of individuals 20-64 years old, but our estimates tell a different story. Both gender increase their average hours and more detailed data also show that this holds for every age group, and that the share of respondents who reported positive hours have increased. The increase in unemployment is not large enough to balance other forces increasing work hours. [3] Independent sources suggest that sickness absence has decreased drastically and that overtime work has increased among those who have a job. These are two plausible explanations to the observed increase in work hours. Decreased sickness and unemployment benefits and an increased unemployment rate have made people more reluctant to take time off work, and the reluctance of employers to hire new people have increased hours of work and overtime hours among those who have a job.

Similarly to Juster & Stafford (1991) we thus get a rather different picture of changes in work hours from time-use data compared to more conventional measures. In their case time-use data showed a stronger decrease and in our case a stronger increase. More work is needed to fully understand why all these estimates differ, but a preliminary conclusion is that time-use estimates more clearly measure hours actually worked and that they are more sensitive to changes in the market than measures based on traditional survey questions. If future research verifies that our interpretation of the time-use results is correct this will have a major impact of our understanding of the economy. In spite of the increased unemployment in Sweden there was a major increase in total hours worked and thus productivity as measured by the ratio of output to work hours, increased much less than we previously thought!

A major disadvantage of time-use data is their short duration, normally just a few days for each respondent. Daily variations in time-use will have a large influence on estimates of weekly hours of work, and much of the greater smoothness of the distribution of time-use estimates of weekly hours compared to alternative estimates derives from variations from day to day in behavior. Assume for instance, the following model,

$$h_{id} = \mu_d + \varepsilon_{id}; \qquad (1)$$

where h_{id} is observed hours of market work for respondent i in day d, μ_d is the expected hours of work and ε_{id} a random error such that $h_{id} \geq 0$. Suppose the sampling design is such that the days of a week are stratified into workdays and weekends and one day is sampled randomly from each stratum. A natural estimate of weekly hours then becomes,

$$h_i = 5h_{iw} + 2h_{is} = \mu_i + \varepsilon_i; \qquad (2)$$

where μ_i and ε_i are implicitly defined by eq. (1). Suppose now that μ_i follows a distribution with three peaks, one at zero hours, one at half-time hours and one at full-time hours. It is then easy to show that h_i will have a distribution, which is much smoother. How clearly the peaks of the underlying distribution of μ_i will show depends on the properties of ε_i. μ_i is thus the average number of hours of market work one would obtain if one could measure hours of work for each day of a whole year (week) without error. ε_i will for instance, depend on

[3] Although it cannot explain the increase in hours it is interesting to note that 5-10 per cent of those who are unemployed report positive work hours in the time-use survey.

whether the respondent has a regular or irregular work schedule, if he is absent because of illness or for another reason, if he works overtime or on a secondary job and if he does more or less nonwork at work. Furthermore, it will depend on any seasonallity in workhours, unemployment, sickness, holidays and on decisions to leave and enter the labor force. It is easy to imagine that the distribution of ε_i depends on the day of the week, season of year and of the type of respondent.

Suppose now that one could get independent and more reliable estimates on weekly (annual) workhours. They could then be used for at least two purposes. First, to estimate the properties of ε_i , which would become useful in designing time-use surveys,[4] and second to improve on the nonmarket time-use estimates. We will return to the latter issue in section 5 below.

3. The household production approach

Following Becker(1965), Gronau(1977) and others the household production model is the framework used in most empirical studies of household behavior which includes household work. This is now a well-known model and there are several excellent reviews, for instance Gronau(1986). Here suffices only a brief summing-up. A more recent survey of household models which also includes bargaining models and focus on intrahousehold distribution is Behrman(1997).

Assume the following model for a single person household,

$$U(X_M, X_H, t_L);\ U'>0,\ U''<0 \qquad (3a)$$
$$X_H = H(X_{M'}, t_H);\ \ H'>0,\ H''<0 \qquad (3b)$$
$$X_M+X_{M'} = Wt_M + Y; \qquad (3c)$$
$$t_M + t_H + t_L = T; \qquad (3d)$$

where U is a utility function which is maximized with respect to market goods X_M, household goods X_H, and leisure time t_L . This maximization is done subject to the household production function H, which uses market goods for household production $X_{M'}$, and time t_H as inputs, and subject to the budget and time constraints (3c) and (3d) respectively. The lagrangian to this problem becomes,

$$\mathcal{L} = U(X_M, H(X_{M'}, t_H), t_L) + \lambda_1(X_M+X_{M'} - W(T- t_H - t_L) - Y) + \lambda_2(T- t_H - t_L); \qquad (4)$$

It is assumed that the first constraint is always binding, while $T- t_H - t_L \geq 0$. If $T- t_H - t_L >0$, that is the household is engaged in market work, then $\lambda_2 = 0$, while if $T- t_H - t_L = 0$ then $\lambda_2 > 0$. From the first-order conditions to this maximization problem one can derive the following marginal conditions,

$$\frac{\partial U}{\partial t_L} = \frac{\partial U}{\partial H}\frac{\partial U}{\partial t_H}; \qquad (5a)$$

[4] See Karlton (1985)

$$\frac{\partial U}{\partial X_M} = \frac{\partial U}{\partial H}\frac{\partial H}{\partial X_{M'}}; \qquad (5b)$$

$$\frac{\frac{\partial U}{\partial t_L}}{\frac{\partial U}{\partial X_M}} = \frac{\frac{\partial H}{\partial t_H}}{\frac{\partial H}{\partial X_{M'}}} = W + \frac{\lambda_2}{\frac{\partial U}{\partial X_M}}; \qquad (5c)$$

From the first two relations we find that t_H, t_L and X_M, $X_{M'}$ are chosen such that the marginal utility of market work equals the marginal utility of household work, and that the marginal utilities of goods for direct consumption and for home production are equal. The last condition says that the marginal rate of substitution between leisure and market goods for direct consumption equals the rate of substitution between time and goods for household production and, if the household chooses to work in the market, the wage rate. If the optimum does not involve any market work the value of leisure (and time in home production) is higher than the market wage rate. An increased wage rate will thus either increase the purchases of market goods, or decrease leisure and household production time (and thus increase hours of market work), or both depending on the properties of the utility function and the household production function.

The solution to the maximization problem is in principle given by the following system of four equations, supplemented by the identity for time in market work.

$t_s = t_s(W, Y \mid U, H)$;   s=L,H. (6a)

$X_m = X_m(W, Y \mid U, H)$; m=M, M'. (6b)

$t_M = T - t_L - t_H$; (6c)

These functions do not only depend on preferences but also on the household production technology.

As noted long ago data on time-use and purchases of consumption goods are in general not sufficient to identify both preference parameters and the household production technology. To do that one will need measures on the output from household production. Some analysts have then ignored the distinction between preferences and household production technology and estimated mongrel time-use equations. The study by Kooreman & Kapteyn(1987) is such an example. They assumed an indirect translog household ”utility” function, which in principle was allowed to represent both preferences and technology. It’s arguments were the male and female time-use in seven leisure activities (but not time in market work) and total market consumption. This function was maximized subject to a time constraint for each spouse and a budget constraint for the whole household. They used the Michigan time-use data, and found relatively strong wage rate and income elasticities for some of the disaggregate activities, but much smaller for the total of all leisure and household activities. Aggregation thus wiped out these effects. They also showed that income had a negative impact on household work which is at variance with the models of Gronau(1977, 1980) and Graham & Green(1984) which by assumption had no such effects. Another interesting finding was that female and male household work activities tended to be substitutes while leisure activities were complements.

Although Koorman & Kapteyn(1987) did recognize that the time-use functions they estimated were a mongrel of preference and technology structures their model did not have any variables which specifically were thought to capture differences in household technology. A study in the same general approach as that of Koorman & Kapteyn (1987) which did include such variables is, for instance, the study of time-use in the Indian city of Madras by Malathy(1994). The presence of labor-saving appliances contributed to the explanation of housework and meal preparation with a strong negative effect. Malathy also allowed for a gender effect of other household members and found that the presence of other women in the household reduced household work and meal preparation for the wife, while other males did not contribute to these activities. This paper also included experiments including hired help as an explanatory variable. There was a strong negative effect, but its interpretation is not straight forward because it was not treated as an endogenous decision variable.

Others have attempted to identify some parameters of the household production technology by introducing assumptions which reduce the complexity of the problem, for instance, perfect substitution between market goods and home produced goods and constant returns to scale in home production. With few exceptions it has not been possible to test these assumptions, so there is not much empirical evidence as to their realism, but *a priori* one might be skeptical. Unfortunately much of the empirical results based on the household production model are driven by strong assumptions about behavior and particular functional forms.

The main advantage of the household production model in empirical work is that it suggests a reformulation the derived demand equations (6) such that properties of the household production process become included. We should not only use variables, which capture heterogeneity in preferences but also variables, which represent heterogeneity in the production process.

