

Using auxiliary data to compensate for noisy time-use data*

N. Anders Klevmarken¹

Department of Economics
Uppsala University

1. Introduction

Time-use data have a high noise to signal ratio and other properties that complicate analysis. These properties include,

- very short spells of data that permit daily deviations from average or normal behavior dominate 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.

These problems are best handled by improved data collection procedures. No statistical or econometric procedure can fully compensate for bad data, but when data have already been collected they should be used as efficiently as possible. Econometric theory suggests that depending on the model structure it will be inefficient to estimate a partial relation rather than a complete model. For instance, to estimate a demand function for leisure rather than a complete model for the allocation of time between leisure, household work and market work is likely to become inefficient because these activities are interdependent. There might, however, be other considerations than efficiency such as sensitivity to misspecification or complexity of the empirical analysis that leads to a partial approach. Time-use surveys usually collect supplementary data about the respondent and the respondent's household. Some of these variables we will use in our economic models to explain time-use behavior, but sometimes there are also variables that contain information about time-use that are typically

*Paper prepared for the Levy Institute time-use conference, October 28-29, 2005

¹ Useful comments and suggestions from Matias Eklöf are gratefully acknowledged.

not used in economic models jointly with the time-use estimates. For instance, in some time-use surveys one asks respondents if the designated day was a typical or atypical day, in others time-use estimates are supplemented by more conventional survey questions about (normal) hours of market work. Although not necessarily part of the economic model this auxiliary information could be used in the econometric model to improve efficiency and possibly also correct for some systematic errors in the parameter estimates. This is the topic of this paper. An analysis of these issues might not only suggest methods that improve on the quality of parameter estimates, but also suggest what kind of information it would become useful to collect in future studies.

Alternative approaches to incorporate auxiliary information

In the statistical literature there are a number of approaches suggested and designed for various combinations of applications and type of auxiliary information.

One might distinguish between general-purpose approaches and applications of model specific approaches. In the first approach the observed value (or missing value in the case of nonresponse) is replaced by an imputed value or adjusted in a reweighing procedure. The new value can then be used for many purposes. The imputed value is obtained from a function that either links the observation to become imputed to another similar observation (for instance the hot deck method) or produces a predicted value that is imputed.

Reweighting originates from the sampling literature and has thus primarily been used in inference to a finite population. Examples are methods like calibration-estimation including post-stratification (DeVilte & Särndal, 1992 and Lundström & Särndal 1999). In this case one uses information about population distributions, for instance, totals or means for auxiliary variables that are also observed in the sample. If these variables are correlated with variables of interest there is a potential gain in using this auxiliary information. It is, however, not necessary to have information about population entities. Even auxiliary information that only pertains to the sample might become useful.

Another approach more well-known to econometricians could be labeled model-based joint estimation. As the name suggests it is model-specific and does not have the general purpose character. In this case the nature of the auxiliary information is formalized in model terms. It could either take the form of deterministic or stochastic constraints on the “deep” model parameters, or the form of a model that formalizes the relation between theoretical variables and measurements by being explicit about measurement errors. In this approach it is possible to have more than one type of measurement on a theoretical variable. An example is LISREL.

In studies of time-use one finds both inferences to finite populations in the form of descriptive tables of time-use distributions, means and totals, and inferences to data generation processes represented by economic models. Most of the approaches mentioned above could thus become of interest.

2. Calibration and imputation approaches

For the reasons given above time-use data are typically noisy. 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. Klevmarken (1999) includes a discussion of a few calibration/imputation approaches to compensate for noisy time-use data.

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 non-market 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. Survey data on, for instance, hours worked in the market in a normal day or survey estimates on annual hours worked might be useful even if population totals (means) are unknown for these variables. If population totals are known that is even better.

Klevmarken(1999) suggested the following simple calibration approach,

$$(1a) \quad \hat{t}_{ij} = \frac{T - T_{i1}}{T - t_{i1}} t_{ij}; \quad j = 2, \dots, k$$

$$(1b) \quad \hat{t}_{i1} = T_{i1};$$

where T is total number of available hours (24 hours in a day), t_{ij} the time-use measure of activity j for individual i., T_{i1} the auxiliary measure of time-use in activity 1 (market work). One immediately finds that this estimator satisfies the adding up constraint,

$$(2) \quad \sum_j \hat{t}_{ij} = T; \quad \forall i$$

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

$$(3) \quad \frac{1}{n} \sum \hat{t}_{ij} = (T - T_{i1}) \frac{1}{n} \sum \frac{t_{ij}}{T - t_{i1}}; \quad \forall i$$

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 (3) will have a larger variance than an unadjusted estimator and should not be used.

There are a few disadvantages of this estimator. First, every activity $j=2, \dots, k$ 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.

Observing these disadvantages and that the auxiliary measures might be error prone too (but less so than the time-use estimates) Klevmarken(1999) continued discussing simple modeling approaches that use auxiliary data to predict time-use. These predictions are then combined with time-use data using appropriate weights to give a joint estimate of an individual's time-use in activity j . A simple sampling experiment illustrated these ideas.

He concluded that it was possible to gain in precision using these simple calibration estimators, and that auxiliary data 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. Using auxiliary data to calibrate time-use data tends to compress the distribution, and in particular reduce the frequency of "zeros". These calibration estimators might thus reduce the need to use Tobit-type models, which rely so heavily on the assumption of one particular distribution.

3. Model-based joint estimation

3.1 A latent variable - indicator approach

Assume there is an economic model that relates average or normal time-use per week in an activity y^* to a number of explanatory variables X ,

$$(4) \quad y^* = X\beta + \varepsilon; \quad E(\varepsilon | X) = 0;$$

Also assume that there are two measures of time-use, one survey measure of normal weekly hours of market work and one from a time-use diary. The corresponding measurement equations are,

$$(5a,b) \quad \begin{aligned} y_1 &= y^* + v_1; \\ y_2 &= g(y^*) + v_2; \end{aligned}$$

where g is a known function and v_1 and v_2 random measurement errors. Substituting the economic model into the measurement relations gives a system of (nonlinear) seemingly unrelated regressions with cross equations parameter constraints. Estimating both relations jointly rather than just the first will give a gain in efficiency.

