The Impact of Weighting on Secondary Outcomes in PIAAC Germany

Susanne Helmschrott, University of Mannheim

SFB 884, “The Political Economy of Reforms”

If a survey is subject to nonresponse, not all survey estimates are necessarily affected by nonresponse bias in the same way. In fact, the magnitude of bias depends on the strength of the association between the survey variable and response propensities of the sample members. When computing weights, survey practitioners have thus to choose (a) “central” study outcome variable(s) for which the reduction of bias is of utmost importance. However, surveys in the social sciences often cover a multitude of topics and “secondary” outcome variables of a survey might not benefit from the weighting adjustments to the same extent as central variables. As a result, researchers using secondary variables risk reporting biased estimates, even when applying the published survey weights. Based on the example of PIAAC, the Programme for the International Assessment of Adult Competencies, this paper shows that weighting adjustments which are effective for central outcome variables can fail to reduce bias in secondary outcome variables.

1. Introduction

During the survey process, several sources of error may occur that restrict the possibility to infer from the data collected to the target population. For example, if certain groups of sampled persons systematically differ in their willingness to participate, nonresponse bias in the estimates may occur (Groves, 2006; Groves & Lyberg, 2010). In order to decrease nonresponse bias, survey methodologists use different correction techniques, notably weighting (e.g. Särndal, Swensson, & Wretman, 1992) and imputation methods (e.g. Rubin, 1987). However, weighting techniques are constrained by the fact that the magnitude of bias differs across survey estimates. In fact, nonresponse bias in the respondent mean does not only depend on the response rate, but also on the strength of the association between the response propensity and the variables measured in the survey (Bethlehem, 2002). Hence, survey weights are only effective when the variables used for their calculation are highly associated with both the survey outcome and response propensity (Little & Vartivarian, 2005). This poses a challenge to survey methodologists as it is often difficult to define for which survey estimate the weights should predominantly reduce bias. In fact, social science surveys often cover several topics, such as politics, well-being, religion, health, or media usage (e.g. ESS, 2016). Hence, it is possible that the computed weights are effective in reducing nonresponse bias in the estimate which has been defined as a central study outcome, but not in others. Researchers using the survey weights supplied in the published data sets may thus unintentionally produce biased estimates when their research interest lies outside the central study outcome. While the general effectiveness of weighting procedures has often been addressed in the literature (e.g. Loosveldt & Sonck, 2008; Stoop, Billiet, Koch, & Fitzgerald, 2010), a closer look at the impact on “secondary outcomes” has been neglected yet.

2. Research question and data

In this paper, I address the question whether survey weights designed to reduce bias[1] in central study outcomes are able to reduce bias in secondary study outcomes as well, using the example of the German PIAAC data. PIAAC, the "Programme for the Assessment of Adult Competencies" is an OECD-led survey that has been implemented in 24 countries in 2011/2012 for the first time. It aims at assessing the level and distribution of skills among the adult population aged 16 to 65, the development and use of skills, and their social and economic benefits (OECD, 2013). In this face-to-face-study, respondents were first asked to complete a questionnaire on their educational background, work-related topics such as the use of skills at work, their income, health, and volunteering activities. In the following, they were asked to complete a proficiency test in two of the three domains literacy, numeracy and problem solving in technology-rich environments (OECD, 2012). While the core interest of PIAAC is the skills assessment, it is only the extensive questionnaire that allows policy makers and researchers to gather insights into how successful labour market, education and training policies are in fostering skills and what economic and social impact the differences in skill levels have (OECD, 2013). Given this design, PIAAC data can be divided into a central study outcome, the proficiency[2] scores, and a multitude of "secondary", mostly work-related outcomes collected in the questionnaire. This division makes PIAAC an excellent data source to investigate the differences in the effectiveness of weighting across central and secondary estimates.