The household production model also suggests that even if output measures are unavailable time-use data could improve on estimates of labor supply and demand for consumption goods. If the household production model is true, omission of household production from the model will in general bias the estimated wage and income effects on labor supply. There are at least two reasons for this. In a conventional labor supply model all nonmarket time is treated as leisure. Following the model above the sum of true leisure and household production time becomes,

$$t_H + t_L = t_H(W, Y \mid U, H) + t_L(W, Y \mid U, H); \qquad (7)$$

If t_H and t_L are relatively complex nonlinear functions, a composite - market work leisure model could easily lead to a misspecification of functional form. Furthermore, household productivity variables would become omitted leading to an omitted variables bias. This is demonstrated empirically in Apps & Rees(1996). Their model differs in several respects from the model outlined above. First, it is a model for a two-person household and it assumed that each spouse maximized his/her utility subject to the constraint that the partner's utility did not fall below a certain threshold. In principle their model allowed for lump sum transfers between spouses, but in their empirical application this was not permitted. Their household production function was simpler than in the model above because household production depended only on time input but not on any input of goods. In the absence of any consumption data and measures of output they also had to assume that the household production function was linear homogeneous. The corresponding cost function was specified as a translog function. Using these assumptions the implicit price of output from household production equals the unit cost function, and a function for the allocation of time to household production can be derived and estimated. Using these estimates the authors could compute estimates of the output price, which were plugged into an almost ideal demand system. The demand system explained the shares of total income allocated to purchases of market goods, household produced goods and services and leisure for each of the two spouses.

Using Australian data the authors compared the estimates obtained for this model with those from a model which omitted household production. They found that the uncompensated wage rate elasticities were smaller in the household production model compared to the more conventional model and that the income elasticities were close to zero too. They then concluded (p. 211) ”A policy implication of the estimates is that using systems which ignore household production to analyze the effects of reforms may lead to results that overstate incentive gains from lowering tax rates on income.”

The assumption of constant returns in household production made it possible to estimate the household cost function and the corresponding marginal cost even without data on household output. The identification problem is further discussed in Apps & Rees(1997), where they show that the model is not identified with a more general production function and time-use data only.

With reference to previous literature on the impossibility of identifying a household cost function in a conventional demand model from demand data alone, Kapteyn(1994) discussed alternative approaches to the identification problem and in particular the use of direct measures of people’s feelings of well-being. He demonstrated that if the subjective measures of well-being obtained in a sequence of survey questions originally designed by Van Praag(1968), actually measure the preferences revealed by conventional demand data, then they can contribute to the identification of preference parameters. The same approach should also work in the household production model, and continuing in the same vein it might be possible to ask people about their preferences for market produced goods compared to household produced goods. For instance, one could ask households to disregard the costs of doing things and in a scale from 1-10 rank how they value a house cleaned by someone hired from outside compared to ”do-it-yourself”. Such data might help separate preferences from production technology in the absence of direct output measures. In a model permitting other activities than leisure to yield direct utility the ”process benefit” scores collected in the Michigan time-use studies and the Swedish HUS studies should also help identify preferences.

Repeatedly it has been suggested that output data would be the best help in identifying preferences separately from production technology, and that all other approaches are second best solutions. There is, to my knowledge, only one unique study, which collected output data and estimated household production functions directly, namely Fitzgerald, Swenson and Wicks (1996). This study was based on a small sample of 135 married couples from Missoula, Montana. Interview data were collected on the number of tasks done in 49 different household production activities and on the time-use devoted to each task. These tasks were then valued at market price and the value of any intermediate goods used subtracted. The measures used were thus of the value-added type. To be able to aggregate some kind of monetary valuation is probably needed but it is not obvious that tasks should be valued at market price. This valuation principle disregards any positive or negative difference in quality between home production and market production. It might be possible to improve on such measures by asking the respondent if the quality of a home produced good is more or less than a market alternative.

Fitzgerald, Swenson & Wicks(1996) estimated a translog production function for each of six aggregate household commodities, and used the estimates for a number of interesting test. For instance, they found that a single production function could not adequately describe the production process of all six commodities. Separate functions were needed. The Cobb-Douglas specification used in previous studies was rejected. They also tested whether productivity in household production depended on age and education and found that ”only in the equation for home repairs were husband’s and wife’s age significant as multipliers of their labor hours”, while ”education never made a difference”. If this is true the significance of these variables in previous studies based on mongrel preference-technology functions should imply that these variables primarily capture heterogeneity in preferences.[5] Additional results were that the hypothesis of constant returns to scale could only be rejected for ”cleaning”, that husbands’ and wives’ time-uses were substitutes in most cases and both tended to be substitutes with the input of capital.

Although these results are very interesting and encouraging the small sample of this study might have reduced the power of the tests such that no rejection is the result of small power rather than the absence of an effect. There is thus a need for more research along these lines.

4. Time-use and investment activities

Investmentsinone-self

In an interesting article Biddle & Hamermesh (1990) analyze sleep and the allocation of time with the motivation that sleep can be seen as an investment which enhances labor market productivity. As they note “if decisions about sleep are not separable from decisions about labor supply – the vast literature on labor supply that ignores sleep contains a difficulty that could have important consequences for understanding the allocation of time between home and the market” (p. 923). In such a model the marginal price of sleep will no longer equal the marginal price of other leisure and the impact of sleep on market productivity and thus on the wage rate makes the wage rate endogenous to labor supply. Their empirical results were, however, not overwhelmingly strong. The wage rate and income effects on sleep and waking nonmarket time were rather uncertain, but for men they concluded that market work is unaffected by an increase in the wage rate, while such an increase will make men shift from sleep to leisure and home production. The interpretation of this result they suggested is that leisure and home production are complements to consumption goods.

Biddle’s and Hamermesh’s idea could be brought even further. Most activities have an investment aspect in addition to a consumption aspect. Not only sleep but nutritious meals, sport activities, certain leisure activities and visits to doctors and dentists can all be seen as investments in good health which, in addition to a general increase in well-being, gives a higher

[5] However, a home production activity like child care, which is likely to depend both on age and education of the parents, was not included among the activities studied in Fitzgerald, Swenson & Wicks(1996).

productivity in market and home production activities. Previous time-use decisions will thus influence current decisions through their effect on people's health capital. An empirical analysis of this mechanism would require panel data of which there is a very short supply[6].

When analyzing the impact of time-use on health not only the total duration of activities, for instance excessively long hours of market work, are important but also the sequence of activities might be important. Very long spells of one and the same activity might have a detrimental effect on health as also many short interrupted spells of an activity, which could be taken as an indication of stress.

Gaining experience in the labor market is in the Becker-Mincer tradition seen as an investment in human capital. Analogously experience of home production can be seen as an investment, which will enhance productivity in the home. To test the relative importance of this type of human capital we would need to collect information about previous spells of household work.

Investmentsinchildren [7]

The care and schooling of children has been the topic of research in several disciplines using many different approaches. In economics child care and schooling are viewed as investment activities. Parents use their own time and purchased goods and services to give their children human capital in the form of knowledge, experiences and good health. In some cultures parents get a return on their investments when their children later work in the market and contribute to family income, particularly when the parents get old. To have many children is thereby old age pension insurance in these cultures. In our modern Western societies this motive to invest in children has no longer the same importance. From a collective point of view it is, however, very important that the parental generations invest in their children and that these investments generate economic growth.

Quite independently of our self interest (individually or collectively) when retiring, most people would probably acknowledge that it is important to have children and to give them a good start in life. If a natural instinct, an altruistic concern for one's children and future generations, or a selfish satisfaction of consuming kids, most people want children and enjoy having them. As shown by Juster (1985) and Flood and Klevmarken (1990) playing with kids and taking care of children is the activity which gives the highest "process benefit", to borrow Juster's terminology. There are thus compelling reasons to invest in children.

These investments take different forms. Home investments consists of the quantity and quality of time inputs and the quantity and quality of goods inputs which jointly with inherited ability, will determine the level of schooling which finally influence post-school investments and income. We may thus distinguish between home investments and investments outside home, and time investments and goods investments.

TimeInvestmentsatHome

We might wish to distinguish between different kinds of time inputs. One is the direct care of a child, an activity in which the child (children) is the primary target, for instance feeding a child,

[6] The Canadian Halifax 1971/81 panel and the Michigan time-use survey 1975/81 are small panels and the Swedish HUS 1984/93 time-use surveys also have a panel subsample.
[7] The remaining part of section 4 relies heavily on Klevmarken & Stafford(1997)

dressing a child, reading to a child or helping out with homework. People commonly do more than one activity at the same time. For instance, a mother might be cooking and helping with homework at the same time. Depending on which activity the mother considers to be her 'primary activity' (in time-use studies) that is classified as the primary activity and the other as a secondary activity.

Both primary and secondary activities could involve investments in children. There are also activities in which the child is not the primary target but the child is present when the activity is done. A family could, for instance, have a meal jointly or they could be doing some kind of sports or outdoor activity together or a child could simply be watching when an adult is doing something. These activities may also include investment aspects, i.e. the child is learning while watching and doing. Finally, there are activities in which the child is not present but which are carried out to the benefit of the child, for instance, the child's dirty laundry is washed and ironed, or a meal is cooked for the child.