An example: Labor supply

Let equation (4) be a conventional labor supply function using net wage rate and virtual income as the key explanatory variables as well as a few additional socio-economic exogenous variables. Further let y_1 in equation (5a) be an estimate of annual hours worked from a conventional battery of survey questions, while y_2 in eq. (5b) is an estimate of weekly hours of work from a time-use survey. Data were collected for one randomly selected week day and one randomly selected weekend day. The estimate of weekly hours was then obtained as the sum of five times the week day estimate and twice the week end estimate. The function

g in eq. (5b) is in this case simply a proportionality factor and we will below use 1/50 because there are approximately 50 weeks of work in a year.² To simplify the interpretation of the results we will just multiply the estimate of the weekly hours by 50 and thus get a time-use data based estimate of annual hours of work. We need, however, also to add a few explanatory variables to try to capture why people might not work in a particular week although they have a job and normally work. The key variable to do this job is just a dummy variable that indicates if the week day interview took place in the summer when many people have long vacations. In addition we will try a few more conventional variables such as the number of children, years of schooling and if anyone in the household has health problems influencing their ability to work. Table 1 gives descriptive statistics for the sample used.

There is a nonignorable number of cases with zero work hours both in survey data and in time-use data. To accommodate this fact the model was estimated as a bivariate tobit model. y_1 and y_2 in eq. (5) above should thus be seen as latent variables while the corresponding observables are censored at zero hours. Table 2 first presents single equation tobit estimates, while Tables 3 and 4 give bivariate tobit estimates using alternative sets of explanatory variables.³

Using annual data the net wage rate and virtual income effects are both significantly negative. Adding more explanatory variables does not change these estimates much but reduces the residual variance. All these exogenous explanatory variables but the bad health dummy contribute significantly and have expected signs.

The corresponding estimates from time-use data are less precise and the net wage rate effect is now positive, but not significantly so. The residual variance is almost twice as large and the number of zero cases more than twice as large as compared with survey data. The large errors in time-use data thus make the parameter estimates uncertain.

Estimating the two functions jointly and allowing for a non zero correlation between them should increase efficiency at least asymptotically. A comparison of the estimates in Table 2 with those in the first columns of Table 3 does not reveal any gain in efficiency. The standard errors in Table 3 are generally larger than those in Table 2.⁴ Most of the estimates are qualitatively the same even if the levels have now changed. We still have negative survey data estimates of the net wage rate and virtual income effects, and a positive but insignificant time-use estimate for the net wage rate effect.

Constraining the wage rate and income effects to be equal in the two equations pulls the estimates closer to zero than the pure survey estimates but not very far. We get significantly negative effects for both variables. The precision of these estimates increases much compared to the pure time-use estimates and just a little compared to the survey estimates. The estimates for the other variables do not change much in level or precision.

Table 4 gives the same kind of results when all exogenous variables are included in both equations. Evaluating the net wage rate and virtual income elasticities at 1550 hours and at the

² In the data used the ratio of the median annual hours and the median (time-use) weekly hours was 52.15 while the corresponding ratio of the means was 48.28.

³ The explanatory variables “net wage rate” and “virtual income” are endogenous to hours of work, but were in fact treated as exogenous.

⁴ The estimates are unfortunately not fully comparable because the results in Table 2 are based on a reduced sample of 820 cases.

means of these two variables we obtain a wage rate elasticity of -0.19 and an income elasticity of -0.17. Among the other results we might note that females work much less than males, that there is a more hump shaped age profile in time-use data than in survey data, and that the time-use based estimates of the other variables are very uncertain. Using the survey based estimates we find that parents with children work fewer hours and that schooling increases hours of work. There is no significant effect of bad health or of being single.

To conclude: In this sample time-use data are so noisy that it is impossible to estimate wage rate and income effects with good precision. Combining these data with regular survey data improves the situation considerably, but most of the information about these two effects comes from the survey data. One might thus conclude that supplementing survey data with time-use data improves the precision of the survey based estimates, but only marginally!

3.2 A constrained estimation approach

Assume that time-use diaries have been supplemented by questions whether the time allocated to specific activities during the designated day was more or less than normal. One could for instance ask: "In this time-use diary you have mentioned that you used XX hours for market work. Is this more or less than normal?"

Assume also that we are interested in estimating a linear regression model regressing time-use in an activity y on a number of explanatory variables X .

$$\begin{aligned} y &= X\beta + \varepsilon; \\ (6) \quad E(y | X) &= X\beta; \\ E(\varepsilon\varepsilon') &= \sigma^2 I; \end{aligned}$$

Define an indicator variable D such that if the response on the supplementary question is

$$\begin{aligned} (7a) \quad & \text{"LESS"} \text{ then } D_i=1 \text{ and } \varepsilon_i > 0; \quad i=1, \dots, n \\ (7b) \quad & \text{"MORE"} \text{ then } D_i=0 \text{ and } \varepsilon_i \leq 0; \end{aligned}$$

where n is the number of observations. For each observation there are data on y_i , x_i and D_i . We now seek estimators of β and σ^2 that do not only use the data y and X but also D_i . This problem can be seen as a least-squares problem subject to inequality constraints.

$$\begin{aligned} (8) \quad & \underset{\beta}{\text{Min}} (y - X\beta)'(y - X\beta) \text{ subject to } k(D_i)\varepsilon_i \leq 0; \forall i \\ & \text{where } k(D_i) = \begin{cases} -1 & \text{if } D_i = 1; \\ 1 & \text{if } D_i = 0; \end{cases} \end{aligned}$$

The inequality is equivalent to

$$(9) \quad k(D_i)(y_i - x_i\beta) \leq 0; \forall i$$

Define a diagonal matrix T such that the first n_1 diagonal elements are -1 and the following $n - n_1$ elements 1 , where n_1 is the number of respondents that responded "LESS" ($D=1$). Then it follows that the inequalities become,

$$(10) \quad T(y - X\beta) \leq 0;$$

If ε is a continuous random variable the equality sign will hold with probability zero and we can treat the inequality as a strict inequality. It then follows from the Kuhn-Tucker conditions (see Judge et.al. 1980, sec. 3.3.1.b) that the constrained estimator is

$$(11) \quad \tilde{\beta} = \hat{\beta} + (X'X)^{-1} X'T'(TX(X'X)^{-1} X'T')^{-1} T(y - X\hat{\beta});$$

where $\hat{\beta}$ is the unconstrained least-squares estimator. Let e be a vector of unconstrained least-squares residuals. The estimator can then be rewritten on the following form

$$(12) \quad \tilde{\beta} = (X'X)^{-1} X'\{y + T'(TX(X'X)^{-1} X'T')^{-1} Te\};$$

This expression demonstrates how the observations on the dependent variable can be adjusted by the auxiliary information to make the unconstrained OLS-estimator optimal given data and external information. There is thus an “imputation strategy” that corresponds to the constrained estimator.

