Weighting in PIAAC-L 2016

Burkhardt, Luise; Silbermann, Tobias; Bartsch, Simone

Leibniz-Institut
für Sozialwissenschaften

GESIS **Papers**

2018|15

Weighting in PIAAC-L 2016

*Luise Burkhardt, Tobias Silbermann, &
Simone Bartsch*

GESIS Papers 2018|**15**

Weighting in PIAAC-L 2016

Luise Burkhardt, Tobias Silbermann, &
Simone Bartsch

GESIS – Leibniz-Institut für Sozialwissenschaften 2018

GESIS Papers

GESIS – Leibniz-Institut für Sozialwissenschaften
Survey Design and Methodology
Postfach 12 21 55
68072 Mannheim
Telefon: (0621) 1246 - 518
Telefax: (0621) 1246 - 100
E-Mail: anouk.zabal@gesis.org

PIAAC-L was funded by the Federal Ministry of Education and Research, Berlin (Grant number 01 JP 1301 A, B, C).

Correspondence: Luise Burkhardt, German Socio-Economic Panel Study (SOEP), DIW Berlin, Mohrenstraße 58, Berlin, Germany, Phone: +49 30 89789-438, lburkhardt@diw.de

Acknowledgments: The authors would like to thank the members of the project team, in particular Silke Martin and Anouk Zabal, for their continuous support as well as for their help as reviewers.

ISSN: 2364-3781 (Online)
Publisher, printing and distribution: GESIS – Leibniz Institute for the Social Sciences
Unter Sachsenhausen 6-8, 50667 Cologne, Germany

Contents

1. Introduction 5
2. Modelling Nonresponse and Nonresponse Weights 7
   2.1 Dependent Variables and Response Rate 7
   2.2 Independent Variables 8
   2.3 Model 1 – Noncontact 11
   2.4 Model 2 – Nonresponse 13
   2.5 Final Nonresponse Weights 15
3. Calibration and Delivered Weights 16
   3.1 Usage of Weights 17
   3.2 Reduction of Bias 18
4. Summary and Outlook 21
References 22
Appendices 24

1. Introduction

In 2011/2012, key adult competencies were assessed in 24 countries (including Germany) as a part of the OECD Programme for the International Assessment of Adult Competencies, PIAAC (Zabal et al., 2014). In order to enrich the analytical power of the PIAAC data, the German PIAAC-Longitudinal Project (PIAAC-L)[1] follows up the original German PIAAC 2012 respondents that could be re-contacted, as well as members of their households, ages 18 and over, with three additional waves of data collection (in 2014, 2015, and 2016). This study is a cooperative project of GESIS – Leibniz Institute for the Social Sciences, the National Educational Panel Survey (NEPS) at the Leibniz Institute for Educational Trajectories (LIfBi), and the Socio-Economic Panel (SOEP) at DIW Berlin and combines research questions and measurement instruments from all three institutes (Martin, Zabal, & Rammstedt, 2018; Rammstedt, Martin, Zabal, Carstensen, & Schupp, 2017; Zabal, Martin, & Rammstedt, 2016).

The present paper describes the weighting process for the last of the three PIAAC-L waves, by illustrating the two classical weighting steps nonresponse adjustment and post-stratification or calibration. For the documentation of the weighting process for the first and second PIAAC-L waves as well as for information on the overall PIAAC and PIAAC-L weighting concept, see Bartsch, Poschmann, and Burkhardt (2017) and Burkhardt and Bartsch (2017). As weighting in PIAAC-L follows a consistent concept across waves, a large part of the present paper was taken literally from these preceding papers. Where applicable, updates or adaptations were undertaken, based on wave 3 weighting processes and data from dataset ZA5989_Weights_16. This dataset is one of twelve sub-datasets that were released for 2016 as part of the PIAAC-L database that encompasses datasets from 2014, 2015, and 2016.[2]

Please note: Weights adjust for biased estimates due to nonresponse. Since PIAAC-L is a follow-up study of PIAAC 2012, it has the complex sampling design of PIAAC 2012. Thus, methods of variance estimation that account for survey design features need to be used when analyzing PIAAC-L data in order to obtain correct variances and standard errors. In addition, when plausible values are included in the analyses, the imputation variance must be taken into account when computing error variance (see Perry, Helmschrott, Konradt, & Maehler, 2017, p. 12-14). There exist two commonly used variance estimation methods that account for complex sample designs: replication and Taylor Series Linearization. For PIAAC 2012 data, the replication approach was used and replicate weights were created by the International Consortium (for further details see Zabal et al., 2014, p. 94-95). For PIAAC-L, no replicate weights are computed. Instead, Taylor Series Linearization can be used to compute variances (for further details see Kreuter & Valliant, 2007), since the PIAAC 2012 scientific use file contains variables on sampling and stratification, such as ID_PSU, STRAT_PSU, Federal_state and GKPOL. Examples of Stata code illustrating how to account for the complex sample design using Taylor Series Linearization and how to compute the imputation variance when plausible values are included in the analysis, can be found in the accompanying data documentation *Notes to the User* as well as in appendix B (Example B1 & Example B2) of the present paper. [3]

---

[1] Commissioned by the Federal Ministry of Education and Research, Berlin, Grant number 01 JP 1301 A, B, C.

[2] Last update 14.12.2017: GESIS – Leibniz Institute for the Social Sciences, German Socio-Economic Panel (SOEP) at DIW Berlin & LIfBi – Leibniz Institute for Educational Trajectories (2017): PIAAC-Longitudinal (PIAAC-L), Germany. GESIS Data Archive, Cologne. ZA5989 Data file Version 3.0.0, doi: 10.4232/1.12925.

[3] Accessible under https://dbk.gesis.org/dbksearch/sdesc2.asp?no=5989&db=e&doi=10.4232/1.12925.

The present paper is subdivided into two parts. The first part deals with the nonresponse analysis in PIAAC-L 2016 and the calculation of the nonresponse weight (bleib_16). The second part covers the modeling of the cross-secitonal weight (hrf_16) for PIAAC-L 2016.

2. Modelling Nonresponse and Nonresponse Weights

The nonresponse analysis in PIAAC-L 2016 comprises two parts: modelling of noncontact and modelling of nonresponse. This is in line with the weighting procedures applied in PIAAC-L 2014 and PIAAC-L 2015 (see Bartsch et al., 2017; Burkhardt & Bartsch, 2017). Hence, the following sections 2.1 to 2.5 are in large parts taken literally from these preceding papers and are, where applicable, updated based on wave 3 weighting processes. For each of the two steps a logistic regression model is estimated, where the dependent variable is a 0/1-variable (nonresponse/response) and the independent variables were selected as explained below. The predicted probabilities (of contact and response respectively) are derived from each of these two models and the inverse of the product of these probabilities yields the nonresponse weights.

