BIBB-FDZ
Data and Methodological Reports

No. 4 / 2013

BIBB Supplemental Task Survey
to the Employment Survey 2012

Version 2.1

Holger Alda
Daniela Rohrbach-Schmidt
Michael Tiemann

BIBB Supplemental Task Survey to the Employment Survey 2012,
authors: Holger Alda, Daniela Rohrbach-Schmidt, Michael Tiemann

Data availability

The data set described in this manual can be requested as a Scientific Use File via our website www.bibb-fdz.de.

Note on version numbering

Changes from the previous version without major thematic relevance are documented by sequentially increasing the number *after* the dot (second level). Changes of thematic relevance, on the other hand, lead to a sequential increase of the number at the first level.

Table of contents

Appendix

1   Motivation behind the Supplemental Task Survey

The research data described in this data and methodological report measure possible occupational changes faced by employed persons, provide supplemental information to the main section of the Employment Survey and measure the share of time that the participants spend performing individual occupational tasks on any working day of their choice. The motivation for conducting the Supplemental Task Survey to the Employment Survey 2012 is originally attributable to discussions regarding measurement problems conducted during the two-day T.A.S.K.S. 2 workshop. T.A.S.K.S. is a long-term series of events jointly organised by the Institute for Employment Research (IAB) and the Federal Institute for Vocational Education and Training (BIBB). At the core of the discussions which led to the Supplemental Task Survey was the actual share of time that employees spend on the individual occupational tasks. The main section of the Employment Survey 2012 measures information on the frequency with which individual occupational tasks are carried out on an ordinal scale (three levels: often, sometimes, never).

Because the Supplemental Task Survey was conducted approximately one year after the main section of the Employment Survey 2012, the research data also contain information on interim (occupational) mobility, as the employed persons may have changed their work schedule, job, company of employment and/or their occupational status (self-employed status vs. salaried employment) since the first survey. In addition to collecting data of actual interest regarding the share of time spent on occupational tasks, the research data described in this data and methodological report consist of a CATI section, which compiles information regarding (non-)mobility since the main survey (interviews conducted between November 2011 and March 2012) until this survey period (October/November 2012), as well as 25 items regarding "generic tasks" that were not a part of the main survey. At the end of the CATI interview, the survey participants were requested to take part in the survey regarding the share of time spent on the individual occupational tasks on any reporting day of their choice or to asked if they would consent to providing this information. The field phase for the second part of the survey ended on 15 December 2012 and in January 2013 for a handful of stragglers.

The next chapter provides a brief overview of the key characteristics of the data set. The third chapter describes the observation design of the study, from the selection of the observation units all the way to possible selection effects of the survey participants' actual response behaviour. It also contains a brief explanation of how the projection factors for the Supplemental Task Survey were determined. The forth chapter describes the contents and special features of the two sub-surveys. Chapter five describes the organisation of the research data and unique aspects of its handling. Chapter six briefly addresses data access and individual data products.

2   An overview of the research data from the Supplemental Task Survey

Survey title	**Supplemental Task Survey 2012 to the main section the Employment Survey 2012**
DOI	doi:10.7803/610.12.1.1.20
brief description	The Supplemental Task Survey measures occupational mobility one year after the main interview for a subsample of the main section of the Employment Survey 2012 and provides information on "generic tasks" for approximately 4,300 persons. Nearly 2,300 persons also took part in a second survey section in which they reported on the share of time they spend on individual occupational tasks and the level of difficulty of these activities on a reporting day of their choice.
observation year	2012 (4th quarter)
observation unit	Persons who took part in the Employment Survey 2012 and who were willing to take part in additional surveys, without persons with an advanced further training certificate (master craftsman, technician, etc.)
key topic	5 (employment)
data access options	SUF, GWA, DFV
number of variables	Part I: 104 (SUF); 5 (VT); 1 (SV) | Part II: 73 (SUF); 1 (VT)
population	Employed persons in Germany who work at least ten hours a week, with the exception of trainees and employed persons with an advanced training certificate (master craftsman, technician, etc.)
weighting / projection	Adaptation of the projection factors from the Employment Survey 2012 to participation cases
representative regional level	East/west
number of cases	Part I: 4,356 ; Part II: 2,272
survey method	Part I: CATI ; Part II: online/written
selection procedure	Layered random sampling (layers: east/west; man/woman; without/with vocational training as well as a degree from a university or university of applied sciences
observation design	Cross-section
note	The full evaluation potential of these research data can usually be exploited only in conjunction with the main section of the BIBB/BAuA Employment Survey 2012.
links	BIBB-FDZ metadata portal: http://metadaten.bibb.de/metadaten/55 Project page: http://www.bibb.de/de/62624.htm
key words	Employment, occupational tasks, generic tasks, occupational mobility, share of time

3 Observation design and selection of the observation units

The supplemental survey "Tasks" to the Employment Survey 2012 has certain connections with other supplemental surveys. This is important in regard to the research data documented here among other things because it required an internal BIBB organisation process for a total of three supplemental surveys. The primary purpose of this process was to prevent negative impacts on the willingness of the participants who took part in the main survey to participate due to repeated strain. This resulted in effects on the research design and realisable sample.

3.1 Observation design

The field phase for the first of a total of three supplemental surveys carried out by BIBB in 2012 took place in July/August of 2012. The goal was to examine differences and similarities in the occupational contents of commercial tasks. All persons with the following characteristics were selected for this first supplemental survey

a) participated in the main section of the Employment Survey 2012,
b) signalled fundamental agreement to supplemental surveys (N ≈ 14,700) and
c) were active in one of ten groups of commercial occupations selected by the project team for content-related reasons (see Chap. 4.4 final report "Similarities and differences in commercial occupations"– *previously unpublished manuscript*).