It is not obvious how one should go assigning investment measures to these diverse forms of activity. Time-use studies give the time input of parents (and others) into direct childcare and sometimes also data on other episodes during which children are present. Time-use for the benefit of children without them being present can usually only be estimated indirectly. Most current time-use studies only give estimates on the time use of the adult, but not on how much time was spent on each child. Simply averaging by the number of children in the family will not do because in some activities a child might benefit as much from the activity doing it jointly with other children as by doing it alone. Indeed, there are activities like certain games, which require more than one participant. Also, parents and school-based caregivers may provide unequal amounts of time to siblings. One apparent motivation is a type of intra-school or intra-family equity (Brown and Saks, 1975, Gustafsson and Stafford, 1997). The literature on child investments also discusses the importance of the order of birth among siblings for the level of parents' time investments as well as the transmission of human capital between siblings. There are, for example results indicating that the first and last child receive more attention from parents than do middle children (Hanushek, 1993).

Ideally one would need child time diaries. Even for diaries of individual children in the same family there are questions of 'scale economies' or public goods. In this case adult time for one child is not at the expense of the other children and, as in the example of organizing a game, the value is from the team element (as defined in economics (Radner, 1986)). One view is that formal school settings consist heavily of those activities where there are scale economies or public goods for adult time in the learning process.

Measures of time input are not necessarily good measures on investments for at least two reasons. First, as noted above, an activity might be done both for its investment contents and for its consumption benefits and it is difficult to separate the two. Second, the quality of the input of hours and minutes might differ. Is one hour of TV-viewing equivalent to one hour of homework? To what extent does content of the TV show or homework matter? It might be possible to rank activities as to their investment contents and more or less arbitrarily weight the time-use proportionally to this. If measures on the return on child investments were available these could be related to the time input in various activities and one could in this way assess their relative importance. In the absence of return measures any such operation involves many ad hoc judgments.

An alternative approach to measuring investments is the ‘cost of time approach’. That is, time inputs are converted into monetary inputs by using either a wage rate of the person doing the activity or, if there is a market alternative, what it would cost to purchase the service or good produced. This approach is easiest to defend when there are true opportunity costs or market alternatives, when this is not the case any monetary evaluation becomes rather arbitrary. If, for instance, a woman gives up her job to care for her children, then there is a recognizable opportunity cost, but if a housewife gives up some of her leisure to care for her children the opportunity cost is less well defined.

A few studies have pointed out that in addition to income forgone while a parent is at home taking care of children there is also a forgone opportunity to invest in additional own human capital and build up future earnings. These estimates (Smith and Ward, 1989; Joshi, 1990, 1994; Calhoun and Espenshade, 1988; Dankmeyer, 1996; Stafford and Sundström, 1997) indicate a substantial cost of this sort. In the Netherlands, this ‘time out’ appears to be very costly for women with less schooling, since the market work experience is even more important for their careers (Dankmeyer, 1996).

TimeInputsOutsidetheHome

From an investment perspective the most important time input outside home is certainly time in preschool and school activities. As with in-home capital formation, there are quality differences between time on task at school, certain tasks at school probably have more investment content than others. Outside school there are also activities which contribute to the human capital of a child. There are more or less organized post-school activities which aim at teaching the child certain skills, but also socializing among other children gives the child useful experiences about group behavior and how to establish relations with others. The playground is an important arena for investments in human capital! For most countries we currently lack measures of time inputs outside the home, except for a few classroom surveys, none of which are based on national samples. Time-use surveys only give data on the time adults interact with children. We need child time-use diaries also for this purpose, particularly of the day at school or other out-of-home activities.

Estimates of the monetary value of these out-of- home investments would have to build on estimates of the value of adult time input, especially for young children, and the value of any goods input. In this way it might be possible to estimate, for instance, the investment value of preschool and school activities. Similarly, for a sports activity one might like to add the time cost of an adult coach and the rental value of the sports ground and any equipment. A problem in this case is though, that one also might like to subtract something for the consumption benefit of sporting. For child activities which neither involve any adults nor any equipment there is no monetary estimate obtainable from the cost side. For children at older ages, their own time and its opportunity cost becomes more important. In higher school grades there are higher pupil/teacher ratios, so the active engagement of the child and forgone market earnings become important (Hansen, 1963).

Thetotalsizeofchildinvestments

Klevmarken & Stafford(1997) give estimates for parents time with children for a number of countries and in particular for Sweden and the United States. They then continue putting data together to arrive at a few summary measures of investments in Swedish children.

Table 2 displays the result. Time inputs are of three kinds: direct child care in which the child is at center of the activity, incremental household work and time with kids. The latter activity includes all activities in which children are present while the purpose of the activity need not be to take care of children. In terms of hours time with kids is by far the largest activity, but we do not know what quality differences hide behind these numbers. We also note that small children take more time than other kids in direct child care and incremental household work. In order to compare input of time with input of goods and services time in direct child care was (arbitrarily) valued at an average wage rate of 80 SEK. The value of household work was put to half of that amount and time with kids to 20 per cent. The result is an input of time corresponding to 87200 SEK per child and year for children below 7 and 65600 for older kids. The incremental input of consumption goods was estimated to about 30 000 SEK per child and year. This is a very rough figure obtained from household budget surveys. Better estimates would most likely show that the incremental input depends both on the age of the child and the total number of children in the family. Finally we added estimates of publicly provided goods and services obtained from others sources. These cover public child care, schools and certain cash benefits to families with children.

These estimates indicate that Swedes on average invest annually between 150000 and 200000 SEK on a young child, which approximately corresponds to an investment in USD between 22000 and 29000. Grossed up to a national total for investments in children these numbers amount to more than twenty per cent of GDP, which is at least as much as gross investment in fixed capital! Given our weighting of time inputs, total time input makes up 49 per cent of the total, private goods and services 17 and publicly provided goods and services 34 per cent for young children. For older children the corresponding shares are 39, 18 and 43.

Whoinvestsandwhichchildrenbenefit?

Analysis of the Michigan time-use data showed that more educated mothers spent more time with their children and did more incremental housework than less educated mothers Corresponding Swedish estimates showed much smaller differences due to education and most of the observed differences were insignificant. Gustafsson & Kjulin(1994) who also used the 1984 HUS time-use data, found a positive effect of education on child care for both males and females when estimating a Tobit model, but a negative effect on other household work for women and no effect for men. The size of the estimated marginal effect was 1 hour per week for women who had passed grade school compared to whose who had not done so, and about half as much for males. Similar results were obtained by Malathy(1994) for India. Why did not the higher market wage potential and presumed higher value to career enhancing activities of well educated mothers make them substitute child care and household production for market work? Do more educated women have higher preferences for child quality as reflected in a higher income elasticity?

Malathy(1994) got a significantly positive effect on time allocated to teaching children of the value of household assets, which can be seen as a proxy for nonlabor incomes, and a strong negative effect of the wife’s wage rate. The study by Gustafsson & Kjulin (1994) did not include any income or assets variable which might explain why they got a significantly positive effect of the female wage rate.

The generous supply of public childcare in Sweden might at least partly explain why incremental child care and household work is less in Sweden than in the United States and why differences in time-use due to education and gender are smaller.[8] Gustafsson & Kjulin(1994) found that in families who had their children cared for outside home women decreased their time in care of infants by about one hour per week, and their time in household work by almost as much. Type of childcare had no effect on the mother's time in childcare for preschoolers but care outside home reduced her time in household work by as much as 4.5 hours per week. For males there were almost no effect of the type of child care on their time with kids, but if their children were cared for outside home they were able to reduce their time in other household work by 2 to 3 hours per week.

To conclude, analysis of the determinants of investments in children and their effects on economic growth will probably become at least as important as analysis of investments in buildings and machinery. The distribution of these investments among children will also greatly influence the distribution of income and standard of living. As evidenced by the studies reviewed above private investments by parents are likely to be as important as public investments in the form of public childcare, schooling and health. Time-use studies tell us about the size of time-investments of parents and about the distribution among households, but to get really good input measures we also need data on the time-use of kids, as are now collected within the PSID surveys. Even if input measures are useful we also need return measures; otherwise we will never be able to evaluate the relative importance of alternative investment strategies. It is not obvious what measures to use, but school performance might be indicative of the return to early childhood investments, while (life-cycle) labor income is a measure on the return to all human capital investments.

5. The distribution of extended income.

Studies of the distribution of incomes have been criticized for ignoring the value of household production. The importance of also valuing household production is perhaps best seen for a farm household, which has a choice between consuming some of its produce or selling it in the market. But also in an ordinary wage earner household there is much household production, which contributes to the well being of the household members.