Assume now that the auxiliary information comes in a slightly different form. The response to the supplementary question now is “LESS”, “ABOUT NORMAL” or “MORE”. In this case the response “ABOUT NORMAL” suggests that the residual ε_i does not deviate much from zero. If all observations were of this kind OLS would become an efficient estimator. One might conjecture that replacing T by a diagonal matrix that has zeros in the diagonal elements corresponding to these observations, will give a useful estimator.

3.3 Robust estimation by reweighing

In the above example we had quality information about specific activities. In some time-use surveys one has only asked if the designated day was normal or atypical without further specification. The response “atypical” only tells us that on this particular day the respondent did not allocate her time as usual. Some of these respondents might have stayed at home because they were sick or had a sick child, and some might have worked unusually long overtime and thus had less time for household work and leisure than normal. To some extent it might be possible to capture this kind of atypical behavior or heterogeneity in behavior if the survey includes relevant explanatory variables, for instance, if the family had small children or if the respondent had a job that typically required overtime. Then heterogeneity in behavior could explicitly be included in the model. However, in most cases we have rather poor indicators of heterogeneity and it will never become possible to capture the multitude of reasons why people deviate from their normal behavior. Most of these deviations are buried in random residuals. In a sample of respondents to a time-use survey there will be respondents that normally work very “short hours” and others that just happen to do it on the designated day. Ideally one would like to capture both types of behavior in the econometric model. In practice our modeling is likely to be rather crude and the data available not good enough to allow us to distinguish the two kinds of observations from each other. When we just look at data (through our economic model) we won’t know which observations are “normal” and which are not. The response on the deviation from normal question will help us distinguish the two and tell us which observations should be given less weight.

The sensitivity of OLS to fat tails and outliers is well-known and the statistical literature has suggested a number of robust estimators. Some of them involve reweighing of tail observations to reduce their influence on the final estimate. Explicitly or implicitly these approaches use criterion functions that give less weight to extreme observations. For instance Huber(1973) suggested an estimator with the following criterion function,

$$(13) \quad \sum_{i=1}^n f(y_i - x_i \beta)$$

where

$$(14) \quad \begin{aligned} f(e_i) &= 0.5e_i^2 & \text{for } |e_i| < c \\ f(e_i) &= c |e_i| - 0.5c^2 & \text{for } |e_i| \geq c \end{aligned}$$

and c is a preassigned constant. This estimator minimizes squared deviations for small residuals and absolute deviations for large residuals. In this case all observations are used. In other approaches extreme observations are assigned the weight zero and dropped from the analysis. With these robust estimators it is a problem to know when an observations should be down weighted, what value should be assigned to c in Huber's approach. It might be possible to use the deviation from normal indicator here. In the Huber case one could use the squared deviation part of the criterion function for normal responses and the absolute deviation part for atypical observations. To derive the properties of this estimator one would have to make assumptions about the distributions of y conditional on x in the two regimes. One might also have to assume that the assignment to regime is exogenous relative to the economic process studied.

Robust estimation is sometimes motivated as a device to down weight the influence of observations that have been generated by a DGP other than the one under analysis. The idea behind the suggestions above is thus that we try to model "normal" behavior while some of the observations have been influenced by mechanisms not captured by the model. In the statistical literature "normal" is defined by some distributional assumptions for the majority of the observations and observations that violate these assumptions are given less weight. Usually it is accomplished in an iterative process in which the properties of residuals are analyzed. In the approach suggested above "normal" is defined by the respondent through the response to the deviation from normal question. This could either imply that one could get by with weaker distributional assumptions or that there is scope for testing these assumptions.

Another way of thinking about a deviation from "normal" is to see it as a deviation from a conditional expectation, say $E(y|x)$. In a regression model the residuals then represent these deviations. The response to the deviation from normal question thus tells us if an observation comes from one of the tails of the error distribution or not, and we could use this information to down-weight the observations that come from the tails. It is, however, not obvious by how much. In a maximum likelihood approach one could assume that there are thresholds symmetric around the mean for each response on the survey question. Suppose for instance, that respondents who had a normal day are coded 1, those who had a very unusual day were coded 3 and those in between 2. The response equal to 2 would then mean that the observation came from a region defined by the thresholds $c1$ and $c2$: $E(y|x)+c1 < y \leq E(y|x)+c2$ or $E(y|x)-c2 < y \leq E(y|x)-c1$. Given the data on the survey question it should both be possible to

use this information to improve on the estimates of $E(y|x)$ as well as to estimate the thresholds.

An example: Which children watch TV?

The PSID Child Development Supplement of 2002 includes time-use diaries collected from children of the age 5-18. The diaries were supplemented with a question: “How typical was this day (for that day of the week)?” and the responses were given on a scale from 1 to 5 (Very typical - Not at all typical). These data have been used to estimate a model to explain how much children watch TV.

The time-use diary was administered to children on randomly selected days, one week day and one weekend day.

Let’s initially look at the raw data. Table 5 displays the mean TV-time in seconds per day by day of the week and the response to the question mentioned above. It also includes the corresponding standard deviations and sample sizes. On average American children used 2.4 hours per day in front of the TV, somewhat more on Fridays and Saturdays than in the rest of the week. According to their own evaluation they use less time for TV watching on typical than on atypical days.

In principle an atypical day could have either more or less TV-time than normal and in a sample one would expect both kinds of event to be represented. For this reason we should expect that the standard deviation will increase with increasing score on the variable “typical”. The second display of Table 5 gives some support to this notion. The standard deviation is larger for days that are not typical than for typical days, but the standard deviation does not increase uniformly.