2.1 Dependent Variables and Response Rate

In the third wave of PIAAC-L, all anchor persons and their household members of age 18 or above were eligible for participation. Please note, that weighting was only performed for anchor persons and not for other household members who participated in PIAAC-L only[4]. Hence, the gross sample for weighting in 2016 (i.e. field phase in 2016) is comprised of n=3,510 eligible anchor persons. Thereof, n=3,263 are participants in PIAAC-L 2015 and n=247 are temporary non-participants in PIAAC-L 2015, i.e. those n=247 cases participated in PIAAC 2012 and PIAAC-L 2014, but did not participate in PIAAC-L 2015. In PIAAC-L 2016 they belong to the gross sample. The gross sample was further divided into four groups: nonresponse due to noncontact (M1), nonresponse due to refusal or due to other reasons (e.g. long-term illness, language problems) (M2), ineligible persons due to death or moving abroad and respondents (see Table 1). The realized anchor person net sample in 2016 amounts to n=2,967. Thereof n=67 are temporary non-participants, that did participate in PIAAC-L 2014 but not in 2015 and rejoined the sample in PIAAC-L 2016. Those n=67 persons will be referred to as temporary dropouts. They will be addressed in terms of weighting in more detail in section 3.1 of this paper.

Table 1 describes the results of the fieldwork for PIAAC-L 2016. The adjusted response rate, i.e. the number of valid interviews with eligible anchor persons (n=2,967) divided by the number of eligible anchor persons in the gross sample for fieldwork excluding neutral (ineligible) nonrespondents (n=3,494) yields 84.9%. The unadjusted response rate, referring to the gross sample including neutral (ineligible) nonrespondents (n=3,510), amounts to 84.5% (Table 1 and Steinacker, Wolfert, & Thümmel, 2017, pp. 36).

A nonresponse rate of 15.1% of the gross sample including nonresponse due to noncontact may lead to a nonresponse bias which should be corrected for (Groves & Peytcheva, 2008). With respect to the net sample of PIAAC-L 2015 (which included 3,263[5] cases) and to the 67 temporary dropouts that did not take part in PIAAC-L 2015, the analysis was undertaken in two steps, as a differentiation between

[4] This is due to the different survey design approaches in PIAAC and PIAAC-L (for further information see Bartsch et al., 2017 and Burkhardt & Bartsch, 2017). Readers not familiar with the PIAAC 2012 weighting procedures are referred to Zabal et al. (2014, p. 80) and Mohadjer, Krenzke, and Van de Kerckhove (2013) for further information.

[5] Please note that the numbers in Table 2 differ because neutral (ineligible) dropouts, i.e. people who moved abroad or deceased (n=16) were not included in the analyses, following the standard weighting approach of the SOEP (Kroh, Käppner, & Kühne, 2014).

nonresponse due to noncontact and nonresponse due to refusal or other reasons is appropriate and is performed in established panel studies like the SOEP (Kroh, 2014).

Table 1 Results of Fieldwork for PIAAC-L 2016

Final results	abs.	%
Interview		
Interview valid – household interview + person interview with PIAAC anchor person available	2,946	83.9
Interview valid – person interview with PIAAC anchor person but no household interview available	21	0.6
Ineligible		
Anchor person moved abroad	8	0.2
Anchor person deceased	8	0.2
Noncontact (M1)		
Declined participation before start of fieldwork	13	0.4
Anchor person moved to unknown address	43	1.2
Anchor person moved to known address	9	0.3
Address no longer exists / address not inhabited	5	0.1
Anchor person unknown at given address	1	0.0
No one home	45	1.3
Nonresponse (M2)		
Interview impossible during fieldwork	41	1.2
Anchor person unable to respond due to long-term illness or other reason	11	0.3
Linguistic problems, inadequate German skills	1	0.0
Anchor person refused to participate in interview	230	6.6
Person in household refused contact to anchor person	3	0.1
Correct address, but anchor not met in person	7	0.2
Other reasons, unusual circumstances	93	2.6
Contact established without final result	7	0.2
Interview invalid – no person interview with PIAAC anchor person[6]	18	0.5
Gross sample	3,510	100.0

2.2 Independent Variables

Modelling of nonresponse aims at consistent estimation of response propensities that can serve as a basis for weighting factors and should compensate for attrition. Nonresponse is modeled based on information, that is available for both, respondents and nonrespondents. As PIAAC-L is designed as a panel survey, those information can be taken from previous waves in order to explain potential selectivity in the following waves. The calculation of weights for PIAAC-L 2016 was mainly based on variables from PIAAC-L 2015. Furthermore, information from the PIAAC-L 2016 fieldwork was used (e.g., information on interviewer changes between waves and on the mode of initial contact). The selection of explanatory variables was based on established assumptions and theories in the field of

[6] In these 18 cases there are only interviews for one or more other persons in the household available. The requirement for a completed case is not met, because no interview with the PIAAC anchor person could be realized (see Steinacker et al. 2017, p. 37).

survey methodology concerning their power to explain nonresponse (also see Kroh et al., 2014) and was also aligned with the selection of explanatory variables that were used to model nonresponse in PIAAC-L 2014 and PIAAC-L 2015 (for further details see Bartsch et al., 2017; Burkhardt & Bartsch, 2017).

All independent variables were checked for association with the outcome variables – noncontact and nonresponse respectively – and those with a statistically significant association at the 10%-level were included in a full model.[7] This full model was subsequently reduced to those variables significant at the 5%-level. Due to the fact that the selection of the independent variables is data-driven and not theory-driven, apparently important explanatory variables such as gender, education, labor force participation and so on are not included in either both or at least one reduced model as they showed no significance at the 5%-level. For details on the final set of variables that were significant at the 5%-level, see Table 2. It is important to note, that there is no focus on an interpretation of the effects as we are mainly interested in the statistical explanatory power of the predictors that were integrated into the model (Spieß, 2010, p. 126).

In PIAAC 2012, for each assessed competency domain, literacy, numeracy and problem solving in technology-rich environments, 10 plausible values were computed per respondent (Yamamoto, Khorramdel, & von Davier, 2013). In order to investigate how these competencies change over time, additional competence data for literacy and numeracy was collected in PIAAC-L 2015 using PIAAC instruments as well as NEPS instruments for the measurement of cognitive competencies. In a next step plausible values were estimated. These plausible values along with the re-scaled PIAAC 2012 competency measures were longitudinally scaled with extended background information from PIAAC 2012, PIAAC-L 2014 and 2015 and are provided in the dataset ZA5989_Persons_15 (for details on the modeling of plausible values in PIAAC-L and on the re-scaling procedures see the accompanying data documentation *Notes to the User*[8]). In order to use the most current information on the competencies of PIAAC-L anchor persons for weighting in PIAAC-L 2016, the proficiency measures for literacy and numeracy that were assessed in PIAAC-L 2015, were used in the nonresponse adjustment models by first calculating the mean across all ten plausible values for each domain. In a second step, quartiles were calculated. For the competency domain problem solving in technology-rich environments, which was not reassessed in PIAAC-L 2015, the re-scaled plausible values from PIAAC-L 2014 were used in the same way. Additionally, a missing category was calculated for cases in which no plausible values on problem solving were available (n=461).[9]

---

7 For a list of all independent variables that were included in the full model, see Table A1 in appendix A.

8 Accessible under https://dbk.gesis.org/dbksearch/sdesc2.asp?no=5989&db=e&doi=10.4232/1.12925.

9 No plausible values were available for respondents who did a paper-based assessment in PIAAC 2012 because the paper instruments did not include the competency domain problem solving.