A few attempts have been made to estimate the value of household work based on time-use studies and extend household disposable income. In the words of Jenkins & O'Leary(1996, p.402), extended income "provides a better measure of a person's access to economic resources, and one which is less contaminated by the effects of differences in preferences". Although it is true that extended income might be less influenced by time allocation decisions than money income, it is still susceptible to decisions influenced by preferences. If market work and household work are differently valued, relative preferences for these two kinds of activities and for market goods compared to home produced goods will determine extended income. Furthermore, someone who prefers more leisure before more output from household production will *ceteris paribus* have a lower extended income but not necessarily be less well off. This kind of argument leads towards Becker's (1965) concept of "full income". However, only if all activities have the same intrinsic value the distribution of full income becomes

[8] Another explanation is that the U.S. time-use survey was done in the mid 1970's while the Swedish surveys were carried out almost 10 and respectively 20 years later.

independent of preference based time allocations. The distribution of full income then reduces to the distribution of the value of time among individuals.

There is an extensive literature on the valuation of household work, primarily for national accounting purposes, which also is of relevance in this context. Without going into details it suffices to note that the two main approaches are valuation by market alternative or "housekeeping wage", and the opportunity cost principle.

Bonke (1992) used the Danish time-use survey from 1987 and applied what he called a modified opportunity cost method. Annual income data were obtained from registers and merged with time-use data in an exact match. Because the individual and not the household was the unit of observation in the time-use survey Bonke did not have time-use observations for both spouses but had to impute for the missing spouse. His main finding was that the Gini coefficient remained virtually unchanged when the value of household work was added to money income of couples. The Lorenz curves crossed, however, which implies that the change in inequality will depended on the particular inequality measure used. Bonke also referred to results from a study of urban U.S. households by Bryant & Zick (1985) which only got a marginal decrease in the Gini coefficient when the value of household work was added.

Jenkins & O'Leary (1996), however, got a clear-cut and large reduction in inequality when they extended the income concept for the U.K. The Lorenz curves did not cross and the Gini coefficient dropped from 0.292 to 0.209 when the opportunity cost principle was used and to 0.170 with the "housekeeping wage" method. The reduction was thus 28 per cent and 42 per cent respectively. They were also able to demonstrate a substantial reranking of households caused by the extension of the income concept.

A problem with the study of Jenkins and O'Leary is that their time-use survey did not include income data. Time-use in household work was thus imputed to the 1986 British Family Expenditure Survey using statistical matching techniques. Although Jenkins and O'Leary thought they got a fairly close match, they could not exclude that the difference between their results and previous results could be attributed to the matching.

Merz & Kirsten (1998) got results for Germany similar to those of Jenkins & O'Leary (1996). The Gini coefficient was estimated to 0.246 for market income and to 0.174 for extended income, a decrease of 29 percent.

On *a priori* grounds it would seem reasonable to believe that household work is relatively more important among households with no or only one member working in the market and among low income households. Extending income should then reduce inequality, unless there is a sufficiently strong positive correlation between money income and the valuation of household work to reverse the effect. If the Jenkins and O'Leary and the Merz and Kirsten results now look plausible is it possible that previous studies have underestimated the equalizing effect of household work? If the estimates of time in household work and its total value were based on observed time-use for just one or a few days without compensation for the large day to day variance, then this variance could have inflated the inequality measures. To see this let's introduce the following notation,

$$h_{it} = h_{i.} + e_{it}; \qquad (8)$$

where h_{it} is time-use in household work for respondent i and day t, $h_{i.}$ is the mean time-use per day and year for i, and e_{it} the deviation from this mean in day t. Assume for simplicity that the standard deviation of e_{it} is σ_e the same for everyone, and that the standard deviation of $h_{i.}$ is σ. Assume also that only one day is observed for every respondent. The estimates of annual household work then become,

$$H_i = 365\, h_{it} = 365 h_{i.} + 365 e_{it} \qquad (9)$$

with coefficient of variation,

$$CV(H_i) = \frac{(\sigma^2+\sigma_e^2)^{1/2}}{\bar{h}_i}; \qquad (10)$$

The inequality holds approximately if the sample of respondents is sufficiently large to make the mean of the individual deviations e_{it} cancel. If money income is given on an annual basis this is, however not the inequality measure we want. It is rather

$$CV(h_i) = \frac{\sigma}{\bar{h}_i}; \qquad (11)$$

Extended income is defined,

$$E_i = Y_i + w_i h_{i.}; \qquad (12)$$

If we simplify by assuming that the valuation of household time is the same for everyone, w_i=w, then the squared coefficient of variation for extended income becomes,

$$CV^2(E) = \frac{\bar{Y}}{\bar{E}} CV^2(Y) + \frac{w\bar{h}_{i.}}{\bar{E}} CV^2(h_{i.}) + 2\rho_{Yh_{i.}} \frac{w\bar{Y}\bar{h}_{i.}}{\bar{E}^2} CV(Y)CV(h_{i.}); \qquad (13)$$

Now, if H_i is used to estimate hours of household work the coefficient of variation (10) will replace the true coefficient (11) in expression (13) and thus contribute to an overestimate of the inequality of extended income. (The correlation between money income and hours in household work will become underestimated, but this is exactly offset by the higher $CV(h_{i.})$ in the last term of expression (13) above.) The same general conclusion also holds if the estimate of hours of household work is based on more than one day of observation, but the additional information will reduce the degree of overestimation.

If this is the explanation to the differences in result between the two more recent studies Jenkins & O'Leary (1996) and Merz & Kirsten (1998) and previous studies is of course impossible to know without going into unpublished details of the procedures used in all studies. But it is interesting to note that Jenkins & O'Leary (1996) used weekly income data and their time-use data referred to an average day obtained from 7 days of diary information, while Bonke (1992) had annual income data and only one day of time-use data per respondent.[9]

[9] In a Norwegian study, Aslaksen & Koren (1995), the Gini coefficient dropped by 22 per cent when the annual disposable income per household was extended by the value of household work. Household work was valued at

We may also note that with two or more observations for every respondent it becomes possible to identify and estimate σ_e , and thus also to correct for the positive bias in the coefficient of variation for the time-use estimates.

6. Special features of time-use data and their consequences for analysis.

Time-use data have certain properties, which make analysis difficult. They include,

- noisy data because we get very short spells of data,
- frequent “zero observations”, respondents have not done an activity on the designated day,
- events of short duration and rare events are missing or underreported
- systematic differences in time-use due to day of the week and season of the year,
- depending on survey design, dependencies on past behavior of longer duration than a day are not observed,
- the quality of the data are much dependent on the interviewers (interviewer effects),
- quality also depends on consistent coding of activities,
- selective nonresponse is likely because people who are rarely at home, or who are very busy either at work or at home more frequently leave no or an incomplete diary,
- members of the same household are not always observed in the same day,
- needed supplementary data are often missing, for instance, wage rates, incomes, expenditures and child care arrangements.

Some of these problems are best handled by improved data collection procedures, but the application of proper econometric models and methods are useful too. This is not the place to discuss all these problems thoroughly, but a few comments and suggestions are offered.

y

The most important problem is probably the noisiness of time-use data. A larger sample of days would reduce this problem but at the expense of higher nonresponse and increased field costs. Alternative approaches, which might reduce the influence of day to day variations in time-use, could thus become helpful.

The constraint that time-use in all activities must add up to 24 hours per day in 365 days per year suggests that independent estimates of hours worked contributes information which could be used to reduce the random variation in the time-use estimates for nonmarket activities. If a respondent to a time-use survey worked exceptionally long hours in the market on the designated day hours of household work and leisure will become unusually short. If we knew that this was an atypical day we should be able to use this information and adjust hours of household work and leisure accordingly.

To fix ideas let t_{ij} be individual i’s time-use in activity j on a randomly selected day.

---

the wage rate of an housewife substitute (husmorsvikarieloen) and the time-use in household work was estimated from two days of measurements per respondent. The details of the computations of extended income were, however, not given in the paper. This study was also missing the spouse‘s time-use which had to be imputed.

$$\sum_j t_{ij} = T; \tag{14}$$

If time-use is measured in hours then T=24. Assume also that

$$E(t_{ij}|i) = T_{ij} \tag{15}$$

T_{ij} is thus individual i's average time-use per day in activity j for an entire year (week), and the deviation $t_{ij} - T_{ij}$ is only due to the random sampling of days. Assume now that there is supplementary information about one of the activities, say T_{i1}, $\forall$i. Using the analog of a ratio estimator from sampling theory one feasible estimator is,

$$\hat{t}_{ij} = \frac{T - T_{i1}}{T - t_{i1}} t_{ij}; \;\; j = 2,...,k \tag{16a}$$

$$\hat{t}_{i1} = T_{i1}; \tag{16b}$$

One immediately finds that this estimator satisfies the adding up constraint,

$$\sum_j \hat{t}_{ij} = T; \; \forall i \tag{17}$$

If n days are sampled by simple random sampling (which is usually not the case) the mean estimator becomes a conventional ratio estimator.

$$\frac{1}{n}\sum \hat{t}_{ij} = (T - T_{i1})\frac{1}{n}\sum \frac{t_{ij}}{T - t_{i1}}; \; \forall i \tag{18}$$

where the summation is done over the sampled days. This estimator has the usual properties of a ratio estimator. It is biased (although the bias is likely to be small) but if t_{i1} is sufficiently highly negatively correlated with t_{ij} it has a smaller variance than the usual uncorrected estimator (sample mean). The adding up constraint suggests that the correlation is negative, at least for broadly defined activities. However, if activities 1 and j complement each other (positive correlation) the estimator (16) will have a larger variance than an unadjusted estimator and should not be used.