From the last panel of Table 5 we note that the sampling design results in relatively more observations on Saturdays and Sundays compared to each of the week days, and that about 17 per cent of the days recorded were not typical or not typical at all.

Table 6 displays the mean TV-time by a number of potentially explanatory variables. We find that children in the age bracket 10-13 have more TV-time, 2.5 hours, than both younger and older kids. On average the youngest use the least time in front of the TV, 2.25 hours. Children who live with a single parent have much more TV-time than children who live with a couple. There is also a negative correlation with the schooling level of the parents and the TV-time of the kids. The more schooling the parents have the less time their children use on TV. We also find a correlation between the quality of the neighbourhood and TV-time, the better quality the less time on TV. Finally we may note that American families typically have three or four TV sets. If they have less their children also have relatively less TV-time.

Figure 1 shows the distribution of TV-time by the indicator “typical”. All distributions have a high frequency of zeros, about 16 per cent of the observations, and they are strongly positively skewed. There are thus a few kids who spend a few very long days, more than 10 hours, in front of the TV. The strong skewness of these distributions made attempts to estimate tobit models and inflated negative binomial models less successful. Even if the estimated model parameters were plausible these models were not able to reproduce the observed distributions in a satisfactory way. It also turned out that the tobit model assumption

of having the same relation determine zero observations as well as nonzero observations was not fully realistic. Instead nonzero time-use observations were logged (while the zero observations were kept at zero), which resulted in nonzero distributions that were much closer to a symmetric distribution, see Figure 2. The following selection model was then estimated,

$$y^* = X\beta + \varepsilon_1; \quad (15a)$$

$$z^* = X\gamma + \varepsilon_2; \quad (15b)$$

$$\begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{Bmatrix} \square IDN\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma^2 & \rho^2\sigma \\ \rho^2\sigma & 1 \end{pmatrix}\right); \quad (15c)$$

$$y = y^* \text{ if } z^* > 0; \quad (15d)$$

$$y = 0 \text{ if } z^* \leq 0; \quad (15e)$$

where y is observed TV-time and X is a vector of explanatory variables. The model was estimated by maximum likelihood. Because most children have contributed two observations (one week day and one weekend day) these two observations are most likely correlated. The above specification is modified to allow for this correlation, while the assumption of independence between individuals is kept.

To avoid identification just by functional form exactly the same explanatory variables should not be used in the first two structural equations. At least one of the parameters in the β and γ vectors should be zero. The sets of variables displayed in Tables 7 and 8 were chosen after some testing. A few comments are in place. Because of the design – every child was asked to leave a diary for one week day and one weekend day – identification prohibits the use of dummy variables for every day. One week day dummy and one weekend dummy must be dropped. The reference days chosen were Monday and Sunday. The reference age group is 14-18, the reference schooling group is less than 12 years of schooling, the reference age group for the age of the head is less than 30 years, and the reference neighborhood group is excellent neighborhood.

The choice of explanatory variables should be almost self explanatory. One might expect that children watch more TV on Fridays, Saturdays and Sundays than on regular school days. Young children might watch TV more frequently but in shorter spells, children in the age group 10-13 are likely to watch more and probably also more than older kids who have more alternatives and can move more freely outside home. The schooling of the parents is included as an indicator of the ambition of parents for their children. More educated parents are expected to provide developing alternatives to TV watching to a greater extent than less educated, but schooling might also serve as an indicator of parents' economic means to provide alternatives. Neighborhood serves the same purpose. Families which live in excellent and very good neighborhoods probably have a good economy and perhaps also higher expectations for their children. Family income is also used to capture differences in economic means. The age of the head of the family and the dummy variable "if single" are both used to capture time constraints on the parents, how much time they can use for their children. Parents who are in the peak of their career around the age of 50 can be expected to have less time for their children than younger (and older) parents. Singles with children are also likely to suffer from time-constraints with little time left for their children.

The responses to the question about how typical the day was suggest that the respondents have a fairly good opinion about what is a typical or normal day. If we want to estimate a

model for “normal behavior”, responses for atypical days might not contribute much information and might even bias the estimates.⁵ A possible estimation approach is to down weight the contribution of atypical days and in the extreme to give them the weight zero and drop them from the sample. The problem is that we don’t know what weight we should give these observations to use the most of any information they might contribute. Tables 7 and 8 give estimates using alternative weighting schemes. The first column of Table 7 gives results with no weights (equal weights), the results in the second column were obtained when the contribution of each observation to the log likelihood function was weighted inversely proportional to the square root of the typical score, in the third column they were weighted inversely proportional to the score, and finally in Table 8 only observations with the scores 1 and 2 were used. The first panel of each table gives the estimates of the time-use equation and the second panel the estimates of the selection equation.

Comparing the estimates across weighting schemes shows that many estimates change systematically as the weight of atypical days become smaller. Some estimates change rather much; see for instance, the estimates of the effect of the head’s age on TV-time and the effect of the number of TV sets on the probability to look at TV, while other estimates only show small changes, for instance the estimates for the age of the child. We might also note that the standard errors increase a little when atypical days are down weighted. We can thus conclude that it matters how we handle atypical days. It is not only a matter of efficiency but there are also systematic differences in the estimates. This suggests that the process that leads to atypical days is different from the process that explains normal or typical time-use.⁶ If we are interested in estimating typical behavior a “safe” way is to drop the atypical observations. The reduction of the sample will, however, increase the standard errors. However, in this particular case the basic qualitative conclusions from the analysis will not differ very much depending on whether one uses the last column of Table 7 or the results from Table 8. In the following the reader is referred to Table 8 for an interpretation of the results.

In testing out the model it was found that some variables did not contribute much to the explanation of the selection or the time-use. They were then dropped from either equation to improve the identification. Age of the head was dropped from the selection equation, and family income from the time-use equation. The number of TV sets seemed to matter more for the choice of watching TV than for the number of hours, so a somewhat richer specification was used in the selection equation.

In the raw data we noted that American children watched more TV on Fridays and Saturdays. After standardizing for the influence of other variables the probability to watch on Fridays is still relatively high. There is a positive effect in the hours equation too, but it is insignificant. The probability to watch on Saturdays is however relatively low and the time-use effect is small too.