A CATI interview was conducted with these persons. At the end of the interview, they were asked about their willingness to take part in the Supplemental Task Survey. If they expressed their willingness, they were asked for their street address so that the task questionnaire presented later in this data and methodological report could be sent to them.

For organisational reasons, it was impossible to avoid conducting the Supplemental Task Survey and an additional supplemental survey on upgrading training at the same time. In extreme cases, this could have led to participants in the main survey being contacted in regard to three supplemental surveys in a period of approximately five months. We therefore decided to exclude persons with upgrading training from the sampling basis for the Supplemental Task Survey[1].

---

[1] More detailed information on the sampling basis for the Supplemental Task Survey as well as on the impacts of the special features discussed above can be found in the next section.

The Supplemental Task Survey was designed ex-ante as a mix of observation designs. While questions regarding occupational mobility and the associated changes in individual task matrices are easy to ask within the scope of a CATI interview, this survey method is not especially well suited to gathering information on the share of time spent on individual occupational tasks. Instead of the originally planned written questionnaire regarding share of time spent on tasks (see Chapter 4), an online survey was conducted in consultation with the survey institute. As mentioned above, the street addresses for a written survey were already available from some of the potential participants from the first, earlier supplemental survey. The questions regarding the share of time spent on activities were therefore asked online for participants who do not perform commercial tasks in the sense indicated above. The resulting distribution of the participants is outlined in the next section.

3.2   Selection of the observation units

The sample design of the Supplemental Task Survey is a random sample quoted according to the percentages in the main survey and layered according to personal characteristics. As a reference value oriented to the available resources, the goal was 3,000 observation units who indicated their fundamental willingness to take part in the second written or online section of the Supplemental Task Survey following the CATI interview. It was assumed here that approximately 25 per cent would not be interested in taking part in a second survey section and that non-responses were possible at any time due to interim unemployment.

Just like the main Employment Survey, the sample for the Supplemental Task Survey was to be representative. In order to achieve this within the scope of the field control for the CATI section without major control effort, a total of 16 layering cells were formed (east/west, men/women, without/with vocational training/master craftsman, technician, etc./university graduates) that were then reduced to 12 cells through exclusion of all employed persons with upgrading training. The observation units are distributed across all 16 layering cells in the Employment Survey as follows:

Table 1:   Distribution of employed persons from the main survey according to the layering characteristics of the Supplemental Task Survey (figures in per cent)

		Vocational training		Trade and technical school	University/ University of applied sciences
		without	with		
East	**Men**	0.47	5.21	0.84	3.06
	Women	0.52	6.73	0.57	4.17
West	**Men**	2.53	20.19	4.83	10.57
	Women	3.80	25.59	2.05	8.86

Source: BIBB/BAuA Employment Survey 2012 (N = 100% = 20,026)

Because the observation units in the column "trade and technical school" have been excluded from the Supplemental Task Survey, the percentages for the main Employment Survey in the following Table 2 refer to population shares when this group of persons is not taken into account (18,371 cases then remain in the main survey).

Table 2: Comparison of the composition of individual layering cells in the Supplemental Task Survey and the main Employment Survey 2012 (figures in per cent)

			Vocational training		University/ University of applied sciences	Total
			without	with		
	East	Men	0.53	5.79	3.31	**9.62**
Supplemental		Women	0.51	7.44	4.50	**12.44**
Task	West	Men	2.78	21.95	11.64	**36.36**
Survey		Women	4.22	27.71	9.64	**41.57**
		Total	**8.03**	**62.88**	**29.09**	**100**
	East	Men	0.51	5.68	3.34	**9.53**
Main		Women	0.57	7.34	4.55	**12.46**
Survey	West	Men	2.76	22.01	11.53	**36.30**
		Women	4.14	27.90	9.66	**41.70**
		Total	**7.99**	**62.94**	**29.07**	**100**

Source: BIBB/BAuA Employment Survey 2012 (N = 100% = 18,371) as well as CATI section Supplemental Task Survey (N = 100% = 4,356)

At a maximum of approximately 0.15 percentage points, the differences between individual cells in the upper and lower sections of Table 2 are negligible. Because the gross sample for the CATI section was compiled through simple random selection, the two samples cannot be expected to differ in regard to other characteristics and characteristics not listed in Table 2. However, this is not the case in regard to the exclusion of employed persons who have completed upgrading training at a trade and technical school during the course of their career history. Based on qualification requirements, such persons – working as master craftsman, technicians, etc. – are likely to have more demanding occupational profiles within their current occupational task than persons, for example, who have completed a vocational training course. The Supplemental Task Survey will therefore underestimate the difficulty level of tasks in Germany to the extent that this level is higher (less probable: lower) for graduates of a trade and technical school than for all employed persons in Germany on average.

3.3 Weighting

In anticipation of the failure analysis in 4.2, there are only slight and in one such case selective failures, so that we believe a conclusion in regard to the population, as described in

the previous section, is legitimate. The recalculation of the weighting factors is based on the assumption that the population is distributed as indicated in Table 2 (lower half). For each of the twelve cells, the projected marginal total is first calculated for all observation units in the Employment Survey 2012 (N = 18,371). When projected, the persons across all twelve layering cells stand for 33,191 million employed persons in Germany. The weighting marginal totals in each of the twelve layering cells are then calculated for all observation units participating in the Supplemental Task Survey. The weighting factor for each unit of the Supplemental Task Survey is the original weighting factor for these observation units plus the projected difference between the marginal total in the main and the supplemental survey, divided by the number of observation units in the Supplemental Task Survey (the calculations are performed per layering cell). This procedure is carried out analogously for the observation units who took part in the second section of the Supplemental Task Survey. Table 3 provides a comparison of the weighting factors.