The auxiliary information available might not be an exact measure of T_{i1} but just a proxy (cf. Table 1). In this case one could use a slightly modified estimator based on the following auxiliary regression,

$$t_{i1} = b_0 + b_1\hat{T}_{i1} + e_{i1}; \; i = 1,..., N. \tag{19}$$

where $\hat{T}_{i1}$ is the error prone auxiliary estimate of T_{i1}. Thus, the predictions from this regression replace T_{i1} in the estimator (16) above. This approach rests on the assumption that t_{i1} is an unbiased estimator of T_{i1}.

There are a few disadvantages of this estimator. First, every activity J=2,…,k is adjusted by the same factor, and second every activity which the respondent did not perform on the designated day is left unadjusted at zero hours. Both these properties will introduce a bias the size of which will depend on the grouping of the activities and on people’s behavior.

If someone had unusually long hours of market work on the designated day it is likely that she would not do all of the household cores or leisure activities she normally does. An adjustment towards a normal situation should thus include an increase in hours at least for some of the observed zero hours activities.

There are many ways to incorporate auxiliary information about an activity, some of which will handle these problems. One approach is to generalize the ratio estimator to a regression estimator, but it would have the disadvantage of not necessarily producing non negative estimates. Another approach, which includes the regression estimator as a special case, is the calibration estimator, see DeVille & Särndal (1992). With a suitably chosen “distance function” negative estimates can be avoided. Still another approach is to note that T_{i1} could be used to predict t_{ij}, j=2, .. K, a prediction which could be combined with the original time-use observation into an estimate of T_{ij}, j=2, .. K.

Assume the following model,

$$t_{ij} = g_j(T_{i1}, \theta) + \varepsilon_{ij} \tag{20a}$$

$$\varepsilon_{ij} = \eta_{ij} + U_{ij} \tag{20b}$$

where g_j is a function to become specified as to its form. It depends on a vector θ of unknown parameters. The random errors ε_{ij} have two components. η_{ij} captures inter individual heterogeneity while U_{ij} is a random error arising because only a few days of a year have been observed. If all 365 days were surveyed for every respondent then U_j would varnish. Thus

$$E(U_{ij} / i) = 0; \tag{20c}$$

$$E(\eta_{ij} / i) = \eta_{ij}; \tag{20d}$$

$$E(\eta_{ij}) = 0; \tag{20e}$$

We also assume that

$$E(U_{ij}^2 / i) = \sigma_u^2 \tag{20f}$$

$$E(\eta_{ij}^2) = \sigma_\eta^2 \tag{20g}$$

Assumption (20f) implies that day to day variability in time-use is the same for everyone. This is probably not a fully realistic assumption.

From the identity

$$\sum_j t_{ij} = T, \ \forall i \tag{21}$$

it follows that,

$$\sum_j \varepsilon_{ij} = 0, \ \forall i \qquad (22)$$

$$\sum_j g_j(T_{i1,}\theta) = T, \ \forall i \qquad (23)$$

These relations imply that the covariance matrix of ε_{ij} is singular and that one time-use activity can be derived from the remaining K-1.

The "parameter" we want to estimate is

$$T_{ij} = g_j(T_{i1}, \theta) + \eta_{ij} \qquad (24)$$

If we could estimate θ and η_{ij} then

$$\hat{T}_{ij} = g_j(T_{i1}, \hat{\theta}) + \hat{\eta}_{ij} \qquad (25)$$

is an estimator of our target parameter. The estimation of η_{ij} requires panel data, i.e. repeated observations in the course of a year. Without panel data it is necessary to assume that T_{i1} carries all information about individual heterogeneity such that $\eta_{ij} = 0$.

$$\hat{T}_{ij} = g_j(T_{i1}, \hat{\theta}) \qquad (26)$$

is then a predictor of T_{ij} with the error $g_j(T_{i1}, \hat{\theta}) - g_j(T_{i1}, \theta)$. The original observation t_{ij} is of course also a predictor of T_{ij}. A weighted combination of these two predictors would then give the final estimator.

$$\tilde{T}_{ij} = W\hat{T}_{ij} + (1-W)t_{ij} \qquad (27)$$

It can be shown that the variance of this estimator is minimized if the weight is chosen such that[10]

$$\mathrm{W} = \mathrm{Var}(\mathrm{t}_{\mathrm{ij}})/(\mathrm{Var}(\hat{\mathrm{T}}_{\mathrm{ij}}) + Var(t_{ij})) \qquad (28)$$

Using (20f) and remembering that the target parameter is T_{ij}, then

$$W = \sigma_u^2 / (Var(g_j(T_{i1}, \hat{\theta})) + \sigma_u^2) \qquad (29)$$

Because $Var(g_j(T_{i1}, \hat{\theta})$ is a function of T_{i1} and conditional on T_{i1}, the weight W is also a function of T_{i1} and it thus depends on each individual observation.

[10] It is here assumed that the two estimators are uncorrelated. The correlation which arises because t_{ij} is used when θ is estimated is likely to be so small that it can be neglected.

The estimator becomes feasible only if the two variance components can be estimated. It is usually possible to get at least a large sample approximation of $(Var(g_j(T_{i1},\hat{\theta}))$ using $Var(\hat{\theta})$, and σ_u^2 can be estimated by the residual variance from the regression behind eq (26).

The estimator (27) is not in general an unbiased estimator. Its bias depends on the functional form g_j, the properties of the estimates of θ and the sampling design. Usually the sampling design is such that t_{ij} is an unbiased estimator, so the bias depends on the bias of $\hat{T}_{ij}$. If g_j is linear in T_{i1}, LS-estimates of θ make $\hat{T}_{ij}$ unbiased, but depending on the properties of ε_{ij} a linear function is not likely to be a realistic assumption because the range of variation of t_{ij} is restricted to (0,T). A linear function might easily give negative predictions.

Neglecting the fact that the weights have to be estimated and introducing the notation σ^2 for the denominator of the weight expression (29), the variance of the estimator (27) becomes,

$$Var(\tilde{T}_{ij}) = \left(\frac{\sigma_u^2}{\sigma^2}\right)^2 (\sigma^2 - \sigma_u^2) + \left(\frac{\sigma^2 - \sigma_u^2}{\sigma^2}\right)^2 \sigma_u^2$$
$$= \sigma_u^2 \left(1 - \frac{\sigma_u^2}{\sigma^2}\right) \qquad (30)$$

The factor in parenthesis is always less than one and it gives the gain in precision compared to using only the original time-use observation. The smaller the relative variance of $\hat{T}_{ij}$ the larger the gain in precision becomes. In practice the weights are estimated, which will contribute to an increased variance, and the resulting gain in precision would have to be balanced against any bias. This is probably best done in sampling experiments.

A simple illustrative sampling experiment was run with only 20 replications using the following exponential model,

$$T_{i2} = 4.4622 + 28.7748\ (0.9592)^{T_{i1}} \qquad (31a)$$

$$t_{i2} = T_{i2} + U_{i2} \qquad (31b)$$

$$U_{i2}\ \textit{is distributed}\ N(0,\sigma_u),\ \textit{where}\ \sigma_u = 2{,}4. \qquad (31c)$$

T_{i1} is normal weekly hours of market work including overtime and secondary jobs, obtained from the 1984 HUS survey. In order to predict T_{i2} for each of 1867 individuals, the estimator (27) was used in two alternative forms, one with the correct exponential functional form and one with a linear predictor. Based on the 20 replications we thus obtain bias and MSE estimates for each individual. The mean, standard deviation of the mean, and min and max values for these estimates are displayed in Table 3. For both estimators the mean bias is very small, just a few per mille of the mean time-use in activity 2, but the MSE is much larger when the erroneous linear predictor is used. However, both estimators give a considerable gain in precision compared to raw time-use data. The explanation is that most weights w are well above 0.9 which implies that the time-use estimates are very close to the model predictions.