The age of the child is important both for the choice of watching and for the number of hours. The probability to watch is significantly higher for the two youngest age groups, while they spend much less time at the TV if they watch, 20-35 per cent less, compared to the teenagers. As we have already seen the joint effect is that the middle age group, 10-13, uses most hours

⁵ Please note that normal or typical behavior is not the same as average behavior. If one is interested in estimating a model that incorporates atypical days, the model should explain when (and why) a day becomes atypical, and use all the data unweighed.

⁶ When the variable “typical” is added as an explanatory variable it turns out that it has a significantly positive effect on the time allocated to TV-watching, while the effect on the probability to watch TV is insignificant.

in front of the TV. The effects of the age variable thus differ between the time-use and selection equations and this illustrates that a selection model is a more realistic model than a tobit model.

The probability to watch is significantly lower if the head of the family (the parents) have graduate training. Children whose parents have a university training use about 20 per cent less time at the TV if they watch compared to other children. The joint result is that TV watching decreases with the parents' increasing schooling. This should probably primarily be seen as the result of these parents' desires to give their children a richer and more meaningful activity than just watching TV. The alternative interpretation is that parents with many years of schooling on average have better means to offer their children alternative activities, but income does not seem to influence TV-time greatly. There is no significant effect in the time-use equation and the effect on the probability to watch is nonlinear. The probability first decreases with increasing income and then increases for the very rich.

The age of the head and marital status has no major effect on the probability to watch, while they both determine how much children watch if they have chosen to watch at all. Children of older heads watch more TV than those of younger heads, in particular children of heads in the age bracket 50-59 watch the most. One interpretation is that these parents are in the peak of their working career and have relatively little time for their children and thus leave them to watch TV. Similarly, children who live with single parents use more time in front of the TV than other children, probably because a single parent has less time to spend with his or her children. There is nobody who can share market work and household work.

There is no strong relation between the probability to watch and the neighborhood, while the time allocated to TV watching, if watching at all, is smallest in excellent and rather poor areas. Children who live in good and very good areas (the middle class?) use more time at the TV. (explanation??)

If the household has two or more TV sets the probability to watch increases compared to having just one TV as expected.

Finally we note that the residual correlation between the two equations is strongly negative. An interpretation is that children with frequent TV sessions do not watch for long, while children who watch more rarely tend to stay watching for a longer time.

Summary of findings

In this particular case it turns out that the response to the question "how typical was this day" was not primarily a quality indicator, signaling a deviant measurement or large measurement error, but primarily an indicator of a systematically deviant behavior. This suggests that atypical days do not contribute much information about typical or normal behavior and that atypical observations should be given zero or a very small weight.

The decision to watch TV at all is different from the decision about the number of hours. A tobit model is thus not a good vehicle to analyze these data.

In summary we have found that young children watch TV more frequently but in shorter spells while older children watch less frequently but in longer spells if they watch. The joint effect is that the total number of hours is the highest for children in the ages 10-13. Parents

with a university education stimulate their children to other activities than TV watching, while parents who work long hours and/or have much household work (as indicated by the head's age, if there is a high family income and if the head is single), allow their children to look at TV relatively long hours.

4. Conclusions

Auxiliary data can improve on the estimates from noisy time-use data. Depending on the nature of the auxiliary data and the desired analysis there are a number of ways these data can be used jointly.

The second numerical example, in particular, suggests that deviations from normal time-use is not a purely random phenomenon, but may shift the whole distribution – in this example towards more TV-time. This implies that a simple variance reducing strategy might not be adequate. Instead it might be necessary to model why and when some observations deviate from normal. Just to fit a simple model structure such as a tobit model to all data will in general lead to inconsistent and biased estimates of normal time-use behavior. To drop all atypical observations from the sample will reduce the risk for bias but increase the uncertainty of the estimates. Depending on the nature of the auxiliary data and the economic problem at hand one will have to choose a good strategy to use all information efficiently. There is no standard recipe.

References

- DeVill, J-C and C-E Särndal (1992) "Calibration estimators in survey sampling", *Journal of American Statistical Association*, Vol. 87, No. 418, 376-382
- Flood, L. N.A. Klevmarken and P. Olovsson (1997), *Household Market and Nonmarket Activities (HUS). Volumes II-VI*, Department of Economics, Uppsala University, Uppsala
- Huber, P.J. (1973) "Robust Regression: Asymptotics, conjectures, and Monte Carlo", *The Annals of Statistics*, I, 799-821
- Judge, G.G., W.E. Griffiths, R. Carter Hill and T-C Lee (1980) *Theory and Practice of Econometrics*, John Wiley & Sons
- Klevmarken, N.A. (1999), "Microeconomic Analysis of Time-use Data: Did We Reach the Promised Land?" in J. Merz and M. Ehling (eds.) *Time Use – Research, Data and Policy*, NOMOS Verlagsgesellschaft, Baben-Baden
- Klevmarken, N.A. and P. Olovsson (1993), *Household Market and Nonmarket Activities. Procedures and Codes 1984-1991*, Vol I and II, The Industrial Institute for Economic and Social Research, Almquist and Wiksell International, Stockholm
- Lundström, S. and C-E. Särndal (1999), "Calibration as a standard method for treatment of nonresponse", *Journal of Official Statistics*, 15,305-327

Table 1 Descriptive statistics of a data set from the Swedish Household Panel Survey, the 1993 wave.

Variable	Mean	Std Dev	Minimum	Maximum
Annual work hours (survey)	1658.6	796.4	0.0000	3715.2000
Annual work hours (time-use)	1449.5	1121.8	0.0000	5229.1667
Net montly wage rate	9.1956	6.3463	0.0000	63.9586
Virtual income	29.5746	27.3778	0.0000	167.4455
If female	0.5058	0.5003	0.0000	1.0000
Age	43.8623	10.9263	19.0000	64.0000
If single	0.1435	0.3508	0.0000	1.0000
Number of children	0.8831	1.0942	0.0000	7.0000
Years of schooling	11.9560	3.2948	6.0000	25.0000
If bad health	0.3229	0.4679	0.0000	1.0000
If weekday inter-view in the summer	0.1458	0.3531	0.0000	1.0000
Capital income	-14.4250	33.0259	-120.1000	264.5000
Monthly wage rate	15.0023	10.2094	0.0000	93.2341

Note: All amounts are given in 1000 Swedish Crowns. The number valid observations is 864.