Table 3:   Comparison of the weighting factors for the main Employment Survey 2012 and the two parts of the Supplemental Task Survey.

	weighting factor for the …		
	Main Employment Survey	Supplemental Task Survey, Part I	Supplemental Task Survey, Part II
N	18,371	4,356	2,272
Mean	1,808.07	7,619.67	14,608.82
Standard deviation	1,507.81	2,069.31	4,035.49
Variation coefficient	0.834	0.272	0.276
Median	1,332.92	7,165.30	12,794.38
10 percentile	545.37	5,616.17	11,088.89
25 percentile	755.26	6,004.84	11,647.07
75 percentile	2,253.22	8,665.04	17,602.86
90 percentile	3,791.95	10,337.91	20,120.76

Source: BIBB/BAuA Employment Survey 2012; Supplemental Task Survey 2012, Part I and II

When the marginal total remains constant, the weighting factors increase as the number of cases decreases. Because a constant value is added within each layering cell for each observation unit, the values are less dispersed than projection factors for the main survey. We therefore recommend using the projection factor calculated for the Supplemental Task Survey for descriptive evaluations and taking the case numbers on which the weighting is based into critical account when making differentiations between groups. The two weighting factors are designated **gew_task1** and **gew_task2** and can be found in the respective SUF data sub-sets I and II of the research data.

4 Contents and special features of the two data sub-sets

4.1 CATI data set[2]

On the one hand, the contents of the CATI section are designed as a screening for the second data sub-set of the Supplemental Task Survey and, on the other hand, as means of presenting additional issues not covered in the main survey. The first area (Section 4.1.1) measures any occupational mobility between the main and supplemental survey in order to ensure that the individual task lists sent to the observation units in the second part are sure to refer to current tasks as of November 2012. The second area (Section 4.1.2) provides additional information on empirical sub-aspects of the task base for the units from the main survey taking part in this survey.

4.1.1 Mobility section (variables F1 to F14 as well as F16)

Following the standardised welcome text, the telephone interview starts with questions in regard to any possible changes in job since the main survey[3]. The reasons for the change were queried individually, yielding the following empirical results:

Table 4: Mobility of employed persons from the main survey who took part in the CATI section of the Supplemental Task Survey (absolute numbers, in parentheses: row percentages)

Reason for change	Men	Women	Total
No change	*1,757 (87.7)*	*2,074 (88.1)*	*3,831 (88.0)*
Employer	74 (3.7)	97 (4.1)	171 (4.5)
Job	52 (2.6)	57 (2.4)	109 (2.5)
Unemployed/retired	99 (4.9)	112 (4.8)	211 (4.8)
Became self-employed/freelancer	12 (0.6)	11 (0.5)	23 (0.5)
Assumption of civil servant status / in apprenticeship training*	9 (0.5)	2 (0.1)	11 (0.3)
Total	2,003 (100)	2,353 (100)	4,356 (100)

* Queried separately in the CATI interview and summarised in one line in Table 3 due to the small number of observations
Note: Percentages have been rounded to one decimal place
Source: Supplemental Task Survey 2012

---

[2] The questionnaire for the CATI section of the Supplemental Task Survey can be found at http://www.bibb.de/de/64652.htm.
[3] The corresponding individual information from the main survey was provided to the individual survey participants prior to each question.

Persons who did not change (first result line in Table 3) were presented with question F9. With the exception of persons who were unemployed or had meanwhile retired or become full-time apprentices (these were guided to the end of the interview), the following information was gathered in the questions F2 to F8 and F16a for persons who changed jobs (N = 305):

- interim unemployment, supervisor function, occupation in current main employment and any change in branch of industry (F2 to F5code as well as F16)
- performing of tasks as in main survey (F303_z to F321_z)
- percentage of usability of knowledge from the old job in the new one, up to eleven reasons for the change, fulfilment of expectations in the new job in up to eleven aspects (F6 to F8xx).

Finally, all respondents (except unemployed / retired workers and full-time apprentices) were interviewed on changes in working time as well as changes in wages and salaries, and on their knowledge in regard to what wages and salaries are typically paid in their current job (F9 to F14).

The calculation of the differential typical for analysis of occupational mobility is formed by comparing the variables listed with the information from the main survey. The questions were asked using the same wording and with the same survey method as in the main survey, and are thus comparable.

4.1.2 Additional questions in the CATI section as compared to the main survey (Variables F15_1 to F16_22)

In addition to the information in 4.1.1, all survey participants were asked for detailed information on the tasks they perform in a total of 30 individual items. The spectrum ranged from the frequency of brief, repetitive activities for the employed persons all the way to how often they use higher mathematics in their area of occupational task. A detailed individual list can be found in the questionnaire for the Supplemental Task Survey.

4.2 Written and online section

Following the questions in the CATI section, the participants were asked to take part in the second part of the survey. The addresses of those persons who had already taken part in another supplemental survey (see Section 3.1) in August were verified. They were sent the questionnaire for entering the amount of time they spend on individual occupational tasks on any working day of their choice in the mail, including a stamped envelope for returning it. People who had not taken part in the first supplemental survey were first asked if they would be willing to provide the same information in an online questionnaire[4]. In this case, they were requested to provide an e-mail address. If they did not provide this information, these persons were asked for their street address so that the written questionnaire could be sent to

---

[4] The decision in favour of the online questionnaire as opposed to a written questionnaire was based on the assumption that participants would be more willing to provide their e-mail address than their street address. In addition, online questionnaires offer better means of control than written surveys. We assumed that also asking for an e-mail address from the persons who had already provided their street address in the first supplemental survey would only lead to irritation, however.

them in the mail. The (few) people who also refused this option then no longer had the option of taking part in the second section of the survey. Table 5 provides an overview of the willingness of the persons contacted to participate.