Table 2 Summary of Independent Variables That Were Included in the Reduced Model[10]

	Model 1 Noncontact	Model 2 Nonresponse
Interview Characteristics		
Interviewer change (from wave 2 to wave 3)	X	X
Partial unit nonresponse (nonresp. of eligible hh-member last wave)		X
Region and Household Characteristics		
Hamburg	X	X
Rhineland-Palatinate	X	X
Baden-Wuerttemberg	X	
Housing area: residential (new)		X
One person in household		X
Income and Employment Characteristics		
Income 4th quartile (>4000 Euros)		X
Income missing		X
Family and Partnership Characteristics		
Married	X	
At least one child in household		X
Education Characteristics		
Education missing for both parents		X
Other Characteristics		
Age 19-24[11]	X	
Age 25-34	X	
Migration background	X	
Cognitive Skills Characteristics		
Literacy: 3rd quartile		X
Literacy: 4th quartile		X
n	3,317[12]	3,245[13]

10 For a list of all independent variables that were included in the full model, see Table A1 in appendix A.

11 Here, respondent age ranging from 19-69 at the time of PIAAC-L 2015 was used as basis. In the post-stratification process, however, the actual age as reported by the respondent in the PIAAC-L 2016 interview is used for calibration.

12 These n=3,317 cases result of the anchor persons' gross sample for PIAAC-L 2016 (n=3,510) minus the temporary nonrespondents in PIAAC-L 2015, who did also not participate in PIAAC-L 2016 (n=180) and thus are not part of the weighting procedure, minus 16 neutral dropouts (hereof 3 who were part of the 247 temporary nonrespondents in 2015).

13 These n=3,245 cases result of n=3,317 respondents, who were included in the non-contact model minus 72 persons who participated in PIAAC-L 2015 and were non-respondents due to non-contact in PIAAC-L 2016.

2.3 Model 1 – Noncontact

Model 1 refers to step one of the analysis: predicting the probability for a successful re-contact in relation to the probability of an unsuccessful attempt to re-contact the eligible anchor persons from prior PIAAC-L waves. In this step an attrition of 116 cases was recorded, thereof 72 persons participated in PIAAC-L 2015 and 44 persons did not participate in PIAAC-L 2015. Model 1 includes all variables that were considered for the attrition analysis and showed significant association to the contact indicator at a 10%-level ($p<0.10$). As mentioned earlier, first a full model with all of these variables was calculated and then stepwise routines were run in order to calculate a reduced model with only significant factors at the 5%-level ($p<0.05$). Table 3 and Figure 1 show the outcomes of the full and reduced model. Figure 1 shows the coefficients and the 95%-confidence intervals for illustration. It can be seen that only few variables remain in the reduced model, which has a rather low statistical power but yet a relative better fit than the full model, according to the adjusted McFadden R-squared.

Table 3 Fit Values for Estimated Models for Step 1

	Full model	Reduced model
Observations	3,317	3,317
Pseudo-R^2	0.102	0.069
Adj. R^2 (McFadden)	-0.010	0.043

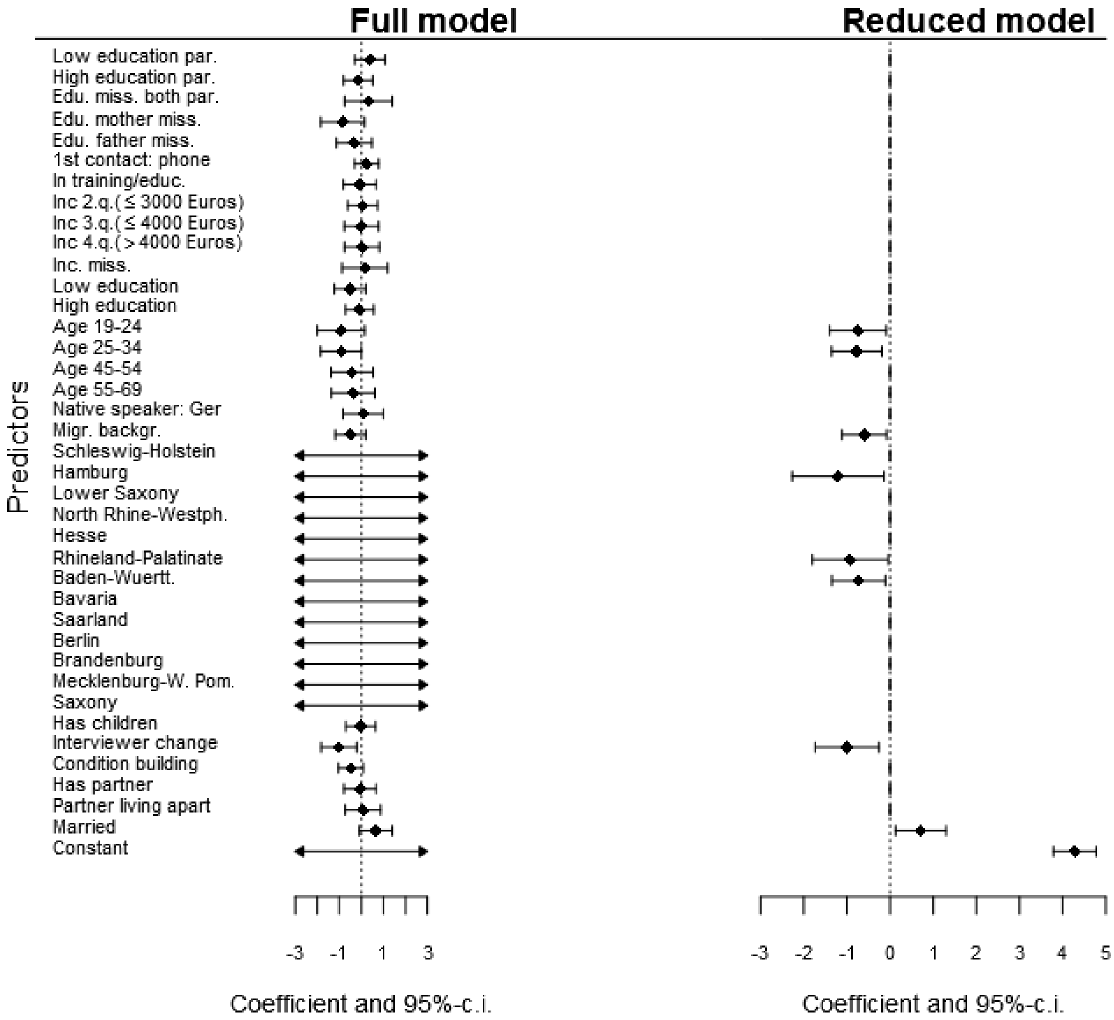

Figure 1 Coefficients and Confidence Intervals for Model 1 (Noncontact)[14]

[14] For a list of abbreviations used in Figure 1, see Table A1 in appendix A.