The ratio estimator (16) and the combined predictions/raw data estimator (27) [11] were also applied to real time-use data for household work excluding care activities, and compared to raw time-use data. Table 4 gives a few statistics. There is not much difference in means but the standard error of the ratio estimator is about 90 per cent of the standard error of raw data and the standard error of the prediction estimator only 60 or 40 per cent depending on the weights. One may also note that with the latter estimator all zero observations have been transformed into nonzero observations. The three sets of time-use data were finally used to estimate a simple linear regression function for household work. Zero observations were treated equally to nonzero observations. The regression model as such should not be taken too seriously but the resulting estimates reflect the properties of the four data sets. The ratio estimator data give a marginally better fit than the raw time-use data. The regression coefficients are approximately the same but they are somewhat better determined when the ratio estimator data are used. With combined predicted/raw data the regression slopes are estimated with an even greater precision, but most of the slope estimates have now been pulled towards zero. At least in this example use of auxiliary data will thus not only increase precision but also give a drastic change in the point estimates. [12] Provided we have used a good predictor these estimates might capture the relation between weekly hours of household work and the exogenous variables better than those obtained from raw data and ratio estimator data. However, another interpretation of the difference in slope estimates is that the error in the predictive relation is correlated with the explanatory variables in the regression model, which implies that there is unmeasured individual heterogeneity left in this error component. Thus, our auxiliary measure of market work does not completely capture all individual heterogeneity in household work. To capture remaining heterogeneity we would either have to use panel data or find additional auxiliary variables. The latter alternative would in fact mean that one moves from an attempt only to reduce random noise in time-use data into causal modeling.

This analysis has only been illustrative and indicative of the potential for using external information to improve on time-use estimates. More work is needed on the design of good approaches and on the exploration of their properties. It is conceivable that approaches tailored to estimate well-defined population statistics or parameters of a particular model will show more successful than attempts to design general purpose calibration methods.

One preliminary conclusion is though that auxiliary information might reduce the problem with zero observations. The larger time-span for which an activity is observed the smaller the probability to observe zero hours become. As we have seen the use of auxiliary data to calibrate time-use also tends to compress the distribution, and in particular reduce the frequency of “zeros”. Although time-use is always nonnegative these calibration estimators might reduce the need to use Tobit-type models, which rely so heavily on the assumption of one particular distribution, the normal distribution.

[11] To simplify the computations optimal weights were not used, but the same weights were applied to all observations, either w = 0.5 or w = 0.9.

[12] The difference in result is not caused by the neglect of the zero observations. They comprise only about 10 per cent of the sample and a Tobit model gives approximately the same results as the LS-regression on raw data.

Zero observations

The frequent occurrence of zero observations in time-use data is usually handled by the application of a Tobit model. This is sometimes but not always a good idea depending on why activities are not observed and how the Tobit model is specified. Suppose there is an activity, which is typically performed on one day of the week, for instance on Fridays. If the survey design is such that all days of the week are surveyed the result will become many ”zeros” from days other than Fridays. As we will see a straightforward application of a Tobit model is not such a good idea, not even if we would introduce a dummy variable for Fridays to capture the day of the week effect.

In designing time-use surveys it is common to build variations in time-use by day of the week and season of the year into the survey design. For instance, many time-use surveys stratify according to season and to workdays and weekends. Suppose n_1 workdays and n_2 weekend days are sampled for each respondent. It is then natural to estimate weekly time-use by $5\hat{t}_1 + 2\hat{t}_2$ where $\hat{t}_1$ is the individual mean for workdays and $\hat{t}_2$ for weekends. By using this stratified estimator one might think that the variation by day of the week in time-use has been taken care of. It works fine for estimates of means and totals, but it does not work in modeling, not even if the sampling design is noninformative in the sense that the selection probabilities are not functions of the model parameters of interest. To see this, consider the rather extreme example already mentioned above in which an activity is only done on Fridays. Assume also that time-use in this activity y (on Fridays) depends linearly on another variable x except for a random disturbance. A sample of data could look like Figure 3. A few Friday observations show a linear association with x while most of the observations lie on the x-axis. A linear regression model in x fitted to this sample would of course seriously underestimate the dependence on x in an attempt to reach a compromise between the true slope and the zero slope. The introduction of a dummy variable for Fridays would not change this. One would also have to use the interaction between x and the dummy variable to allow the slope to depend on day of the week. Similarly, a Tobit model fitted to the whole sample will give inconsistent estimates of the effect of x because the model will try to make x explain all the zeros. If a dummy variable is introduced for non-Friday observations the estimate of the effect of this dummy will tend towards infinity. In this case a Tobit model should be fitted only to Friday observations.

Let’s now turn to a slightly different example. Assume that there is heterogeneity in people’s time-use such that some people never do an activity while others do it depending on how much of the property X they have. For instance, some people never go to the opera independently of the charge for a ticket and independently of their wage rates and incomes. Similarly some people do not have any taste at all for attending sports events, and some do never take car trips because they do not have a driver’s license or do not own a car. The difference between this example and the previous one is that now we do not know to which group a respondent belongs.

To fix ideas assume the following model.

$$y^* = X\beta+\varepsilon; \quad E(\varepsilon|X)=0; \quad F_\varepsilon \qquad (32a)$$
$$y = y^* \quad \text{if } (y^*>0)\cap(D=0) \qquad (32b)$$
$$y = 0 \quad \text{if not } (y^*>0)\cap(D=0) \qquad (32c)$$

$$P(D=1) = P_1 \qquad (32d)$$

where F_ε is the distribution for ε and P_1 the probability that a randomly observed respondent does not fancy this particular activity. P_1 is an unknown parameter. The likelihood function then becomes,

$$L = \prod^{n_0}\left[P_1 + p(y^* \leq 0 | D = 0)(1 - P_1)\right]\prod^{n_1} f_\varepsilon(y - X\beta)(1 - P_1) \; ; \qquad (33)$$

and the first order condition with respect to β can be written,

$$\sum^{n0} \frac{1 - P_1}{P(y=0)} \frac{\partial F_\varepsilon(-X\beta)}{\partial \beta} + \sum^{n1} \frac{1}{f_\varepsilon} \frac{\partial f_\varepsilon(y - X\beta)}{\partial \beta} = 0 \, ; \qquad (34)$$

The first factor after the first summation sign in (34) is a weight signifying the importance given to zero observations. The larger share of respondents who fancy the activity the bigger this weight becomes and if everyone would belong to this group, then the likelihood function simply becomes the likelihood for a common Tobit model.

All this might appear rather trivial, and it is, because all these two examples suggest is that a Tobit model cannot be applied routinely to time-use data. One has to consider why zero observations are generated and model these mechanisms in a realistic way.

There are also other well-known problems common to all applications of the Tobit model. For instance, the model assumes that the latent variable is normally distributed with a constant variance. I have not found much testing of these two assumptions in the time-use literature. Additional problems arise when limited dependent variables enter as endogenous explanatory variables in interdependent equation systems. For instance, in a model with two spouses who in general contribute time both to market work and household work in such a way that they maximize household utility subject to budget and time constraints, and to a household production function of the CES or Cobb-Douglas type, the first order conditions for maximum can be used to derive a two-equation interdependent model for the spouses' allocation of time to household production. Time input from one spouse depends on the other spouse's input. Although we have learned how to estimate such models, estimation and testing is not standard.

These comments illustrate the general need to carefully integrate into the econometric model the economic process analyzed with the properties of the measurement procedures. In practice the obvious is sometimes overlooked.

7. Conclusions

In my review of the literature as a preparation for this paper I found less work using time-use data than I expected to find. Assuming this is not just the result of misguided expectations and insufficient search, what is the explanation? Have economists interested in household behavior seen the identification problem and refrained from using time-use data as long as there are no measures of output from household production? Have the high noise to signal ratio of time-use data and all other problems with these data contributed to the same decision? In spite of the difficulties my personal evaluation is that time-use data are underutilized. As indicated above there are ways to handle the statistical properties of time-use data and at least partly overcome the difficulties with noisy data, zero observations etc. The fact that time-use data alone do not carry enough information to allow identification of the household production structure is more difficult to come around. I do not have much sympathy for approaches, which p impose strong untested assumptions about preferences and household technology just for the sake of identification. We will never know if the results are driven by these assumptions or by data. But I do believe that we can learn much more about time-use behavior from estimating "mongrel" time-use functions.

There is information in time-use data about the timing and spacing of activities which has hardly been used at all. Most users of time-use data transform the diaries into files containing the total time used by type of activity and day. This aggregation destroys the original information about timing and spacing in a day. Such data could be useful in a study of the "fine structure" of household behavior, in analyzing the length and allocation of market work hours, and in studies of the distribution of stress, passiveness and solitude from a social policy perspective, just to mention a few examples. A recent exception is the Hamermesh (1996) study of who is doing market work in the evening.[13]

The few empirical results we have indicate that one cannot disregard household production and leisure activities even if one is primarily interested in labor supply. Models, which assume that utility is just a function of one aggregate leisure activity, will give biased estimates of wage rate and income incentives to work in the market. Common sense and the measures of "process benefits" also suggest that a model which assumes that only leisure contributes to well-being has become an unrealistic representation of people's behavior. In a modern society market work does not only yield income, and household production does not only yield domestic output, they both give direct utility (process benefits) as well. People enjoy doing these activities. Any model, which does not recognize this, has the danger of biasing estimates of economic incentives.