Table 5: Willingness to participate in the second part of the Supplemental Task Survey

	Number	in %
CATI in total *not including* addresses already on hand and persons unemployed at the time of the survey	2,977	100
Of which willing to participate via e-mail address	2,337	78.5
Respondents hesitant	233	7.8
Respondents do not have an e-mail address	362	12.2
Respondents do not want to participate	45	1.5
Respondents hesitant/do not have an e-mail address (N=595):		
Of which:		
Willing to participate by post	584	(98.2)
Refused to participate	11	(1.8)

Source: Paradata from the Supplemental Task Survey 2012

Only 56 of the nearly 3,000 respondents (1.9 per cent) expressed unwillingness to take part in the second part of the supplemental survey in advance. The actual return rates for the second part of the supplemental survey differ according to the survey method, so that a special failure analysis was performed here[5].

For the written section, it was sufficient to differentiate between people who returned the written questionnaire and those who did not. An additional category was formed for the online section, however, because respondents on the CATI section could log onto the online platform but then not enter any information there – for whatever reasons. Expressed in figures, that results in the following statistics for the second part of the Supplemental Task Survey:

Table 6: Actual participation in the second part of the Supplemental Task Survey

	Number	in %
CATI in total *not including* unemployed /retired persons, full-time apprentices	4,136[6]	100.0
Of which written response	1,067	25.8
Online response	1,224	29.6
Logged in online, but did not provide any (usable) information	249	6.0
No response	1,596	38.6

Source: Supplemental Task Survey 2012

---

[5] The contents and adjustment steps in the second part of the supplemental survey are described in the next chapter.

[6] Thus there were no respondents who initially refused to take part in the second part of the study.

The response rates were approximately 67 per cent for the written and approximately 51 per cent for the online section. The differences between the two survey methods are attributed less to the medium used than to the different nature of the respective groups of participants. Approximately two-thirds of the participants who received a written questionnaire were persons who had already participated in a survey in late summer. From among all of the potential participants, they therefore immediately represent a certain positive selection, while the same does not hold true of those completing the survey online.

It is possible that the responses by survey method and by personal or job characteristics different significantly among each other as well as in regard to persons not responding. We used a multinominal logit model (MNLM) with the following specifications for more detailed analysis.

Dependent variable (=y): Four categories (=n) for written response, online response, those who logged in online but did not provide any usable information and no response. The base category m are persons who did not provide us with a response to the second part of the Supplemental Task Survey.

Independent variables (=x): Exclusively dummies with – if necessary – the respective reference category in parentheses. Characteristics female yes/no, German citizen yes/no, Saturday work yes/no, Sunday work yes/no, qualification level (completed vocational training). Added to this are vectors for occupational status (employee), German federal state (North Rhine-Westphalia), age categories (35 to 44 years old), wage categories (EUR 1,500 to under EUR 3,000), working time categories (35 to 44 hours), company size (1 to 4 employees) as well as the BIBB occupational fields (commercial office occupations).

Tests: Likelihood ratio and Hausman test for merging the attributes of y, Hausman and Small/Hsiao test for review of the irrelevance of independent alternatives (IIA) assumption and a Wald and likelihood ratio test for all x vectors.

In presenting the results, we will limit ourselves to the significant effects. If needed, the corresponding Stata output with all original results can be sent formlessly via e-mail at any time. In regard to the tests for the n attributes of y, only the Wald test showed that the coefficients of x for the category pair "did not participate" and "logged in online but did not enter any information" (with the exception of the intercept) are not different from zero. The likelihood ratio test (LR test), which is more suitable according to the specialist literature (Long/Freese, 2006: 236 ff.), does *not* arrive at this result, however. According to this test, the four attributes formed for y are not redundant in relation to the x variables. Table 7 presents the significant x-coefficients, also confirmed by an LR test, for the x variables as marginal effects for an average sample observation unit[7].

---

[7] In order to provide a better overview, Table 7 shows only the number of occupational fields that deviate significantly from the commercial office occupations. In order to avoid long research times, the discrete changes were calculated with prchange.ado from Long/Freese (2006), which is why no standard errors are indicated. Information on the average effect of an x variable on y has not been provided Table 7.

Table 7: Selection effects as compared with the base category "did not participate in the second part of the Supplemental Task Survey" (information as percentage probability change for an average observation unit)

	Written	**Online**	**Online without entries**
German citizen (yes=1)	11.3	5.8	
Family worker (ref. employee)	20.7		1.7
German state (ref. NRW)			
Lower Saxony			- 0.2
Saarland		13.1	
Brandenburg:	-7.4	-5.9	
Saxony-Anhalt			-0.3
Age (ref. 35 to 44 years old)			
45 to 55	16.2		
56 and older	31.1		
Monthly wages/salary (ref. EUR 1,500 to EUR 2,999)			
EUR 401 to EUR 800	-9.2		
EUR 3,000 to EUR 4,999		7.3	
Over EUR 5,000		10.4	
Working time >= 45 h (ref. 35h to 44h)	-3.8		
Company size (ref. 1 to 4 employees)			
5 to 19 employees		- 8.3	
20 to 99 employees		- 8.3	
Not specified/missing/filter		- 10.0	
Number of significant occupational fields (ref. commercial office occupations)	8	6	3

reporting only coefficients with an error probability < .05
Source: BIBB/BAuA Employment Survey 2012, Supplemental Task Survey 2012

Persons who logged in online but did not make any entries are especially similar to non-participants in regard to their characteristics included in the regression equation[8]. Differences between written and online respondents are not necessarily attributable to effects of the survey method but in part describe the special features in the selection of persons for the written survey method (commercial tasks in the occupational profile). This can be assumed in particular when the respective effect is significant for only one of the two response groups.

---

[8] This was to be expected in keeping with the test results (Hausman vs. LR test).