3. Analyses

a. Metrics

In this paper, two metrics are used to assess the impact of weighting on study estimates. First, I compute the *average relative bias* ($RB_{\emptyset}$) of both central and secondary study outcomes and compare changes in both groups after weighting. For this metric, I first compute the relative bias for each category k of the variables used in the analysis[3]. The relative bias per category k is defined as the absolute difference between the respondent proportion y_{rk} and the corresponding population[4] proportion y_{pk}, divided by the population proportion in that category. Following, I average the results by summing up the relative bias of the variable categories and dividing the sum by the number of variable categories k. Furthermore, this value is multiplied by 100, enabling the interpretation of the bias as percentage of the population value:

$$RB_{\emptyset} = \sum_{i=1}^{k} 100 * \left|\frac{y_{rk}-y_{pk}}{y_{pk}}\right| / k \quad \text{(adapted from Groves, 2006)}$$

[1] "Bias" stands for nonresponse bias throughout the paper. Bias due to noncoverage is considered negligible in PIAAC, as only an estimated 2,5% of the target population was not included in the sample frame (Zabal et al., 2014).

[2] Throughout the paper, "proficiency" means proficiency in literacy.

[3] Only categorical and ordinal variables are used. For example, variable categories for gender are male and female.

[4] The population data is taken from the German Microcensus, which is a mandatory survey of a representative sample of 1% of households in Germany (DESTATIS, 2016).

Compared to the absolute bias, i.e. the net difference between y_{rk} and y_{pk}, the relative bias has the advantage to account for different sizes of the proportions across variable categories. In addition, taking the average of the relative biases (instead of the sum) allows for comparisons between central and secondary outcomes, as both groups include a different number of variables. As a second key figure, I assess the *statistically significant differences* between the survey and the population data by computing t-tests for each variable category. Since the number of variable categories in the central outcomes is distinct from the number in the secondary outcomes, I report the *share of statistically significant differences* out of all categories for each group. In order to assess the impact of weighting on the estimates, changes in this share will be examined after weighting.

b. Variables

For PIAAC's central study outcome, the proficiency scores, no benchmark data exist. Hence, proxy variables had to be chosen for which a) population (i.e. Microcensus) data exist, b) that are significantly correlated with proficiency and c) that do not classify as secondary outcome. The variables satisfying these conditions are age, citizenship, the level of education (with $p<0.001$ for each correlation), gender, the size of the municipality the respondent lives in (with $p<0.01$ each), as well as household size ($p<0.05$). For the secondary outcomes, I chose seven basic work-related variables, for which Microcensus data are available. These are the respondents' ISCED[5] level of education, employment status, occupation, the industry and size of the business they work in, the work contract they hold, as well as their usual working hours.

c. Weights

Weighting in PIAAC included several weighting steps. First, design weights were created, which were then adjusted for unknown eligibility, followed by an adjustment for nonresponse[6] and a calibration to population totals, yielding a final weight. The calibration technique employed in PIAAC Germany was post-stratification. Each weighting step was computed on data weighted with the respective preceding weight. Consequently, the final weights, which are intended for use in estimation and analysis, are the result of all weighting steps conducted (Mohadjer, Krenzke, Van de Kerckhove, & Hsu, 2013). In order to reduce bias in the central study outcome, i.e. proficiency, the international PIAAC consortium required that the nonresponse adjustment variables had to be related to both response status and proficiency[7] (OECD, 2014). In PIAAC Germany, these were age, citizenship and municipality size. For post-stratification, age and gender were required. For an additional reduction of bias in proficiency, the German national team further chose to add the level of education and a regional variable to the weighting adjustments (Zabal et al., 2014).

[5] International Standard Classification of Education.
[6] This weighting step was conducted separately for literacy- and nonliteracy-related nonrespondents (Mohadjer et al., 2013).
[7] At the time the weights were computed, the proficiency scores had not been available yet. Hence, for the choice of appropriate nonresponse weighting variables, the highest level of education was used as a proxy (Zabal et al., 2014).

