CZECH HOUSEHOLD PANEL SURVEY

Time-Use Diary Documentation

Waves 1 (2015) to 4 (2018)

Michaela Kudrnáčová (translation Jan Morávek)

Institute of Sociology, Czech Academy of Sciences

www.promenyceskespolecnosti.cz

promeny.spolecnosti@soc.cas.cz

English version 1.0 (10 October 2020), based on Czech version 4.0 (21 November 2019)

PROMĚNY ČESKÉ SPOLEČNOSTI

1. Introduction

This documentation relates to the Czech Household Panel Survey (CHPS) datasets from child and adult time-use diaries.

Information about the panel survey, the sample, data collection, processing and weighting methods are available in the main dataset's documentation.

2. Time-use diary and its completion

The CHPS time-use diary takes the form of a table with activity categories on the vertical axis and timeslots on the horizontal axis. Different versions of the diary were prepared for adult respondents and child respondents aged 10–17 years, with different numbers and contents of the activity categories prompted. The data was collected by means of a paper-and-pencil self-administered questionnaire, and a CAWI mutation of the instruments was prepared for the purposes of online interviewing.

The diary consists of 48 slots of 30 minutes. The first slot begins at 6 AM on the previous day, and the last slot ends at 6 AM on the present day. Thus, the diary was completed in retrospect, for the past day. The respondent drew a horizontal line to mark activities he/she had performed at the times specified. The entire length of the timeslot was to be marked even if the activity had started or ended at another time than the boundary moment (e.g., at 4:15 PM). Timeslots that the respondents only marked in part (e.g., by drawing a line from 4:15 to 4:30 PM) were entered in the same way as timeslots marked throughout.

The respondents were allowed to mark several simultaneous activities in a timeslot, without distinguishing between primary and secondary activities. Several activities in the same slot also appeared when a respondent had performed several consecutive activities that had not lasted for the slot's entire duration. Simultaneous and consecutive activities cannot be told apart in the data.

Adult respondents completed the diaries independently, whereas children aged 10–17 could be assisted by adult household members. Instructions for completing the diary were given on the first page. During in-household interviewing, the instructions were given by the interviewer. A dedicated section of the mandatory interviewer training dealt with diary completion. The interviewer also verified the completeness and accuracy of the diaries completed, asking the respondents for any additional information or corrections necessary.

The basic parameters of the diaries were the same across all waves of the survey. However, for Wave 2, the question about diary completion date was modified, the completion instructions were expanded, and more precise wording of activities 13, 14 and 16 was applied in the adult diary. In Wave 3, the diary completion date question was modified again, and that change was kept for Wave 4. The diary completion date question was modified in line with fieldwork experience to prevent respondents from filling in the previous day's date instead of the present day's date.

3. Dataset content and structure

Diary data are provided in two separate files for adults and children, respectively, because children were given different answer options than adults.

Table 1: Variables in the diary datasets

w(*x*)_hid	unique household ID in Wave (*x*)
w(*x*)_pid	unique personal ID in Wave (*x*)
pno	person's order in the household
diarymon	diary completion month
diaryday	diary completion day
orday	day described was ordinary or extraordinary
daytype	type of day described (workday, weekend etc.)
tx_yy_yy_zz_zz	activity in the timeslot
impusle	imputed sleep activity

The datasets include ID variables *w(*x*)_hid*, *w(*x*)_pid* and *pno* for linking to the given wave's main dataset.

The *diarymon* and *diaryday* variables indicate when the diary was completed. The respondents were to fill in the current date, not the date of the previous day that the diary accounted for.

In addition to the completion date question, the first page of the diary also contained two questions about characteristics of the day accounted for (*orday*, *daytype*).

Due to anonymization, the *ordayun* variable (a text specifying why the respondent considered the previous day extraordinary) was excluded from the dataset.

The variables *t1_06_00_06_30* to *t15_05_30_06_00* contain information about activities performed in the different timeslots. In the t*x_yy_yy_zz_zz* format, *x* refers to the number of the activity in the slot, *yy_yy* indicates the slot's start time and *zz_zz* its end time. During data entry, activities marked in the same timeslot were entered in the slot's variable in ascending order of their numerical codes. In timeslots with *x* activities marked, the variables *tx+1* to *t15* equal 96 (no other activity). The number of variables per timeslot (15) was derived from the maximum number of activities Wave 1 respondents marked in a single slot.

The *impusle* variable indicates whether Activity 1 (sleep) was imputed for the diary at hand. Sleep imputation was done as part of manual data entry in Waves 1 and 2. Sleep was imputed in empty timeslots preceding the first marked timeslot between 6 AM and 10 AM, and in empty timeslots following the last marked timeslot between 8 PM and 6 AM, unless the sleep activity was marked in any of those slots. In Wave 2, imputation was only required to a very limited extent because of increased quality of respondent completion. In Waves 3 and 4, imputation in line with the above rules was performed during data cleaning.

The weight variables for diary respondents are included in the given wave's main dataset. Weighting was performed using the *rake* function in R (*survey* package) on the basis of sex, age, education and region of residence for adult respondents, and on the basis of sex and region of residence for child respondents. The basic set of weights (*W*) takes the above sociodemographic variables into account, while the extended set of weights (*Wd*) not only considers sociodemographics but also ensures a uniform distribution of days of the week. The main dataset's documentation contains more details on weighting.

4. Dataset versions

By 25 April 2018, the time-use diary datasets from CHPS Wave 1 had been published in two versions.

Table 2: Wave 1 diary dataset versions

	publication date	changes made
v1.0	18 Jul 2017	-
v1.1	25 Apr 2018	ID variables *hid* and *pid* renamed to *w1_hid* and *w1_pid*

The time-use diary datasets from CHPS Wave 2 were published in a single version on 25 April 2018.

Table 3: Wave 2 diary dataset versions

	publication date	changes made
v1.0	25 Apr 2018	-

The time-use diary datasets from CHPS Wave 4 were published in a single version on 31 January 2019.

Table 4: Wave 3 diary dataset versions

	publication date	changes made
v1.0	31 Jan 2019	-

The time-use diary datasets from CHPS Wave 4 were published in a single version on 21 November 2019.

Table 5: Wave 4 diary dataset versions

	publication date	changes made
v1.0	21 Nov 2019	-