A more systematic difference between the participants (no matter which type of survey) is therefore the personal characteristic "German citizen". The average effect for this characteristic shows a 11.3 (online) and 5.8 (postal) per cent higher probability of response. With the exception of the regional effects, the contents of which are not easy to interpret, all other effects are identifiable for only one of the two survey methods. Because a content-related interpretation of the significant effects is not usually conducted at a general level but in relation to specific, content-related issues, we have chosen not to pursue a content-related interpretation of the results of the failure analysis.

5   Time information on the reporting day (second part)

There is no difference between the two survey methods written and online in terms of content. As long as they were not unemployed at the time and had not refused to take part in the second section of the survey, participants in the CATI interview were each given an individual questionnaire listing all of the tasks for the individual and at the time point of the survey current occupational profile according to the corresponding items in the main survey (F303 to F320) or, in the case of a change, from the CATI section (F303_z to F320_z). Respondents could also enter up to three additional tasks of their choosing. The second section gathered the following information:

- On which day of the week the reporting took place
- Total working time on the reporting day
- For each individual task: number of hours/minutes, level of difficulty (subjective from 1: especially easy to 4: especially difficult) and frequency of the task (from 1: daily to 4: almost never).

A sample of the written survey for one observation unit can be found in Appendix A1, a sample (screenshot) of the online survey in Appendix A2[9].

5.1   Structure of the data set for the second part of the Supplemental Task Survey

The data set for the second part of the supplemental survey contains a total of 2,566 data lines and the service data set for 2,273[10] cases described in 5.2 and cleansed of data lines that could not be meaningfully evaluated. The responses in total are somewhat higher than this because in some cases (number in the low two-digit range) people sent a response when they were on holiday, were off of work or similar[11]. In the case of a written response, these questionnaires were not electronically recorded unless they contained any further information. An indicator was formed for the online section encompassing cases such as the one described and others. Expressed in figures, the statistics for the online section are presented in Table 8.

---

[9] For reasons of data privacy, the one-to-one identifier that permits a link between the CATI section and the main survey is not printed.
[10] The original data include on case with an identifier that does not match with the main survey. This case was removed from the research data at the end of the data preparation.
[11] People explicitly indicated this on the questionnaire and did not make any further entries in such cases.

Table 8: Qualitative evaluation of the online section of the Supplemental Task Survey

	Number	in %
Online responses in total	1,517	100.0
Of which at least one entry regarding time/task	1,224	80.7
Daily working time zero	136	9.0
Daily working time not indicated	60	4.0
No time information on any task	44	2.9
Number of hours, day of week missing	53	3.5

Rounding error possible for percentages
Source: raw data set for the online section of the Supplemental Task Survey

If only the daily working time (or the day of the week) was missing and time information was provided for the individual tasks, an attempt was made to reconstruct missing information (see Section 5.2). If this was not possible, the observation units that fell into the lower four lines of Table 8 were removed from the evaluation data set[12].

For data capture reasons, the raw versions of the total of approximately 2,300 data lines have been given the variable names T1 to T21. T1 to T18 represent the standardised task items from the main survey (F303 to F320), although it should be noted that due to the individually prepared task lists, T1 for one person could mean a different task item than for other persons. T19 to T21 here always represent the up to three additional tasks of the participant's choosing, to be entered as needed. In the initial data preparation step, these variable names were again adapted to the designation of the task items from the main survey and the CATI section of the supplemental survey:

Hours:	F303_z_Std to F323_z_Std
Minutes:	F303_z_Min to F323_z_Min
Task frequency:	F303_z_typ to F323_z_typ
Level of difficulty:	F303_z_grad to F323_z_grad

In the case of the last two variable groups, some respondents were unable to decide on a whole number value, meaning that in some cases intermediate figures (1.5; 2.5; 3.5) represent valid values[13].

Comments and notes (the latter, for example, when something was not (easily) legible) from the data capture process are kept separately in the SUF and can be evaluated and viewed with a usage agreement as part of a guest stay or via remote data processing.

---

[12] On request and upon submission of an application for usage, the excluded cases can gladly be made available.
[13] Another illustration of this is the phrase "five times a week" in the case of activity frequency. Accordingly, variables for activity frequency and difficulty level are formated as strings in the raw version of the data.

However, the BIBB-FDZ offers the research data from the second part of the Supplemental Task Survey in this raw version to external scientists only on explicit request, because evaluating this raw version of the data usually requires several somewhat extensive and in particular complex data preparation steps prior to actual data analysis. The BIBB-FDZ has performed several of these steps as standard. They are discussed in the next section and must be independently reviewed by users prior to processing of the raw data.

5.2 Data preparation steps for a service data set

The following adjustment steps build on each other in the order listed. Accordingly, the following individual steps represent the history of the transformation from the raw data set to the service data set[14].

Transformation of string variables into numeric variables

The first preparation step consists of recoding information on task frequency and level of difficulty such as "1 to 3," "2/3", etc. into useful numerical values (such as "1 to 2" into "1.5"). The string data format is then transformed into a numerical one.

The day of the week was also originally coded as a string and is transformed into a numerical variable (from 1 = Monday to 7 = Sunday).

Resorting the variable sequence

The sequence of the variables in the service data set has been resorted so that all hour information (F303_z_Std to F323_z_Std) is followed by all minute information (F303_z_Std to F323_z_Std) and then variables for task frequency (F303_z_typ to F323_z_typ) and level of difficulty (F303_z_grad to F323_z_grad).

Correction of individual cases

Corrections were made to individual cases in five of the data lines. The corrections related almost exclusively to the correct entry of hours and minutes for individual or multiple tasks (for example transforming hour information "270 min." into "3 hours and 30 minutes").