4. Results

As Figures 1 and 2 show, the PIAAC weights achieve only minor reductions in bias in the secondary outcomes, whereas bias in the central outcomes is strongly reduced. Indeed, for the secondary outcomes, the average relative bias drops only from an average of 18.3 to a score of 17.8 in the final weighted data. Similarly, the share of significant differences between the respondent and the Microcensus data decreases only slightly. Indeed, 65.5% of the variable categories comprised in the secondary outcomes (i.e. 19 out of 29) are significantly different to the Microcensus data when unknown eligibility weights were used. This share drops only to 55.2% (i.e. 16) in the final weighted data, yielding a reduction of 10.3 percentage points. Contrary to that, the average relative bias in the central study outcomes drops from an average of 11.7 when unknown eligibility weights are used to 7 after the application of final weights. This is a reduction of 4.7 points. The analysis of changes in significant differences between PIAAC and Microcensus data substantiates this finding. In fact, their share drops from 68.2% of the variable categories comprised in the central outcomes (i.e. 15 out of 22) to 27.3% (6) after applying the final weight, yielding a reduction of 40.9 percentage points.

Figure 1. Changes in average relative bias

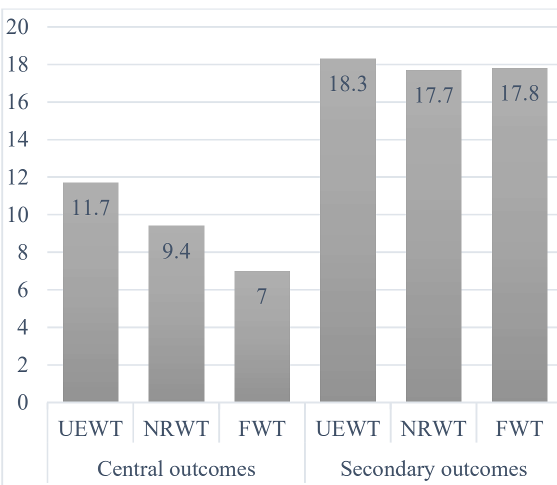

Figure 2. Changes in significant differences

Legend: UEWT: Unknown eligibility weight; NRWT: Nonresponse weight; FWT: Final weight

When comparing the differential impact of the nonresponse and final weights, we can observe, particularly for the central outcomes, a stronger decrease of bias when using the latter compared to the former. For example, in the group of the central outcomes, the share of significant differences decreases only about 4.6 percentage points when nonresponse weights are used to 40.9 percentage points after applying the final weights. This observation illustrates the multiplicative nature of the PIAAC weights. In fact, each weighting step serves a different purpose, and each weight is computed on data weighted with the previous weight, yielding a final weight with a maximum reduction in bias. The nonresponse weights exclusively aim at reducing nonresponse bias. Hence, in this weighting step, the respondent data is adjusted to the eligible sample data. As the eligible sample is subject to sampling error itself, population totals are not perfectly met yet. It is only the post-stratification that yields a major reduction of bias when comparing the survey distributions to population totals. It has to be noted, however, that the strong effect of the final weights on the central study outcomes is to a large part due to the fact that some of the analysis variables were used for post-stratification as well.

5. Conclusion and discussion

In this paper, I show using the example of PIAAC Germany, that weighting that is designed to reduce bias in a specific (set of) outcome(s), can fail to reduce bias in further study outcomes. Researchers studying those secondary outcomes may thus unintentionally report biased estimates, even when using the published survey weights.

Survey methodologist should not accept this problem as an inevitable trade-off inherent to weighting techniques. Instead, the impact of weighting on all study estimates should routinely be examined. If necessary, alternative weights could be provided, that focus on reducing bias in a pre-defined set of secondary outcomes. If imputation yields less biased data, researchers could be provided with a data set including imputed values. In the case of the first round of PIAAC in Germany, no major data updates are to be expected, as the project has been completed in 2014. However, a possible solution for PIAAC Germany could be to provide researchers with a code allowing them to calculate an alternative weight themselves.