Table 2 Single equation Tobit estimates of labor supply functions using survey and time-use data.

<i>Variable</i>	<i>Specification 1</i>		<i>Specification 2</i>	
	<i>Parameter estimates</i>	<i>Standard errors</i>	<i>Parameter estimates</i>	<i>Standard errors</i>
<u>Survey data function</u>				
Intercept	2456.13*	65.72	-1088.47*	389.68
Net monthly wage rate	-18.59*	5.15	-24.22*	5.10
Virtual income	-13.06*	1.11	-12.18*	1.05
If female	-578.14*	55.99	-530.80*	57.20
Age			150.90*	19.04
Age squared			-1.70*	0.23
If single			-236.41*	79.69
No of children			-107.04*	27.90
Years of schooling			45.89*	8.23
If bad health			-92.35	61.05
Residual Std.	779.92*	21.39	725.05*	19.85
No of left censored cases	102		102	
<u>Time-use function</u>				
Intercept	-3250.16*	771.92		
Net monthly wage rate	8.29	8.36		
Virtual income	-9.07*	2.02		
If female	-370.35*	110.30		
Age	219.80*	37.53		
Age squared	-2.40*	0.44		
If single	-109.64	154.64		
No of children	-74.27	53.70		
Years of schooling	31.31*	15.83		
If bad health	-195.79	118.26		
If summer week day	-924.10*	147.98		
Residual Std.	1367.19*	43.01		
No of left censored cases	231		231	

Note: Total number of cases is 820

Table 3 Bivariat Tobit estimates of labor supply functions using survey and time-use data.

<i>Variable</i>	<i>Unconstrained estimates</i>		<i>Constrained estimates</i>	
	<i>Parameter estimates</i>	<i>Standard errors</i>	<i>Parameter estimates</i>	<i>Standard errors</i>
<u>Survey data function</u>				
Intercept	2609.79*	88.89	2515.92*	54.50
Net monthly wage rate	-34.11*	7.47	-26.63*	4.21
Virtual income	-9.50*	1.39	-8.70*	1.26
If female	-483.12*	63.54	-471.09*	61.69
Residual Std.	573.80		573.41	
<u>Time-use function</u>				
Intercept	-837.50	1133.09	-607.66	1162.21
Net monthly wage rate	6.30	13.71	-26.63*	4.21
Virtual income	-2.96	3.11	-8.70*	1.26
If female	-265.44	154.44	-325.27*	156.07
Age	121.07*	54.57	121.09*	56.02
Age squared	-133.30*	64.06	-129.16*	65.72
If single	199.63	208.15	226.20	213.85
No of children	10.86	74.80	16.29	77.39
Years of schooling	-2.51	22.94	10.71	22.75
If bad health	-138.41	164.64	-137.39	169.24
If summer week day	-887.94*	201.36	-910.57*	206.85
Residual Std.	1230.38		1286.79	
Rho	0.207		0.188	

Table 4 Bivariat Tobit estimates of labor supply functions using survey and time-use data. Alternative specification

<i>Variable</i>	<i>Unconstrained estimates</i>		<i>Constrained estimates</i>	
	<i>Parameter estimates</i>	<i>Standard errors</i>	<i>Parameter estimates</i>	<i>Standard errors</i>
<u>Survey data function</u>				
Intercept	393.84	468.01	325.22	467.23
Net monthly wage rate	-41.31*	7.68	-32.10*	6.60
Virtual income	-9.52*	1.34	-8.70*	1.29
If female	442.43*	65.73	-428.13*	65.48
Age	88.81*	22.75	89.02*	22.73
Age squared	-97.87*	26.80	-98.57*	26.78
If single	-84.84	89.69	-84.92	89.62
No of children	-67.42*	32.51	-67.63*	32.43
Years of schooling	38.64*	9.55	35.68*	9.46
If bad health	-124.96	70.44	-126.02	70.38
Residual Std.	541.77		541.79	
<u>Time-use function</u>				
Intercept	-1845.65	1118.48	-1523.28	1145.04
Net monthly wage rate	4.46	13.60	-32.10*	6.60
Virtual income	-2.91	3.10	-8.70*	1.29
If female	-245.16	154.48	-308.03	157.01
Age	161.09*	54.36	158.99*	55.81
Age squared	-177.41*	64.01	-170.95*	65.64
If single	161.95	211.83	193.61	217.87
No of children	-19.57	76.27	-11.71	77.80
Years of schooling	14.52	22.28	26.46	22.72
If bad health	-194.85	166.86	-189.90	171.69
If summer week day	-888.63*	201.40	-911.59*	207.64
Residual Std.	1254.48		1290.29	
Rho	0.195		0.183	

Table 5 TV-time (mean, std and number of observations) by day of the week and indicator of how typical the day was.

Day of week	Typical					All
	1	2	3	4	5	
Mean time-use						
Sunday	8011	8376	8738	8964	8877	8382
Monday	7712	8835	8999	8550	9349	8455
Tuesday	8533	8279	9106	8777	7824	8535
Wednesday	7685	8167	8289	9794	10798	8336
Thursday	7492	8865	8456	9021	8399	8206
Friday	9907	8909	8873	9308	9723	9399
Saturday	8596	8804	8810	9150	9530	8817
All	8277	8579	8787	9074	9239	8595
Standard deviation						
Sunday	7446	7069	7473	7691	8186	7437
Monday	7555	7368	7339	8498	8603	7646
Tuesday	7797	6953	7840	7242	7541	7512
Wednesday	7188	6941	7038	8915	9588	7490
Thursday	7666	7688	7587	8011	7178	7658
Friday	7903	6815	7673	7219	8675	7623
Saturday	7941	7229	7533	8218	8561	7767
All	7672	7139	7513	7962	8376	7598
Number of observations						
Sunday	969	608	439	207	173	2396
Monday	373	239	212	95	69	988
Tuesday	432	288	175	88	66	1049
Wednesday	370	272	181	81	72	976
Thursday	378	211	176	88	80	933
Friday	376	225	189	95	84	969
Saturday	944	621	491	238	197	2491
All	1931	1237	938	447	371	4924