Allocation of free text entries to the standard task items F303 to F320

In many cases, free text entries regarding what respondents feel are additional tasks as compared to the main and supplemental survey can be easily allocated to standard items. In total, there were approximately 400 responses in which additional tasks were listed (approximately 300 in the online survey and some 100 in the written survey). Each of these units was reviewed to determine if the information could be allocated to the standard items. If this was the case, the time for these tasks was added to the already existing time calculation.

---

[14] The BIBB-FDZ does not currently archive "interim data sets", i.e. data sets on which only some of the data preparation steps have been performed. On justified request, they can be generated at any time from the raw data. It should be noted here that a certain adjustment step contains all previously mentioned adjustments (example: in an interim data set in which missing values are replaced by a zero, the string variables have already been transformed into numerical variables. In addition, the sequence of variables has been changed and individual cases have been corrected).

The mean was calculated for the task frequency and the level of difficulty if the corresponding standard item already contained information. If the standard task item did not yet contain any information, the corresponding value for the free text entry was transferred into these two fields.

Approximately 80 per cent of free text entries could be allocated to the standardised task items using this procedure. The remaining free task entries are represented by the variables F321_z_task and F322_z_task; i.e. the third free text field was no longer needed after the allocation procedure. The research data set contains a code korr_freitext (equal to one) that indicates if free texts were allocated to standard items for the corresponding observation unit and, accordingly, the time information and the two other pieces of information regarding this task were corrected.

Missing values for time information set to zero

For many of the calculations and applications, information on tasks in one form or another is added or grouped. Missing values for tasks not carried out on the reporting day quickly lead to problems in data preparation[15].

Converting time information to industry time

Adding together information on minutes, for instance, requires corrections of both minute and hour information for values over 60. Conversion to industry minutes (formula: number of minutes divided by 60) and merging with the hour information makes it possible to carry out arithmetic operations with the individual time data without additional processing steps. The number of variables is reduced as well. The new variables are labelled F303_z_zeit to F323_z_zeit and replace the variables F303_z_Std to F323_z_Std as well as F303_z_Min to F323_z_Min.

Harmonising time information on individual tasks with daily working time

This requires several consecutive tests. The test results can usually be read in additional variables in the data set. The first such additional variable is the qualitative code kz_zeit that summarises statements on total time information for individual tasks with the daily working time indicated. It is distributed as follows:

---

[15] Nevertheless, missing values always occur when a standard activity item is not part of an individual's occupational profile.

Table 9:   Qualitative ratio of total working time and the sum of the individual tasks

	Number	in %
Total	2,273[16]	100.0
Of which daily working time and sum of the individual tasks present	2,233	98.2
Time information for individual tasks missing	9	0.4
Daily working time not indicated	26	1.2
No time information on any task No weekly working time indicated	5	0.2

Source: Supplemental Task Survey, Part II

In cases in which information on daily working time is missing, this data is replaced by the sum of the time information for the individual tasks. If the sum of the individual tasks is less than 24 hours, the calculated value is transferred (23 cases). If the sum of the time information for the individual tasks is more than 24 hours, it is assumed that the respondents are providing time information for one week. In this case, the sum of the time information for the individual tasks is divided by five and transferred as daily working time (3 cases).

In order to harmonise the daily working time corrected in this manner and the sum of the times for individual tasks, the deviation is first determined as an absolute number (range: from minus 96 to plus 11 industry time hours[17]). If the deviations are greater than 7.5 per cent and thus considerable, this can be ascertained in the research data set based on the code kz_abw = 1. A correction factor is then calculated on an individual basis for all cases by dividing the daily working time by the sum of the time information for individual tasks. The distribution of the correction factor is as follows:

Table 10:   Distribution measures of the correction factor for harmonisation of daily working time with the sum of the time information for individual tasks

Distribution measure	Value
Mean	0.98
Standard deviation	2.37
Median	1.00
10 percentile	0.48
25 percentile	0.74
75 percentile	1.00
90 percentile	1.01

Source: Supplemental Task Survey, Part II

---

[16] The research data include 2,272 cases (see sec. 6.3).
[17] In the case of negative values, the sum of the time information for individual activities is greater than the daily working time.

The values are above one in a minority of cases only. In this case, the sum of the time information for the individual tasks is smaller than the daily working time indicated. Much more frequently, the sum of the individual time data is larger than the daily working time (the correction factor is nevertheless smaller than one). In order to harmonise the sum of the time information for the individual tasks with the daily working time information for all cases, the individual time data is multiplied with the correction factor. The corresponding new variables are labelled F303_z_zeit_korr to F320_z_zeit_korr.

If needed, the times originally indicated can easily be reproduced by dividing the time information for the individual tasks by the correction factor, which has also been included in the service data set. In addition, the correction factor can help users exclude implausible daily working times and individual time sums from the analyses.

A complete overview of all variables in the research data from Part I and Part II of the Supplemental Task Survey is available in Appendix A3.

6   Data access and data products

6.1   Data protection

Access to BIBBs' SUF data is exclusively granted in line with the European data protection regulations, in particular Chapter 89 of the Regulation (EU) 2016/679: General Data Protection Regulation (GDPR) and its national amendments, in particular § 27 of the German Data Protection Act (BDSG). Accordingly, data may be left for independent scientific research if it is impossible to establish a reference to a unique survey unit ("anonymity"). In order to achieve this goal without any exception, the organization and/or the researcher(s) have to establish specific technical and organizational data security arrangements, making unauthorized access to BIBBs' SUF data impossible, respectively. BIBBs' Research Data Centre (BIBB-FDZ) recommends that researchers with data access be specifically obliged to complying with the GDPR (c.f. our respective SUF using guidelines).

BIBB-FDZ staff get insights into research questions, methods, and analyses of researchers only for the purpose of providing advice, improving the BIBB-FDZ service, and ensuring compliance with the guidelines of the GDPR. BIBB employees who are not belonging to BIBB-FDZ staff get no insights into any activities of researchers.