2.4 Model 2 – Nonresponse

Model 2 refers to step two of the analysis: The final nonresponse to the request of the interviewer to take part in PIAAC-L 2016. This is, in numerical terms, the biggest step in the attrition process, with 411 individuals refusing to participate again. The resulting statistical power of the reduced model is again rather low, considering the adjusted McFadden R-squared (see Table 4). Even though more variables from the PIAAC-L 2015 wave remain in the model, only few variables show statistically significant explanatory power at the 5%-level (see Figure 2). As mentioned earlier, the content-related interpretation of the models is not in the focus when modelling nonresponse.

Table 4 Fit Values for Estimated Models for Step 2

	Full Model	Reduced Model
Observations	3,245	3,245
Pseudo-R^2	0.066	0.043
Adj. R^2 (McFadden)	0.004	0.030

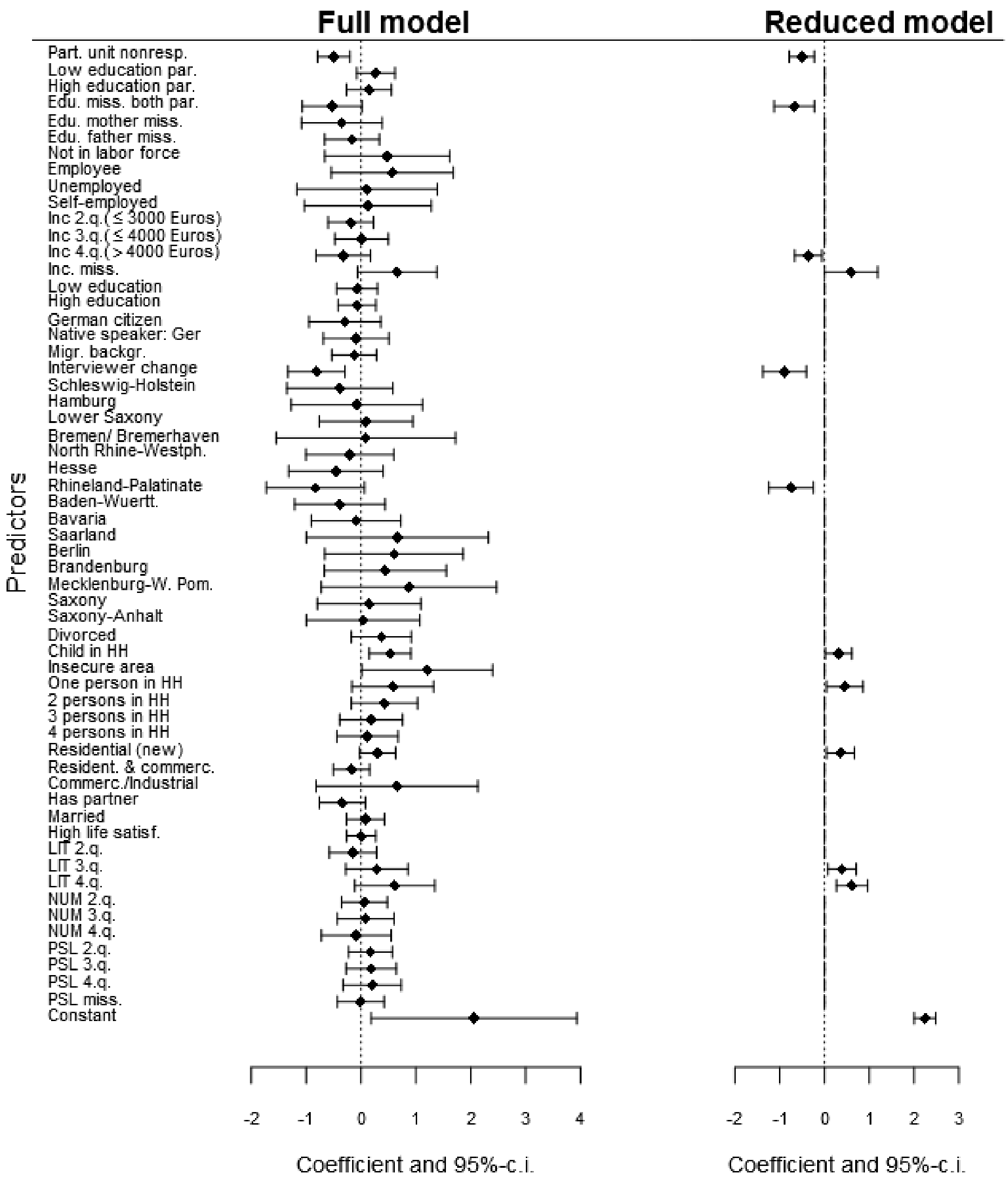

Figure 2 Coefficients and Confidence Intervals for Model 2 (Nonresponse)[15]

[15] For a list of abbreviations used in Figure 2, see Table A1 in appendix A.

2.5 Final Nonresponse Weights

To calculate the nonresponse weight, the staying probability for each of the two models was calculated. The individual's participation probability (P(WB=1)) is the product of the probability to be successfully recontacted for PIAAC-L 2016 and the probability to participate in the study: P(WB=1) = P(M1=1)*P(M2=1). The inverse of this participation probability yields the individual's raw nonresponse weight.

In PIAAC 2012 as well as in the SOEP, weights were trimmed when exceeding certain thresholds. The criteria used in PIAAC-L is to trim weights exceeding two times their median. The selection of the cut-off point is basically driven by the decision not to trim more than 1% of the derived weights and thus maintain efficacy. Since the nonresponse weights in PIAAC-L 2016 showed only low dispersion, no trimming procedure had to be applied (see Table 5). Their distribution shows an elongated right tail, indicating the few cases that lie outside the 95% percentile (see Figure 3).

Table 5 Estimated Nonresponse Weights (bleib_16)

	Min	10%	50%	75%	Max	Mean	SD
Final nonresponse weights	1.025	1.061	1.098	1.141	1.675	1.117	0.064

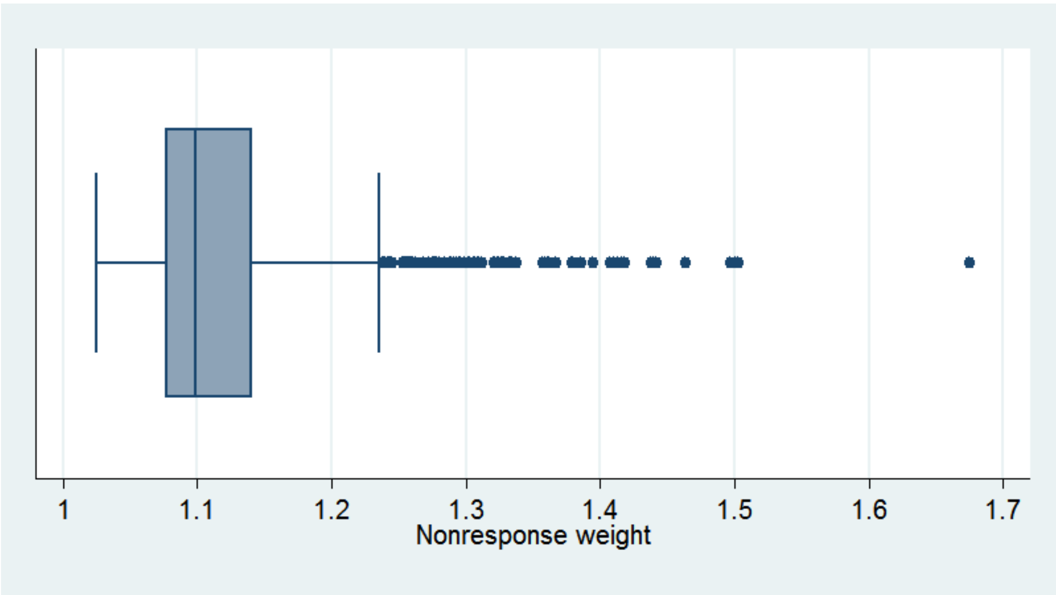

Figure 3 Distribution of Estimated Nonresponse Weights (bleib_16)