Table 6. Mean TV-time by age group, marital status, schooling and number of TV-sets

Variable	Mean	Number of observations
Age of child		
5-9	8132	1675
10-13	9001	1534
14-19	8682	1715
Marital status of parents		
Not single	8163	3387
If single	9546	1537
Schooling of head		
Less than 12 years	9208	944
12 years	9126	1552
13 – 15 years	8048	1123
16 years	7896	635
More than 16 years	6270	345
Schooling of wife		
Less than 12 years	9061	539
12 years	8384	966
13 – 15 years	7810	889
16 years	7484	543
More than 16 years	7029	231
Quality of neighbourhood		
Excellent	7830	1513
Very good	8465	1547
Good	9086	1036
Fair or poor	9646	808
No of TV-sets at home		
1	7763	289
2	8460	996
3	8677	1557
4	8806	1248
5	8749	458
6	8945	205
7	8503	125

Table 7. Estimates of selection models using alternative weights

Explanatory variables	Slope estimates			Std. of slope estimates		
	No weights	Weights = 1/sqrt(typical)	Weights= 1/typical	No weights	Weights = 1/sqrt(typical)	Weights= 1/typical
Log(TVtime)						
Tuesday	-0.038	-0.024	-0.009	0.046	0.046	0.049
Wednesday	-0.031	-0.060	-0.086	0.046	0.047	0.051
Thursday	-0.048	-0.044	-0.038	0.048	0.050	0.054
Friday	0.031	0.043	0.056	0.046	0.047	0.051
Saturday	0.024	0.031	0.040	0.029	0.030	0.032
Age of child						
5-9	-0.307*	-0.307*	-0.304*	0.038	0.039	0.042
10-13	-0.176*	-0.172*	-0.167*	0.036	0.037	0.039
Schooling of head						
12	0.024	0.029	0.037	0.040	0.041	0.044
13-15	-0.031	-0.043	-0.052	0.045	0.046	0.049
16	-0.113*	-0.136*	-0.155*	0.052	0.053	0.057
17-	-0.060	-0.087	-0.107	0.068	0.072	0.080
Age of head						
30-49	0.065	0.086*	0.109*	0.036	0.038	0.042
50-59	0.173*	0.183*	0.195*	0.051	0.053	0.058
60-68	0.032	0.110	0.179*	0.089	0.087	0.089
Marital status						
If single	0.101*	0.103*	0.104*	0.027	0.028	0.030
Neighbourhood						
Very good	0.090*	0.094*	0.094*	0.037	0.038	0.041
Good	0.049	0.055	0.058	0.041	0.042	0.045
Fair/poor	0.001	0.013	0.026	0.047	0.047	0.050
At least 2 TVs	0.072	0.065	0.061	0.066	0.068	0.072
Constant	9.147	9.113	9.072	0.096	0.099	0.107

Explanatory variables	Slope estimates			Std. of slope estimates		
	No weights	Weights = 1/sqrt(typical)	Weights= 1/typical	No weights	Weights = 1/sqrt(typical)	Weights= 1/typical
If viewing TV						
Tuesday	0.107	0.124	0.139	0.071	0.074	0.080
Wednesday	0.042	0.072	0.102	0.070	0.073	0.079
Thursday	0.058	0.058	0.055	0.072	0.076	0.083
Friday	0.100	0.127	0.153	0.070	0.074	0.082
Saturday	0.037	0.046	0.055	0.045	0.047	0.052
Age of child						
5-9	0.326*	0.323*	0.316*	0.054	0.057	0.062
10-13	0.308*	0.313*	0.314*	0.056	0.059	0.064
Schooling of head						
12	0.005	-0.016	-0.034	0.063	0.066	0.072
13-15	-0.079	-0.094	-0.111	0.068	0.072	0.079
16	0.074	0.084	0.092	0.085	0.089	0.098
17-	-0.271*	-0.286*	-0.301*	0.105	0.110	0.120
Neighbourhood						
Very good	-0.058	-0.072	-0.082	0.056	0.059	0.064
Good	0.055	0.052	0.047	0.064	0.067	0.073
Fair/poor	0.158*	0.132	0.105	0.075	0.078	0.085
Number of TV sets						
2	0.096	0.134	0.165	0.104	0.107	0.116
>2	0.153	0.182	0.202	0.100	0.102	0.109
Family income						
Income	-1.27e-06*	-1.65e-06*	-2.05e-06*	5.18e-07	5.74e-07	6.51e-07
Income_sq	1.27e-12*	1.61e-12*	1.98e-12*	6.49e-13	7.52e-13	8.61e-13
Constant	0.712	0.714	0.727	0.134	0.138	0.148
Rho	-0.934	-0.925	-0.914	0.009	0.011	0.014
Sigma	0.923	0.916	0.908	0.013	0.014	0.015
lambda	-0.862	-0.847	-0.831	0.018	0.020	0.023

Note: Total number of observations is 4535, of which 674 were censored. The model was estimated by maximum likelihood and the standard errors were adjusted for any correlation between observations on the same individual.

Table 8. Estimates of a selection model using a sample reduced to “typical days”

Explanatory variables	Slope estimates	Std. of slope estimates
Log(TVtime)		
Tuesday	-0.031	0.053
Wednesday	-0.117*	0.054
Thursday	-0.047	0.058
Friday	0.037	0.056
Saturday	0.048	0.035
Age of child		
5-9	-0.307*	0.045
10-13	-0.188*	0.043
Schooling of head		
12	0.015	0.047
13-15	-0.097	0.053
16	-0.179*	0.061
17-	-0.201*	0.086
Age of head		
30-49	0.124*	0.046
50-59	0.207*	0.067
60-68	0.185	0.098
Marital status		
If single	0.107*	0.033
Neighbourhood		
Very good	0.128*	0.044
Good	0.072	0.049
Fair/poor	0.005	0.055
At least 2 TVs	0.056	0.082
Constant	9.095	0.116

Explanatory variables	Slope estimates	Std. of slope estimates
If viewing TV		
Tuesday	0.172*	0.088
Wednesday	0.151	0.086
Thursday	0.107	0.092
Friday	0.198*	0.093
Saturday	0.050	0.057
Age of child		
5-9	0.336*	0.069
10-13	0.330*	0.072
Schooling of head		
12	-0.040	0.079
13-15	-0.059	0.090
16	0.085	0.107
17-	-0.293*	0.130
Neighbourhood		
Very good	-0.131	0.069
Good	0.076	0.083
Fair/poor	0.136	0.095
Number of TV sets		
2	0.163	0.130
>2	0.242*	0.123
Family income		
Income	-2.46e-06*	7.65e-07
Income sq	2.61e-12*	1.10e-12
Constant	0.675	0.165
Rho	-0.917	0.014
Sigma	0.901	0.016
lambda	-0.826	0.024

Note: Total number of observations is 2895, of which 412 were censored. The model was estimated by maximum likelihood and the standard errors were adjusted for any correlation between observations on the same individual.