6.2 Data access

Part I (CATI) and Part II (online/written) of the Supplemental Task Survey can be procured via the usual application procedure through the Research Date Centre of the Federal Institute for Vocational Education and Training for scientific research purposes. Unless otherwise specified, the service data set is supplied for Part II of the research data. In addition, the BIBB-FDZ checks whether an application for usage has been submitted for the main data set of the employment survey because usage of the research data for the Supplemental Task Survey is usually dependent on the main data set for the Employment Survey 2012. Only the latter contains, for example, information on socio-demography, job characteristics and occupational education processes of the respondents in the Supplemental Task Survey[18].

6.3 Data products

Part I and Part II of the Supplemental Task Survey are made available in the form of Scientific Use Files. The (remaining) free text entries regarding individual tasks are contained in the SUF Part II. All other free text entries (Part I: unaided labelling of the occupation, industry as well as unaided labelling of reasons for change; Part II: respondent comments) are not included in the SUF for reasons of data privacy. As usual, they can be evaluated via remote data processing or at the safe centre at BIBB-FDZ in Bonn.

The data collecting institute recoded the information on respondents' current occupation and industry (if different to the one of the main interview) into the 5-digit level *Klassifikation der Berufe 2010* (German classification of occupations) and the 2-digit level of the *Wirtschaftszweigeklassifikation WZ2008* (German industry classification, corresponds to NACE Rev.2 / ISIC Rev. 4 at this level). Both recodings are included in the research data (occupational codes at the 3-digit level are included in the SUF, 4-/5-level codes can be accessed via remote access or at the safe centre in Bonn. Similarly, full texts on further task performed at the workplace and reasons for changing the workplace were recoded into numeric variables. In case of very specific responses or multiple answers such a recode was not feasible and thus these cases were recoded into a category named "Sonstige" (other). As stated above the full texts on further tasks performed are included in the SUF so that the recodings can be reconstructed. Open answers for the question on reasons for job change can only be analysed via remote access or at the safe centre in Bonn.

---

[18] In particular for the analysis of occupational mobility, the initial status of occupationally mobile persons can only be determined using data from the main survey.

6.4   Changes in SUF version 2.0 and Errata

Part I:

Imputed / newly generated working time variable: az_t2_imp, F12xx_neu, az_t2, indic_korr

Important measures of the BIBB Supplemental Task-Survey include wage changes and changes in contractual (*F11Std*, *F11Min*) and real working hours (*F12Std*, *F12Min*). After field work was finalized it became obvious that the absolute amount but not the signs of the change in working hours were collected. Thus, from the data it cannot be derived whether the working hours have increased or decreased by the respective hours. On the basis of measures of the main survey and the Supplemental Task-Survey the missing signs for real working hours were imputed. The procedure is described in detail in the BIBB-FDZ data and methodological report 1/2015 (*forthcoming*). The research data of the BIBB-Supplementary Task-Survey to the Employment Survey 2012 (version 2.0) includes the imputed and newly generated variables *az_t2_imp*, *F12xx_neu*, *az_t2* und *indic_korr* in addition to the original variables.

F5code

The SUF (2.0) includes the occupational code of the *Klassifikation der Berufe 2010* (German classification of occupations) at the 3-digit level.

F7_99 (reasons for job change: no answer)

The value labels of F7_99 were not labelled in version 1.0 of the SUF.

F10 (change in gross monthly earnings in Euro)

This variable includes 116 system missings for respondents who before stated that their earnings have increased or decreased, respectively.

F11-F13Std,Min

Some missing values were not coded according to the filter instruction in the questionnaire but as system missings. This was corrected in version 2.0.

F13Std, F13Min

Because of a wrong filter instruction in the CATI program at the beginning of the field work, for all workers without a change in working hours the actual working hours (*F13Std*,*Min*) were surveyed. After this was noticed the filter was corrected. The 767 cases without changes in working hours as stated in *F12* but with information on *F13Std*,*Min* were recoded as filter missing in version 2.0.

F9-F16_22

Respondents who were unemployed or had meanwhile retired or become full-time apprentices were guided to the end of the interview (see questionnaire). The original data nevertheless have some (mostly 1 to 3) valid cases on variables F9-F16_22. These cases were recoded as filter missings in version 2.0.

Part II:

Deletion of one case

One case in part II has an identifier that is not to be found in the main survey. We suppose that the identifier was wrongly recoded during the manual entry of the postal responses. This case was dropped in the research data.

Appendix

Appendix A1: Sample of the written questionnaire

Question 1: What day of the week is it today? _____ (please enter)

Question 2: How long did you work today in total? ____ hours ____ minutes

Question 3: What tasks did you perform today and how long did each of them take you?

In the <u>first column,</u> please enter <u>how much time (in hours and/or minutes)</u> you spent on performing the task indicated in the list. If you cannot find an task that you have performed today in the list, you can add tasks in the last lines. If you performed several tasks at the same time, please divide the time equally between the tasks in question**.** Please check to make sure that the sum of the times you have entered is approximately equivalent to your total working time today.

In the <u>second column,</u> please indicate <u>how typical</u> this task is for your occupation. If this task is performed daily, please enter a 1; if it is performed relatively often, please enter a 2; if it is performed relatively seldom, please enter a 3; if it is almost never performed, please enter a 4.
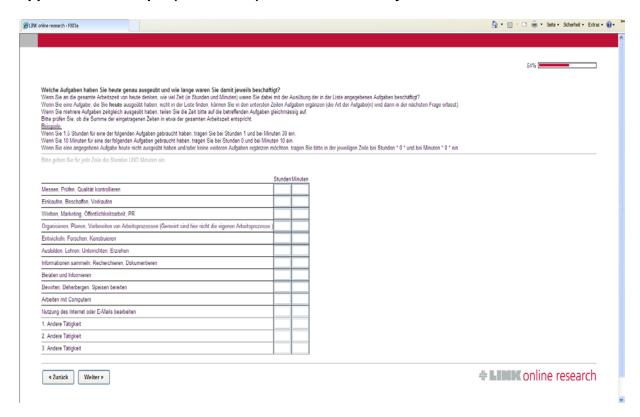
In the <u>third column,</u> please indicate <u>how difficult</u> it is for you – based on your skills and knowledge – to perform this task. 1 stands for especially easy, 2 for relatively easy, 3 for relatively difficult and 4 for especially difficult.