In this review we have repeatedly found studies in which people used various imputation strategies to compensate for data which were not fully adequate for their purpose. Either important variables were missing or potential respondents were missing. For instance, in some cases only one of the spouses was interviewed while both were needed. Although imputations may give useful results, there are always assumptions involved, which cannot be tested. The only satisfactory approach is to collect richer data. We need time-use data for both spouses and also for their children, we need joint time-use and expenditure data, and above all a much greater effort is needed in collecting output measures from household production.

To conclude, even if time-use data did not provide all the opportunities we might have hoped for, there is still territory to cover using the existing type of data. If further time-use studies become better integrated in the research community and based on theory and clear research objectives, the future for time-use research looks even brighter. Time-use data are too valuable to be left only to statistical agencies and national accounting people.

---

[13] Using the German Socioeconomic Panel and the U.S. Panel of Income Dynamics he found, for instance, that evening work was more common among workers with relatively little human capital and low incomes, and more common among men than among women. He also showed that young children increased the likelihood that one spouse would work in the evening and that women then took a disproportionate share of such work.

References

Apps, P.F. and Rees, R.,1996, "Labour supply, household production and intra-family welfare distribution", *JulfPulcEcmc* 60 (2), 199-219.

Aslaksen I. and C. Koren, 1995, "Det ubetalte husholdsarbeidet – omfang og fordelning" (Unpaid Household Work – Measurements for National Accounts and Income Distribution), *Tdkffmfufkg* No. 1 pp 3-30.

Becker G.S., 1965, "A Theory of the Allocation of Time", *EcmcJul* 75,493-517

Behrman, J.R., 1997, "Intrahousehold distribution and the family" in *HdkfPpul dFmlyEcmc,* edited by M.R. Rosenzweig and O. Stark, Elsevier Science B.V.

Biddle, J.E. and D.S. Hamermesh, 1990, "Sleep and the Allocation of Time", *Julf PlclEcmy* 98(5:1), 922-943.

Bonke, J., 1992, "Distribution of Economic Resources: Implications of Including Household Production", Review of Income and Wealth, 38(3), 281-293

Brown, B. and D. Saks, 1975, "The Production and Distribution of Cognitive Skills Within Schools", *JulfPlclEcmy* , Vol. 83, June, 571-593.

Bryant, W.K. and C.D., Zick, 1985, "Income Distribution Implications of Rural Household Production", *AmcJulfAgcululEcmc* , 67 (5), 1100-1104.

Calhoun, C.A., and T.J. Espenshade, 1988, "Childbearing and Wives", *PpulSud* , Vol. 44, 41-60.

Carlin, P.S. and L. Flood, 1997, "Do children affect the labor supply of Swedish men? Time diary vs. survey data.", *LuEcmc.AIlJul* . 4(2), 167-183.

Coleman, J. S., 1994, "Social Capital in the Creation of Human Capital," *AmcJulf Sclgy* , (supplement): S95-120.

Dankmeyer, B., 1996, "Long Run Opportunity Costs of Children According to the Education of the Mother", *JulfPpulEcmc* , Vol. 9, 349-361.

DeVille, J-C and C-E Särndal, 1992, "Calibration Estimators in Survey Sampling", *J.fh AmcSclAc* 87, 418, 376-382.

Fitzgerald, J.M., M.S. Swenson and J. H. Wicks, 1996, "Valuation of Household Production at Market Prices and Estimation of Production Functions". *RwfIcmdWlh,* 42(2), 165-180.

Flood, L. and N.A. Klevmarken, 1990, "Tidsanvändningen i Sverige 1984" in *Kll km* . Bilaga 23 till LU90, Nordstedts Tryckeri, Stockholm

Flood, L. N.A. Klevmarken and P. Olovsson,1997, *HuhldMkdNmk Ac(HUS).VlumII-VI* , Department of Economics, Uppsala University, Uppsala

Graham, J.W. and C.A. Green, 1984, “Estimating the parameters f a household production function with joint products”, *ThRwfEcmcdSc* , 66, 277-282.

Gravesen, E.K. & N. Smith,1997, ”Labour Supply, Overtime Work and Taxation in Denmark”, mimea, CLS, Aarhus University.

Gronau, R., 1977, “Leisure, Home Production, and Work – The theory of the Allocation of Time Revisited”,*JulfPlclEcmy* 86(6), 1099-1123

Gronau, R., 1980, “Home production a forgotten industry”, *ThRwfEcmcd Sc* , 62, 408-416.

Gronau, R.,1986, ”Home Production - A Survey” in *HdkfLEcmc* , Vol. I, edited by O. Ashenfelter and R. Layard, Elsevier Science Publishers BV

Gustafsson, B. and U. Kjulin, 1994, “Time Use in Child Care and Housework and the Total Cost of Children”, *JulfPpulEcmc* , Vol. 7, 287-306

Gustafsson S. and F. P. Stafford, 1997, “Childcare and Economic Efficiency in Western Industrial Countries, 1900 and 2000”, forthcoming conference volume of the Arne Ryde Symposium, Lund University

Hamermesh, D.S., 1996, ”The Timing of Work Time: Evidence from the U.S. and Germany”, *Kjukuplk(AppldEcmcQuly),* Vol 42, 1-22.

Hansen, W. Lee, 1963, “Total and Private Rates of Return to Investment in Schooling”, *JulfPlclEcmy* , Vol. 76, April, 128-140

Hanustreck, E.A., 1992, “The Trade-Off between Child Quantity and Quality”*Julf PlclEcmy* , 100, 84-117.

Joshi, H, 1990, “The Cash Alternative Costs of Childbearing: An Approach to Estimation Using British Data”,*PpulSud* , Vol. 44, 41-60.

Joshi, H, 1994, “The Forgone Earnings of Eurpoe’s Mothers”, in Efert-Jaff O. (ed) *Sdd fLgdFml:OdAly* , John Libbey, London.

Jenkins, S.P. and N.C. O’Leary, 1996, “Household Income Plus Household Production: The Distribution of Extended Income in the U.K.”*RwfIcmdWlh* , 42(4), 401-419.

Juster, F. T. , 1985, ”Preferences for Work and Leisure”, in F. Thomas Juster and Frank P. Stafford, (eds.), *TmGddWll-Bg* , Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor.

Juster, F.T. and F. P. Stafford,1985, *Tm,GddWll-g,* Ann Arbor, Institute of Social Research, University of Michigan

Kapteyn, A., 1994, "The measurement of household cost functions. Revealed preference versus subjective measures", *JulfPpulEcmc,* 7(4) 333-350.

Karlton, G.,1985, "Sample Design Issues in Time-use Diaries", in Juster and Stafford(1985) pp 93-131.

Klevmarken, N.A. and P. Olovsson, 1993, *HuhldMkdNmkAc. PddudCd1984-1991,* Vol I and II, The Industrial Institute for Economic and Social Research, Almquist and Wiksell International, Stockholm

Klevmarken, N.A. and F. P. Stafford, 1997, "Time Diary Measures of Investment in Young Children", Working Paper 1997:8, Department of Economics, Uppsala University

Koreman, P. and A. Kapteyn, 1987, "A Disaggregated Analysis of the Allocation of Time within the Household", *JulfPlclEcmy* 95, 223-249.

Leete, L and J.B. Schor, 1994, "Assessing the Time-Squeeze Hypothesis: Hours Worked in the United States, 1969-89", *IdulRl* . Vol.33.No1, 25-43.

Malathy, R., 1994, "Education and Women's Time Allocation to Nonmarket Work in an Urban Setting in India", *EcmcDlpmdCululChg* , 42(4), 743-60

Merz, J. and D. Kirsten, 1998, "Valuing Household Production and its Implication to the Extended Welfare Distribution – A Panel Analysis for Professions, Self-Employed, and Unemployed", Paper presented at the International Conference on Time Use, Lüneburg, April 22-25, 1998.

Quah, E., 1993, *EcmcdHmPduc.Thydmum* . Aldershot, U.K.; Brookfield, Vt. and Sydney: Ashgate, Avebury.

Radner R., 1986, *Tm(AlumHfJcMchk*), Second Edition, Minneapolis, University of Minnesota Press,

Reid, M.G., 1934, *EcmcfHuhldPduc* . New York: John Wiley

Reid, M.G. , 1947, "The Economic Contribution of Homemakers", *AlfhAmc AcdmyfPlcldSclScc* , May, 61-69

Smith, J. and M. Ward, 1989, "Women in the Labor Market and the Family", *Julf EcmcPpc* , Vol. 3, Winter.

Stafford, F. P., and M. Sundström, 1997, "Time Out for Childcare: Signalling and Earnings Rebound Effects for Men and Women", forthcoming, *Lu* r Vol. 10(3), 609-629..

Szalai, A., 1966, Trends in Comparative Time Budget Research", *Amc Bhl Sc* , 9(9), 3-8.

Szalai, A. (ed.) 1972, *Thufm,* The Hague: Mouton.

Van Praag B.M.S., 1968, *Iddulwlffucdcumh* . North-Holland, Amsterdam.