3. Calibration and Delivered Weights

The process of calculating weights described so far reflects the nonresponse adjustment. Calibration aims at bringing the sample in closer alignment with the underlying population, at least with regard to the distribution of some central variables. This is generally done by using data from official statistical sources. In the case of Germany, the Microcensus is the source for the reference data. The calibration process described here is in line with the procedures undertaken in PIAAC-L 2014 and 2015 (Bartsch et al., 2017 and Burkhardt & Bartsch, 2017). Hence, the following sections 3, 3.1 and 3.2 are in large parts taken literally from theses preceding papers and are, where applicable, updated based on wave 3 weighting processes.

Since the PIAAC-L sample had no refreshment of sample members, the reference population is described as *non-institutionalized adults born between November 1946 and November and have lived in Germany at least since 2012* and is the same as the reference population for PIAAC 2012. For this reference group, a separate count of the most up-to-date data from the 2016 Microcensus (Statistisches Bundesamt, 2017) serves as the basis for calibration. This means that the population to which the weights refer, consists of 51.42 million persons (Statistisches Bundesamt, 2017).[16]

Considering the reduced sample size due to the transition of PIAAC into a panel study, it was not possible to replicate the PIAAC post-stratification process one-to-one. As an alternative, a mixed approach between raking and post-stratification was chosen for the PIAAC-L post-stratification process (see Bartsch et al., 2017).

In PIAAC-L 2016 the combined table for raking used the variables gender (2 categories), age (5), and education (3)[17] as one raking reference at the individual level and the variables region (3), size of household (5), and size of municipality (7) as additional variables at the household level.[18] The additional variables were used according to the weighting approach of the German Socio-Economic Panel (Kroh et al., 2014), keeping in mind that the number of variables adjusted for should still be kept to a minimum. This technique is in line with the weighting procedures applied in PIAAC-L 2014 and 2015.

The basis for the raking procedure was the product of the nonresponse adjustment factor (bleib_16) and the cross-sectional weight from PIAAC-L 2015 (hrf_15). The descriptive statistics of the resulting weights are given in Table 6.

---

16 Please note: Due to its design it could not be ensured that the sample of the Microcensus precisely resembles the structure of the PIAAC-L sample, which, by definition, only includes persons in non-institutionalized households – both in 2011 and 2016. The reference sample might include a presumably small number of people that lived in institutions before 2016 as the Microcensus is a cross-sectional survey.

17 Since the subsample of pupils in school naturally decreases over time, the categories "Still in school" and "Low education level" were combined for calibration in PIAAC-L 2015 and 2016. Thus, the number of categories for education was reduced from four to three compared to PIAAC 2012 and PIAAC-L 2014.

18 For all variables, up-to-date information from 2016 was used.

Table 6 Descriptive Statistics of the Final PIAAC-L 2016 Weights

	Min	10%	50%	75%	Max	Mean	SD
Final weights	3169.87	7968.28	14585.74	20829.07	113364.37	17330.91	10597.74

3.1 Usage of Weights

In the dataset ZA5989_Weights_16, the two weighting factors bleib_16 and hrf_16 are included. The weighting factor hrf_16 can be used for cross-sectional analysis with data from PIAAC-L 2016 as it aims at adjusting the figures to the population benchmarks in 2016, at least with regard to some central variables. Bleib_16 is the product of the factors derived from the nonresponse analyses. For longitudinal analysis across multiple waves the cross sectional weight of the starting wave of interest should be multiplied with the nonresponse weights of the following waves. Table 7 provides an instruction on the appropriate combination of weights for longitudinal analysis of PIAAC and PIAAC-L.

Table 7 Usage of Weights for Longitudinal Analysis with PIAAC and PIAAC-L

Longitudinal Analysis with PIAAC and PIAAC-L	Combination of Weights
PIAAC 2012 and PIAAC-L 2014	SPFWT0 * bleib_14
PIAAC 2012 and PIAAC-L 2015, or PIAAC 2012, PIAAC-L 2014, and 2015	SPFWT0 * bleib_14 * bleib_15
PIAAC 2012 and PIAAC-L 2016, or PIAAC 2012, PIAAC-L 2014 and 2016, or PIAAC 2012, PIAAC-L 2015 and 2016, or PIAAC 2012, PIAAC-L 2014, 2015, and 2016	SPFWT0 * bleib_14 * bleib_15 * bleib_16
PIAAC-L 2014 and 2015	hrf_14 * bleib_15
PIAAC-L 2014 and 2016, or PIAAC-L 2014, 2015, and 2016	hrf_14 * bleib_15 * bleib_16
PIAAC-L 2015 and 2016	hrf_15 * bleib_16

Please keep in mind that the reference population as described above is limited to a certain age group and excludes people who moved to Germany after 2012. Also, only anchor persons–those who had participated in PIAAC 2012–have a weighting factor. The information of the other persons in the household can be used as context information in the analyses. Examples of the application of the PIAAC-L weights for cross-sectional and longitudinal analyses and their effects on the reduction of bias are depicted in Tables 8 and 9.

Excursus: temporary dropouts in PIAAC-L 2015

As mentioned in chapter 2.1, in PIAAC-L 2015 some anchor persons temporarily dropped out of the sample (n= 247 non-participants). Some of these cases participated again in PIAAC-L 2016 (n=67). These temporary dropouts can have various reasons of noncontact or refusal (see Kalton, 1986; de Leeuw, 2005) and create problems for longitudinal analyses as sample size decreases and estimates can be biased if the wave nonresponse is not properly addressed (Haunberger, 2011). Temporary dropouts are also problematic in terms of calculating weights, due to the fact that longitudinal weights are calculated as the product of the cross-sectional weight for the initial wave and *all* nonresponse weighting factors between the initial and the current wave. As information from PIAAC-L 2015 is naturally missing for these cases, the inverse staying probability of these temporary dropouts is zero for PIAAC-L 2015 – the year these persons dropped out temporarily; these cases are by definition not included in the PIAAC-L 2015 weighting dataset.