As no population data for PIAAC's central study outcome, the proficiency scores, is available, proxy variables are used in this paper. Given that nearly all variables qualifying as proxies are used in one of PIAAC's weighting steps as well, the finding that bias in the central study outcomes is reduced after weighting comes as no surprise. However, if population data for the proficiency scores were available, it could be expected that the analyses would yield similar results. As pointed out, all variables used in this analysis (and in the weighting procedures as well) were highly significantly related to proficiency. The stronger the relationship between weighting variables and study outcome, the stronger the reduction in the bias of this estimate (Bethlehem, 2002; Little & Vartivarian, 2005). Further research is needed to substantiate a need for action for data users focussing on non-central study outcomes. For example, in the case of PIAAC, bi- and multivariate analyses with PIAAC's secondary outcomes could be examined to explore whether bias in point estimates impairs estimates of relationships as well. If this were the case, it would be crucial to experiment with alternative weights and scrutinize whether they represent a true improvement to the published weights.

Points for discussion

- What do you think are the practical implications for large scale survey projects? Should this problem be addressed by survey practitioners and how?
- Given that the analyses of the central outcomes are affected by the fact that the analyses variables are used in the computation of the weights as well, do you think that results are robust? Do you have any ideas of improvement?

Literature

Bethlehem, J. G. (2002). Weighting Nonresponse Adjustments Based on Auxiliary Information. In R. M. Groves, D. A. Dillman, J. Etlinge, & R. J. A. Little (Eds.), *Survey Nonresponse* (pp. 275 – 288). New York: Wiley & Sons.

European Social Survey. (2016). ESS Themes. Retrieved July 7, 2016, from http://www.europeansocialsurvey.org/data/module-index.html

Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, *70*(5), 646–675.

Groves, R. M., & Lyberg, L. (2010). Total Survey Error: Past, Present and Future. *Public Opinion Quarterly*, *74*(5), 849–879.

Little, R., & Vartivarian, S. (2005). Does Weighting for Nonresponse Increase the Variance of Survey Means? *Survey Methodology*, *31*, 161–168.

Loosveldt, G., & Sonck, N. (2008). An Evaluation of the Weighting Procedures for an Online Access Panel Survey. *Survey Research Methods*, *2*(2), 93–105.

Mohadjer, L., Krenzke, T., Van de Kerckhove, W., & Hsu, V. (2013). Survey Weighting and Variance Estimation. In OECD (Ed.), *Technical Report of the Survey of Adult Skills (PIAAC)* (pp. 1–30). Paris: OECD. Retrieved from http://www.oecd.org/site/piaac/_Technical Report_17OCT13.pdf

OECD. (2012). *Literacy, Numeracy, and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*. OECD Publishing.

OECD. (2013). *OECD Skills Outlook 2013: First results from the Survey of Adult Skills*. OECD Publishing.

OECD. (2014). *PIAAC Technical Standards and Guidelines*. Retrieved from http://www.oecd.org/skills/piaac/PIAAC-NPM(2014_06)PIAAC_Technical_Standards_and_Guidelines.pdf

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley & Sons.

Särndal, C.-E., Swensson, B., & Wretman, J. (1992). *Model Assisted Survey Sampling*. New York: Springer Verlag.

Stoop, I., Billiet, J., Koch, A., & Fitzgerald, R. (2010). *Improving Survey Response*. New York: Wiley & Sons.