Wagman, B. And N Folbre, 1996, “Household Services and Economic Growth in the United States 1870-1930”, *FmEcmc* 2(1), 43-66

Table 1. Alternative estimates of weekly work hours by gender 1983/84 and 1992/93

	1983		1984		1992		1993	
	m	f	m	f	m	f	m	f
<u>"Survey estimates"</u>								
HUS data								
Annual hours/weeks worked	41.8	40.3			41.9	35.0		
	(0.2)	(0.3)			(0.3)	(0.3)		
5*hours latest workday			43.0	38.5			44.5	38.9
			(0.3)	(0.4)			(0.4)	(0.5)
Normal hours incl. secondary jobs*			41.8	35.0			42.5	37.5
			(0.5)	(0.5)			(0.2)	(0.2)
Latest week worked							41.9	33.6
							(0.4)	(0.4)
Labor force Survey. Statistics Sweden								
All in work and temporarily absent, 16-64 years old			35.1	25.7			34.6	26.1
All in work 16-64 years old			40.6	31.7			39.9	32.5
Employed 16-64(74) years old			39.7[#]	31.4[#]			38.6	32.2
<u>Time-use estimates</u>								
Excluding break time								
All 20-64 years old			22.4	19.1			34.1	24.4
			(0.8)	(0.7)			(0.9)	(0.8)
If in work or temporarily absent, 20-64 years old			29.6	25.8			39.0	29.6
			(0.9)	(0.8)			(0.9)	(0.8)
Including break time								
All 20-64 years old			25.4	21.8			43.1	31.1
		(0.9)	(0.8)			(1.1)	(1.0)	
If in work or temporarily absent 20-64 years old			33.7	29.5			49.2	37.8
			(1.0)	(0.9)			(1.2)	(1.1)

*Employed only. The questions used were phrased: "On average, how many hours per week are you currently working at your primary job, including both paid and unpaid overtime", "Do you have another job in addition to your primary job?" and if YES, "How many hours do you spend on your other job(s)?" (Replies given per day, week, month or year).

[#] 16-74 years old

<u>Note 1</u>. The estimates for "Annual hours/weeks worked" were obtained using a sequence of questions about weeks worked in full-time and part-time work last year and about the average number of hours during those full-time and part-time weeks respectively.

<u>Note 2</u>. The hours of work questions in the Labor Force surveys were: "The questions which follow apply to a certain week, Monday the … to Sunday the …, that is week no …. How many hours did you work that week in your main job? How many hours in any secondary job?"

<u>Note 3</u>. Time-use estimates include the sum of work hours in primary and secondary activities, but market work as a secondary activity is very small. Secondary jobs are also included. Breaks include lunch, coffee breaks, other breaks, personal errands and telephone calls while at work. The sample is limited to respondent who gave two complete time-use interviews (one work day and one weekend day). If the respondent had a job at the time of the workday time-use interview and had not been a way for more than 8 week the respondent was classified as in work or temporarily absent.

Table 2. Total Value of Child Investments, Time and Goods. Average per Child, Two Parents, Sweden, 1993

	Children less than 7 years old		Children 7-17 years old	
	Hours/year	SEK/year	Hours/year	SEK/year
Time input				
-direct child care	250	20000	180	14400
-household work	200	8000	70	5600
-time with children	3700	59200	2850	45600
	-------		-------	
-all time input	4150	87200	3100	65600
Family Goods Input		30000		30000
Publicly Provided Goods and Services Input		62000		73000
		--------		--------
Total		179200		168600

Source: Klevmarken & Stafford (1997), Table 12

Table 3 Bias and MSE from a small sampling experiment

	Mean	*Std.dev. of mean*	*Min*	*Max*
Bias				
$\sigma_u = 2$				
Nonlinear predictor	-0.0019	0.0003	-0.1222	0.0155
$\sigma_u = 4$				
Nonlinear predictor	-0.0044	0.0002	-0.0102	0.0804
Linear predictor	-0.0100	0.0004	-0.0517	0.1166
MSE				
$\sigma_u = 2$				
Nonlinear predictor	0.0054	0.0003	0.0011	0.2598
$\sigma_u = 4$				
Nonlinear predictor	0.0059	0.0003	0.0022	0.3692
Linear predictor	0.0246	0.0010	0.0106	0.9408

Note: Number of observations 1867, and number of replications 20.

Table 4 Alternative time-use estimates for household work (hours/week)

	Raw time-use data	*Ratio estimates*	*Combined predictions/raw data* *w=0.5*	*w =0.9*
Mean	12.44	12.03	12.42	12.42
Std.error	11.21	10.01	6.64	4.25
Min	0	0	2.14	3.86
Max	63.41	50.31	39.63	35.49

Note 1: The combined estimator used an exponential predictor and the same weight (w) for every observation.
Note 2: These statistics are based on 1475 observations

Table 5 Alternative regression estimates of a simple function for time-use in household work

	Raw time-use data			Ratio estimates			Combined model predictions/raw data w=0.5			w=0.9		
Explanatory variable	*Coef.*	*Std. err.*	*P-value*	*Coef.*	*Std. err.*	*P-value*	*Coef.*	*Std. err.*	*P-value*	*Coef*	*Std err.*	*P-value*
age	.061	.021	0.005	.056	.020	0.006	.038	.012	0.002	0.021	0.008	0.012
koen	3.56	.646	0.000	3.28	.609	0.000	2.16	.366	0.000	1.056	0.248	0.000
school	-.027	.093	0.769	-.013	.087	0.880	-.043	.052	0.411	-0.056	0.036	0.116
univ	.079	1.05	0.940	.105	.989	0.915	-.175	.594	0.768	-0.380	0.403	0.347
single	4.72	.741	0.000	4.59	.698	0.000	2.84	.420	0.000	1.346	0.285	0.000
antbarn	1.04	.275	0.000	1.04	.260	0.000	.502	.156	0.001	0.070	0.106	0.506
cons	-.914	1.77	0.606	-.393	1.67	0.814	4.52	1.00	0.000	8.873	0.862	0.000
R-square	0.098			0.100			0.112			0.073		
No of obs	858			856			858			858		

Note: The combined estimator used an exponential predictor and the same weight (w) for every observation.

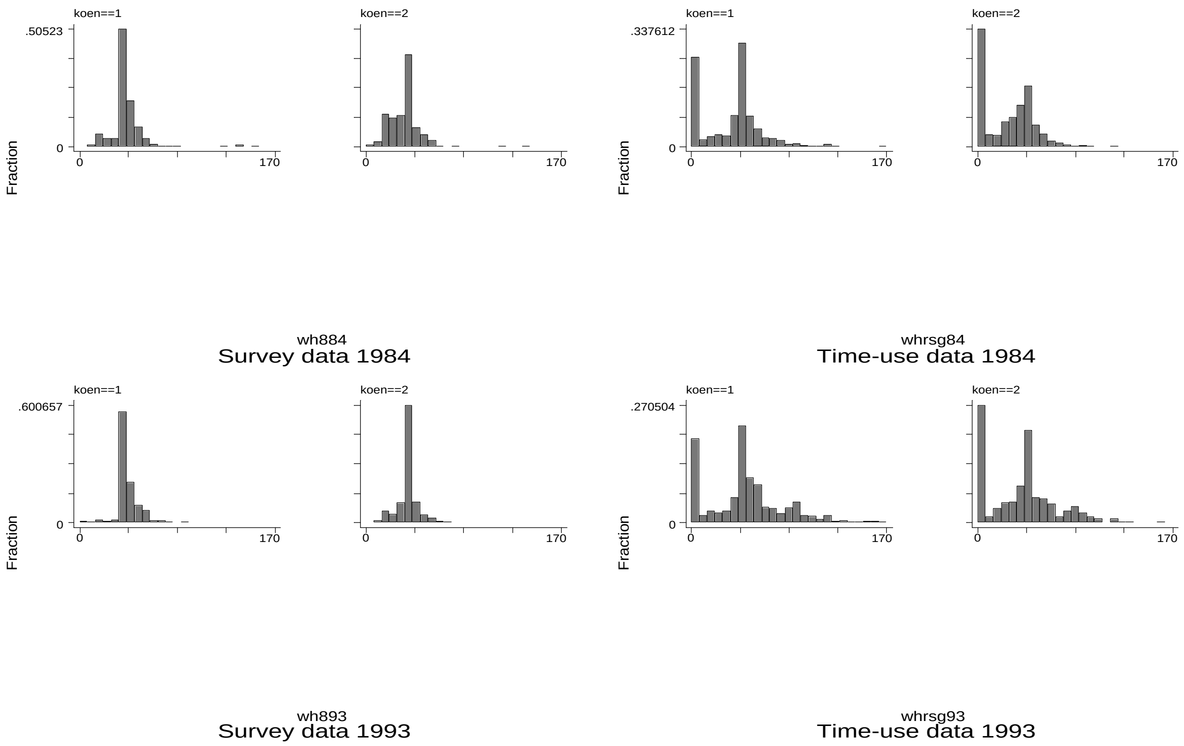

Figure 1 Weekly workhours by gender and data type

Figure 2 Diagrammatic illustration of the "Zero-observation" problem