In order to adequately estimate the cross-sectional weight for PIAAC-L 2016 (hrf_16), which is based on the product of the nonresponse weight (bleib_16) and the cross-sectional weight from PIAAC-L 2015 (hrf_15), we applied augmentation strategies to model these weights for the temporary dropouts in order to include them in the nonresponse analysis. For relative time invariant variables such as sex, year of birth, migration background and the education level of the respondents' parents, we used information from PIAAC-L 2014. For information that might change over time, such as the respondents' citizenship, marital status, education level and household size, we also used data from PIAAC-L 2014 and aligned it with information from PIAAC-L 2016. When the information did not contradict one another, we used the information from PIAAC-L 2014. In order to fill the missing information from 2015, we additionally made use of retrospective data from PIAAC-L 2016, e.g. to check whether a respondent changed his or her job or employment status or whether a major life event happened before the field phase of PIAAC-L 2015. Modeling of nonresponse also included variables for which no reliable information was available for the temporary dropouts. For instance the respondents' income. As income might change drastically between PIAAC-L waves one to three, deriving the income for the 67 cases from information on the respondents' income in 2014 and 2016 would be based on many uncertain assumptions. Hence, for these variables the 67 temporary dropouts were included in the missing categeory. For the proficiency measures for literacy and numeracy that were assessed in PIAAC-L 2015, the temporary dropouts were also included in the missing category, since no up-to-date data was available for these cases. For the competency domain problem solving in technology-rich environments, which was not reassessed in PIAAC-L 2015, the re-scaled plausible values from PIAAC-L 2014 were used for the temporary dropouts.

By means of these augmentation techniques, we were able to model the inverse staying probability (bleib_16) and the cross-sectional weight (hrf_16) for the 67 temporary dropouts. The weights are included in the dataset ZA5989_Weights_16. However, the inverse staying probability for PIAAC-L 2015 (bleib_15) is still zero. Therefore, these cases cannot be included in longitudinal analyses based on a balanced panel covering the survey year 2015. Thus, they are automatically excluded from the analyses. Temporary dropouts are a common problem all panel studies face and which increases over time (see Kalton & Citro, 1995).

3.2 Reduction of Bias

In this section of the paper, examples of the application of the PIAAC-L weights for cross-sectional and longitudinal analyses and their effects on the reduction of bias are given.

Table 8 gives an overview of the distributions of selected population characteristics: gender, level of education and age in PIAAC-L 2014, 2015, and 2016 with and without the cross-sectional weights obtained via calibration. These distributions are compared to those in the German Microcensus of the respective years which serves as the basis for poststratification. From this table one can see, that the weighted estimates are in all cases closer to the estimates of the Microcensus than are the estimates obtained without weighting.

Table 8 Comparison of PIAAC-L and Microcensus Weighted Estimates of % Population by Gender, Education Level and Age, 2014-2016[19]

Gender

Year	PIAAC-L unweighted		PIAAC-L weighted		Microcensus weighted		Ratio	
	Male (1)	Female (2)	Male (3)	Female (4)	Male (5)	Female (6)	Male (3/5)	Female (4/6)
2014	49.0	51.0	50.1	49.9	50.1	50.0	1.00	1.00
2015	48.7	51.3	50.0	50.0	50.0	50.0	1.00	1.00
2016	48.6	51.4	50.1	49.9	50.1	49.9	1.00	1.00

Highest education level

Year	PIAAC-L unweighted			PIAAC-L weighted			Microcensus weighted			Ratio		
	Low (1)	Medium (2)	High (3)	Low (5)	Medium (6)	High (7)	Low (9)	Medium (10)	High (11)	Low (5/9)	Medium (6/10)	High (7/11)
2014	25.2	36.4	38.5	32.5	33.2	34.4	32.5	33.2	34.3	1.00	1.00	1.00
2015	22.2	36.2	41.7	32.0	33.0	35.1	32.0	33.0	35.1	1.00	1.00	1.00
2016	21.8	36.2	42.0	31.3	32.8	35.9	31.3	32.8	35.9	1.00	1.00	1.00

Age[20]

Year	PIAAC-L unweighted					PIAAC-L weighted					Microcensus weighted					Ratio				
	min-25 (1)	26-35 (2)	36-45 (3)	46-55 (4)	56-max (5)	min-25 (6)	26-35 (7)	36-45 (8)	46-55 (9)	56-max (10)	min-25 (11)	26-35 (12)	36-45 (13)	46-55 (14)	56-max (15)	min-25 (6/11)	26-35 (7/12)	36-45 (8/13)	46-55 (9/14)	56-max (10/15)
2014	15.9	18.2	19.9	25.1	20.9	12.7	18.5	20.5	25.1	23.2	12.4	18.8	20.1	25.4	23.2	1.02	0.98	1.02	0.98	1.00
2015	14.1	17.4	18.5	26.1	23.8	10.6	18.6	19.6	25.6	25.6	10.5	18.8	19.5	25.7	25.5	1.01	0.99	1.01	1.00	1.00
2016	12.2	17.2	18.3	26.4	26.0	9.1	18.2	19.3	25.9	27.6	8.7	18.9	19.0	25.9	27.6	1.05	0.96	1.02	1.00	1.00

[19] The underlying sample size in each survey year equals the number of anchor persons in the respective year. Thus, in 2014 N=3,758, in 2015, N=3,236 and in 2016, N=2,967.

[20] PIAAC-L is a panel study and follows up the original German PIAAC 2012 respondents that could be re-contacted, as well as members of their households, ages 18 and over. The reference population for PIAAC 2012 and all follow-up waves of PIAAC-L is described as non-institutionalized adults living in Germany, who are born between November 1946 and November 1995 and have lived in Germany at least since 2012. Hence, the minimum and maximum age of the respondents differs in each panel wave (2014: min=18, max=67; 2015: min=19, max=68; 2016: min=20, max=69).

Table 9 gives an overview of the reduction of bias by means of weighting in PIAAC 2012, PIAAC-L 2014, 2015 and 2016. In a first step, central sociodemographic indicators were estimated for the PIAAC 2012 cross-sectional sample by applying the PIAAC 2012 final full sample weight (SPFWT0). In a second step, a longitudinal estimation for the same indicators was conducted for the anchor persons' net sample in 2014, multiplying the final full sample weight of PIAAC 2012 with the nonresponse weight of PIAAC-L 2014 (SPFWT0 * bleib_14). In addition a longitudinal estimation for the anchor persons' net sample of PIAAC-L 2015 was performed, multiplying the final weight of PIAAC 2012 with the nonresponse weights of PIAAC-L from 2014 and 2015 (SPFWT0 * bleib_14 * bleib_15). Finally, a longitudinal estimation for the anchor persons' net sample of PIAAC-L 2016 was performed, multiplying the final weight of PIAAC 2012 with the nonresponse weights of PIAAC-L 2014, 2015, and 2016 (SPFWT0 * bleib_14 * bleib_15 * bleib_16). For PIAAC-L 2016 both, the unweighted results (raw) and the results after weighting are displayed in Table 9. For the selected variables, it can be seen that bias caused by nonresponse is in most cases reduced through weighting.