Figure 1. TV-time by the indicator “typical” (seconds)

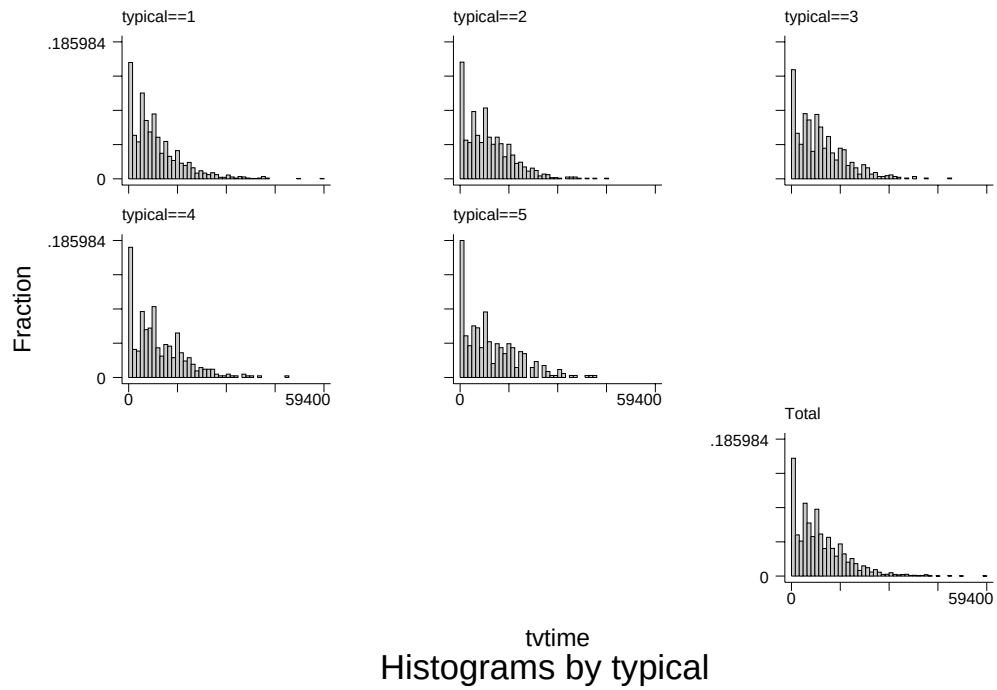


Figure 2. Log TV-time by the indicator “typical”

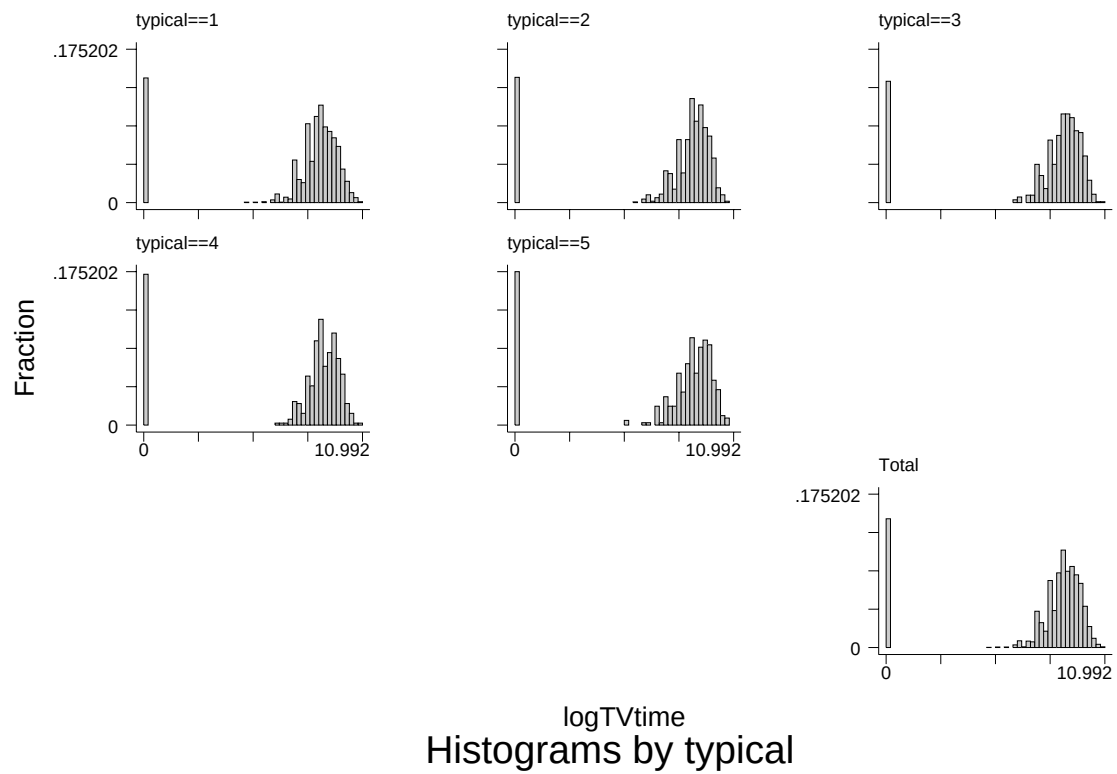


Figure 3. Simulated log TV-time distributions by the variable “typical” from a selection model estimated without weights.