	Column 1		**Column 2**	**Column 3**
Task	<u>How much time</u> did you spend on this task?		<u>How typical</u> is the task? It occurs … 1= daily 2 = relatively often 3 = relatively seldom 4 = almost never	<u>How difficult</u> is the task? 1=especially easy 2=relatively easy 3=relatively difficult 4=especially difficult
	Hour(s)	Minutes		
T1				
T2				
T3				
T4				
T5				
T6				
T7				
T8				

(Notes and additions can be made on the reverse side.)

Appendix A2: Sample (screenshot) of the online survey

Appendix A3: Overview of variables

Variable name(s)	Meaning	Data product
intnr	Artificial ID as key variable for Part 1 and 2 as well as the main Employment Survey 2012	All data products
F1	Change of job and/or employer	Part 1, SUF
F2	Interim unemployment	Part 1, SUF
F3	Target person is supervisor	Part 1, SUF
F3a	Number of employees for which target person is supervisor	Part 1, SUF
F4	Other occupation than in the main survey	Part 1, SUF
F5	Current main occupation	Part 1, SUF
F5code	KldB 2010 (5-digit) of current main occupation[19]	Part 1, SUF
F303_z to F320_z	Query regarding individual occupational tasks as in the main survey in items F303 to F320	Part 1, SUF
F321_offen	Free text field for designation of additional tasks	Part 1, SUF
F321_z	Other additional tasks	Part 1, SUF
F7_01 to F7_08	Reasons for change in job/employer	Part 1, SUF
F7_09 to F7_11	Unaided labelling of change reasons yes/no	Part 1, SUF
F8_01 to F7_08	Fulfilment of expectations in change	Part 1, SUF
F8_09 to F7_11	Unaided labelling: expectations in a change yes/no	Part 1, SUF
F9	Change in gross monthly earnings with change of job	Part 1, SUF
F10	Amount of change in earnings	Part 1, SUF
F10a	Earnings significantly higher/lower than before (alternative to F10)	Part 1, SUF
F11	Change in contractual weekly working time	Part 1, SUF
F11Std	Amount of change (hours) in weekly working time	Part 1, SUF
F11Min	Amount of change (minutes) in weekly working time	Part 1, SUF
F12	Change in actual weekly working time	Part 1, SUF
F12Std	Amount of change (hours) in weekly working time	Part 1, SUF
F12Min	Amount of change (minutes) in weekly working time	Part 1, SUF
F13Std	Average weekly working time (hours)	Part 1, SUF
F13Min	Average weekly working time (minutes)	Part 1, SUF
F14	Knowledge of industry wages and working conditions	Part 1, SUF
F15_1 to F15_8	Query regarding – "requirements" (see questionnaires)	Part 1, SUF
F16_1 to F16_22	Query regarding – "situations"(see questionnaires)	Part 1, SUF

---

[19] Although individual qualifications are not indicated here for space reasons, one term a piece was chosen for the labelling and more accurately expresses a qualification than a requirement.

F16a	Change in industry since main survey	Part 1, SUF
F16a_code	New industry, coded in accordance with WZ 2008	Part 1, SUF
az	Working time information from the main survey	Part 1, SUF
gew_task1	Projection factor for adaptation to the population	Part 1, SUF
F5offen	Unaided labelling of current occupation	Part 1, GWA/DFV
F7_09offen to F7_11offen	Unaided labelling of reasons for change	Part 1, GWA/DFV
F16aoffen	Unaided labelling of the new industry	Part 1, GWA/DFV
az_tag	Working time on reporting day in industry time	Part 2, SUF
Wochentag	Reporting day for time information	Part 2, SUF
Stunden_ges	Total number of hours on reporting day	Part 2, SUF
Minuten_ges	Total number of minutes on reporting day	Part 2, SUF
kz_zeit	Code indicating if sum of individual data on tasks and/or daily working time was indicated by the respondent	Part 2, SUF
kz_abw	Code indicating if sum of individual data on tasks and daily working time vary significantly (>7.5 per cent)	Part 2, SUF
F303_z_zeit_korr to F320_z_zeit_korr	Information in industry time regarding how much time was spent on the standardised task items F303 to F320 from the main survey on the reporting day (if necessary revised upwards through allocation of free text entries)	Part 2, SUF
F303_z_typ to F320_z_typ	Frequency with which standardised task items are carried out over an extended time period	Part 2, SUF
F303_z_grad to F320_z_grad	Subjectively perceived level of difficulty of the standardised task items	Part 2, SUF
summe_einzeltätigkeiten	Total working time according to sum of time information for individual tasks	Part 2, SUF
korrekturfaktor	Factor with which the sum of the time information for the individual tasks is multiplied in order to harmonise with the daily working time indicated	Part 2, SUF
quelle	Origin of the data line online/written	Part 2, SUF
gew_tasks2	Projection factor for adaptation of the participants in Part 2 to the population of the main Employment Survey 2012	Part 2, SUF
anmerkung	Comments by respondents regarding Part 2 of the Supplemental Task Survey	Part 2, GWA/DFV

28

Research Data Center

BiBB FDZ

**Federal Institute for Vocational
Education and Training**

Legal notice