Table 9 Reduction of Bias by Means of Weighting

	PIAAC 2012 (N=5,379)[21]	PIAAC-L 2014 (N=3,758)	PIAAC-L 2015 (N=3,263)	PIAAC-L 2016 (N=2,900)[22]	
	weighted	weighted	weighted	raw	weighted
Gender					
Male	50.5	50.2	50.1	48.5	49.6
Female	49.5	49.8	49.9	51.5	50.4
Total	100.0	100.0	100.0	100.0	100.0
Highest education level including those still in school*					
Low education level	31.6	32.0	31.7	22.4	31.4
Middle education level	34.4	33.6	34.2	36.9	34.4
High education level	30.7	30.9	30.7	37.2	30.9
Still in school	3.3	3.4	3.5	3.5	3.3
Total	100.0	100.0	100.0	100.0	100.0
Birth cohort					
1946-1961	32.0	32.7	32.7	31.0	33.3
1962-1976	34.3	35.1	35.5	34.6	35.7
1977-1995	33.7	32.2	31.8	34.4	31.0
Total	100.0	100.0	100.0	100.0	100.0

*based on information from PIAAC 2012

[21] Eighty-six cases that are classified as literacy-related nonrespondents belong to the PIAAC 2012 net sample and thus have a weighting factor in 2012. They are included in the nonresponse adjustment in PIAAC-L. In 2012, only the information on age and gender was collected for them. As information on education as a variable of interest for the analyses of the reduction of bias is missing for these 86 cases, the PIAAC 2012 sample used here consists of N=5,379 (of N=5,465) cases.

[22] These 2,900 cases do not include temporary dropouts (n=67), i.e. anchor persons who participated in PIAAC-L 2014 and 2016 but did not participate in PIAAC-L 2015, due to the fact that their inverse staying probability in PIAAC-L 2015 yields zero.

4. Summary and Outlook

This documentation describes the weighting procedure for the third wave of PIAAC-L. Similar to PIAAC-L 2014 and 2015, the weighting strategy basically follows the approach of the SOEP as a panel study, but also takes into account the weighting procedure in PIAAC 2012. The results are weights that can be used both for longitudinal and cross-sectional analyses. These weights were delivered as part of the scientific use file, including data from the third wave of PIAAC-L which was released in mid December 2017.

Selectivity in PIAAC and PIAAC-L was detected for birth cohorts and in the area of education among others. The use of weights is thus recommended for analysis. As mentioned earlier, to account for selectivity between PIAAC-L 2015 and PIAAC-L 2016, for instance, users should multiply the respective weighting factor of PIAAC-L 2015 with the weighting factor bleib_16 (hrf_15 * bleib_16). Using hrf_16 will adjust the figures to the population benchmarks in 2016 and should be chosen for cross-sectional analyses of the data from PIAAC-L 2016.

References

Bartsch, S., Poschmann, K., & Burkhardt, L. (2017). *Weighting in PIAAC-L 2014.* GESIS Papers 2017|06. GESIS: Cologne.

Burkhardt, L., & Bartsch, S. (2017). *Weighting in PIAAC-L 2015.* GESIS Papers 2017|30. GESIS: Cologne.

de Leeuw, E. (2005). Dropout in longitudinal studies. Strategies to limit the problem. In D.C. Howell. & B.S. Everitt (Eds.). *Enyclopedia of Statistics in Behavioral Science* (pp. 515-518). Volume 1. Chichester: John Wiley & Sons.

Groves, R. M., & Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse-Bias. A Meta-Analysis. *Public Opinion Quarterly. 72*. (pp. 167-189).

Haunberger, S. (2011). *Teilnahmeverweigerung in Panelstudien*. Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-17710-6

Kalton, G. (1986). Handling wave nonresponse in panel surveys. *Journal of Official Statistics*; 2. (pp. 303-314).

Kalton, G., & Citro, C. (1995). Panel surveys. Adding the fourth dimension. *Innovation: The European Journal of Social Science Research. 8*. 25-39.

Kreuter, F., & Valliant, R. (2007). A survey on survey statistics: What is done and can be done in Stata. *The Stata Journal,* 7(1), 1-21.

Kroh. M. (2014). *Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2012).* SOEP Survey Papers 177. Series D. DIW/SOEP: Berlin.

Kroh, M., Käppner, K., & Kühne, S. (2014). *Sampling. Nonresponse. and Weighting in the 2011 and 2012 Refreshment Samples J and K of the Socio-Economic Panel.* SOEP Survey Papers 260. Series C. DIW/SOEP: Berlin.

Martin, S., Zabal, A., & Rammstedt, B. (2018). *PIAAC-L Data Collection 2016: Technical Report.* GESIS Papers 2018|05. GESIS: Cologne.

Mohadjer, L., Krenzke, T., & Van de Kerckhove, W. (2013). Survey Weighting and Variance Estimation. In OECD (Ed.). *Technical Report of the Survey of Adult Skills (PIAAC)* (Chapter 15). Paris: OECD.

Perry, A., Helmschrott, S., Konradt, I., & Maehler, D.B. (2017). *User Guide for the German PIAAC scientific use file*. GESIS Papers 2017|23. Cologne: GESIS.

Rammstedt, B., Martin, S., Zabal, A., Carstensen, C., & Schupp, J. (2017). The PIAAC Longitudinal Study in Germany – Rationale and Design. *Large-scale Assessments in Education* (5:4). doi: http://dx.doi.org/10.1186/s40536-017-0040-z.

Spieß. M. (2010). Der Umgang mit fehlenden Werten. In C. Wolf & H. Best (Eds.). *Handbuch der sozialwissenschaftlichen Datenanalyse* (pp. 117-142). Wiesbaden: VS Verlag für Sozialwissenschaften.

Statistisches Bundesamt (2017). *Ergebnisse des Mikrozensus 2016.* Wiesbaden: Statistisches Bundesamt.

Steinacker, G., Wolfert, S., & Thümmel, K. (2017). *Durchführung der 3. Erhebungswelle von PIAAC-L (Kooperative längsschnittliche Weiterverfolgung der PIAAC-Studie in Deutschland). Feldbericht zur Erhebung 2016.* GESIS Papers 2017|27. GESIS: Cologne.

Yamamoto, K., Khorramdel, L., & von Davier, M. (2013). Scaling PIAAC Cognitive Data. In OECD (Ed.). *Technical Report of the Survey of Adult Skills (PIAAC)* (Chapter 17). Paris: OECD.

Zabal, A., Martin, S., Massing, N., Ackermann, D., Helmschrott, S., Barkow, I., & Rammstedt, B. (2014). *PIAAC Germany 2012. Technical report.* Münster: Waxmann.

Zabal, A., Martin, S., & Rammstedt, B. (2016). *PIAAC-L Data Collection 2014: Technical Report. Follow-up to PIAAC Germany 2012.* GESIS Papers 2016|17. GESIS: Cologne.

Appendices

Appendix A

Table A1 List of Abbreviations of Independent Variables Used in the Nonresponse Analyses

Variable	Label	Value
Education Characheristics		
Education (school degree. grouped)	Education level at the time of the interview (grouped)	
Low education	Low education level	1/0
High education	High education level	1/0
Pupil in school	Pupil in school	1/0
In training/education	In education or training	1/0
Interview Characteristics		
Interviewer change	The interviewer changed from PIAAC-L 2015 to PIAAC-L 2016	1/0
Phone contact	Number of initial contact attempts via phone	
Part. unit nonresp.	Nonresponse of eligible household member	1/0
HHmoved	Respondent moved to another address	1/0
Region and Household Characteristics		
Federal state		
Schleswig-Holstein	Schleswig-Holstein	1/0
Hamburg	Hamburg	1/0
Lower Saxony	Lower Saxony	1/0
Bremen / Bremerhaven	Bremen / Bremerhaven	1/0
North Rhine-Westph.	North Rhine-Westphalia	1/0
Hesse	Hesse	1/0
Rhineland-Palatinate	Rhineland-Palatinate	1/0
Baden-Wuertt.	Baden-Wuerttemberg	1/0
Bavaria	Bavaria	1/0
Saarland	Saarland	1/0
Berlin	Berlin	1/0
Brandenburg	Brandenburg	1/0
Saxony	Saxony	1/0
Saxony-Anhalt	Saxony-Anhalt	1/0
Thuringia	Thuringia	1/0
Housing / building type		
Farm house	Farm house	1/0
Row / duplex house	Row house or duplex (with one dwelling next to the other)	1/0
Building: 3-4 flats	Residential building containing 3 or 4 dwellings	1/0
Building: 5-8 flats	Residential building containing 5 to 8 dwellings	1/0

Variable	Label	Value
>8 flats <9stories	Residential building containing 9 or more dwellings (up to 8 stories)	1/0
>9 stories	High-rise building (9 or more stories)	1/0
Building: miss	Missing information on housing / type of building	1/0
Home owner	Owner of dwelling	1/0
Subtenant	Subtenant of dwelling	1/0
Housing area	Household is located in insecure housing area	1/0
Condition building	Condition of the building at the time of the initial contact attempt	1/0
Household size (anchor)		
One person in HH	One person in household	1/0
2 persons in HH	Two persons in household	1/0
3 persons in HH	Three persons in household	1/0
4 persons in HH	Four persons in household	1/0
Income and Employment Characteristics		
Income (quartiles)		
1q. (<1900 euros)	Monthly net household income less 1900 euros	1/0
2q. (<=3000 euros)	Monthly net household income less 3000 euros	1/0
3q. (<=4000 euros)	Monthly net household income less 4000 euros	1/0
4q. (>4000 euros)	Monthly net household income more than 4000 euros	1/0
Inc. miss.	Information on monthly net household income is missing	1/0
Labor force participation		
Not in labor force	Not participating in labor force at all	1/0
Employee	Respondent is employed	1/0
Unemployed	Respondent is unemployed	1/0
Occupational position		
Self-empl.	Self-employed (also: Working for a self-employed relative)	1/0
Family and Partnership Characteristics		
Has partner	Respondent has a steady partner	1/0
Married	Respondent is married	1/0
Divorced	Respondent is divorced	1/0
Family in HH	Family members live in househould (incl. spouse)	1/0
Has Children	Respondent has at least one child	1/0
Partner living apart	Partner of respondent is not living in same household	1/0
Education parents (school degree. grouped)	Parental education level at the time of the interview in PIAAC-L 2014	
Edu. miss. both. par.	Information on education is missing for both parents	1/0
Low education par.	Both parents have a low education level	1/0
High education par.	Both parents have a high education level	1/0
Edu. mother miss.	Information on education is missing for the mother	1/0
Edu. father miss.	Information on education is missing for the father	1/0

Variable	Label	Value
Other Characteristics		
Age (grouped)	Age at the time of the interview	
Age 19-24	19-24 years old	1/0
Age 25-34	25-34 years old	1/0
Age 45-54	45-54 years old	1/0
Age 55-69	55-69 years old	1/0
German citizenship	Respondent has German citizenship	1/0
Native speaker: Ger.	Respondents' native language is German	1/0
Migr. backgr.	Respondent has migration background	1/0
High life satisf.	Respondent reports high life satisfaction	1/0
Cognitive Skills Characteristics		
Literacy (quartiles)		
Lit 1.q.	Literacy scale score - Competence data 2015. background data 2012/14/15– first quartile	1/0
Lit 2.q.	Literacy scale score - Competence data 2015. background data 2012/14/15– second quartile	1/0
Lit 3.q.	Literacy scale score - Competence data 2015. background data 2012/14/15 – third quartile	1/0
Lit 4.q.	Literacy scale score - Competence data 2015. background data 2012/14/15 – fourth quartile	1/0
Numeracy (quartiles)		
Num 1.q.	Numeracy scale score - Competence data 2015. background data 2012/14/15 – first quartile	1/0
Num 2.q.	Numeracy scale score - Competence data 2015. background data 2012/14/15 – second quartile	1/0
Num 3.q.	Numeracy scale score - Competence data 2015. background data 2012/14/15 – third quartile	1/0
Num 4.q.	Numeracy scale score - Competence data 2015. background data 2012/14/15 – fourth quartile	1/0
Problem solving (quartiles)		
PSL q. miss	No plausible values for PSTRE in PIAAC Germany 2012	1/0
PSL 1.q.	PSTRE scale score - Competence data 2012 / background data 2012. 2014 – first quartile	1/0
PSL 2.q.	PSTRE scale score - Competence data 2012 / background data 2012. 2014 – second quartile	1/0
PSL 3.q.	PSTRE scale score - Competence data 2012 / background data 2012. 2014 – third quartile	1/0
PSL 4.q.	PSTRE scale score - Competence data 2012 / background data 2012. 2014 – fourth quartile	1/0

Appendix B

Example B1 Stata Code For Analyses With Plausible Values Based on ZA5989_Persons_15, Anchor Persons

Step 1: Declare multiple imputation of plausible values

capture drop PVPSL PVNUM PVLIT

gen PVPSL=.

gen PVNUM=.

gen PVLIT=.

capture mi unset

mi import wide, imputed(PVNUM=PVNUM1-PVNUM10 PVLIT=PVLIT1-PVLIT10) clear drop

Step 2: Declare survey design

mi svyset ID_PSU [pw=hrf_15], strat(GKPOL)

Notes: Use ID_PSU (clustering variable) and GKPOL (stratification information) from the PIAAC scientific use file, and variable hrf_15 from ZA5989_Weights_15. For longitudinal analyses, please use the longitudinal weights (see Section "Usage of Weights").

Step 3: Run analyses

- Mean level of literacy:

```
mi estimate: svy: mean PVLIT
```

- Mean level of literacy by age group:

```
mi estimate: svy: mean PVLIT, over(agegroup)
```

  Note: The variable agegroup can be derived from variable AGE_R_15, for example.

- Regression analysis of education (in years), employment experience, gender, and literacy on income:

```
mi estimate: svy: reg income YRSQUAL_15 C_Q09_15 gender PVLIT
```

  Note: The dependent variable income is the logarithm of monthly net houshehold income (hinc_15). Use the variable gender from ZA5989_Registry.

Example B2 Stata Code For Analyses Without Plausible Values Based on ZA5989_Persons_15, Anchor Persons

Step 1: not required

Step 2: Declare survey design

```
svyset ID_PSU [pw=hrf_15], strat(GKPOL)
```

Step 3: Run analyses

- Mean level of monthly net household income:

```
svy: mean hinc_15
```

- Mean level of montly net household income by age group:

```
svy: mean hinc_15, over(agegroup)
```

- Regression analysis of education (in years), eployment experience, and gender on income:

```
svy: reg income YRSQUAL_15 C_Q09_15 gender
```