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

Annex

Table 1. Comparison of central study estimates from PIAAC and Microcensus 2011

	PIAAC respondent data – Unknown eligibility weight				PIAAC respondent data – Nonresponse weight				PIAAC respondent data- Final weight				MC 11
	%	*SE*	*Relative bias*	*p-value t-test*	*%*	*SE*	*Relative bias*	*p-value t-test*	*%*	*SE**	*Relative bias*	*p-value t-test*	*%*
Age													
16-25	19.6	0.56	15.3	0.000	16.3	0.35	4.1	0.003	17.5	0.48	2.9	0.713	17.0
26-35	18.2	0.64	0.0	0.859	18.8	0.43	3.3	0.083	18.0	0.57	1.1	0.910	18.2
36-45	21.5	0.51	3.6	0.002	22.2	0.39	0.4	0.020	23.0	0.60	3.1	0.823	22.3
46-55	23.6	0.58	2.1	0.984	25.0	0.38	3.7	0.001	23.8	0.58	1.2	0.778	24.1
56-65	17.1	0.58	7.1	0.162	17.7	0.41	3.8	0.699	17.7	0.56	3.8	0.800	18.4
Gender													
male	48.8	0.66	2.4	0.017	48.5	0.68	3.0	0.007	50.5	0.71	1.0	0.879	50.0
female	51.2	0.66	2.6	0.017	51.5	0.68	3.2	0.007	49.5	0.71	0.8	0.879	49.9
Citizenship													
German	93.3	0.44	3.2	0.000	91.8	0.38	1.5	0.000	91.4	0.41	1.1	0.000	90.4
Not German	6.7	0.44	30.2	0.000	8.2	0.38	14.6	0.000	8.6	0.41	10.4	0.000	9.6
Highest school leaving degree													
Low	25.6	0.66	21.7	0.000	25.9	0.65	20.8	0.000	31.6	1.10	3.4	0.389	32.7
Medium	35.9	0.87	11.1	0.016	35.8	0.84	10.8	0.018	34.4	1.10	6.5	0.594	32.3
High	35.7	0.83	14.1	0.000	35.9	0.82	14.7	0.000	30.7	0.77	1.9	0.677	31.3
Pupil	2.8	0.20	22.2	0.022	2.4	0.17	33.3	0.000	3.3	0.27	8.3	0.911	3.6
Municipality size													
1 – 4.999 inhabitants	16.5	2.15	11.5	0.437	15.6	2.01	5.4	0.686	15.9	1.96	7.4	0.517	14.8
5.000 – 49.999 inh.	46.7	2.96	5.7	0.392	45.3	2.98	2.5	0.725	45.6	3.00	3.2	0.650	44.2
50.000 – 99.999 inh.	9.3	1.69	4.5	0.797	9.3	1.70	4.5	0.814	9.2	1.67	3.4	0.843	8.9
100.000 – 499.999	14.8	2.10	3.9	0.758	15.5	2.12	0.6	0.960	15.4	2.11	0.0	0.983	15.4
500.000 – 99.999.999	12.7	1.56	24.0	0.012	14.3	1.76	14.4	0.181	13.9	1.70	16.8	0.108	16.7

	PIAAC respondent data – Unknown eligibility weight				PIAAC respondent data – Nonresponse weight				PIAAC respondent data- Final weight				MC 11
	%	*SE*	*Relative bias*	*p-value t-test*	*%*	*SE*	*Relative bias*	*p-value t-test*	*%*	*SE**	*Relative bias*	*p-value t-test*	*%*
Household size													
1	14.4	0.48	25.8	0.000	14.8	0.51	23.7	0.000	14.5	0.51	25.3	0.000	19.4
2	31.4	0.70	2.2	0.047	31.8	0.71	0.9	0.184	31.4	0.69	2.2	0.044	32.1
3 to 4	43.5	0.73	6.9	0.000	43.2	0.74	6.1	0.001	43.4	0.72	6.6	0.000	40.7
5 or more	10.7	0.47	37.2	0.000	10.2	0.04	30.8	0.000	10.7	0.47	37.2	0.000	7.8
Average rel. bias			**11.7**				**9.4**				**7**		
Red. compared to UEW							***-2.3***				***-4.7***		
Number of significant differences (p<0.05)				**15**				**14**				**6**	
Share of sig. diff.**				**68.2%**				**63.6%**				**27.3%**	
Reduction in pp. compared to UEW***								***-4.6 pp.***				***-40.9 pp.***	

* For the estimation of standard errors of the final weighted respondent data of age, gender and education, Taylor Series Linearization was used instead of PIAAC's 80 replicate weights (for details, see Mohadjer, Krenzke, Van de Kerckhove & Hsu, 2013). These variables were used for post-stratifying the data. During the final weighting step, each subsample was recalibrated to population totals. Hence, when using the replicate weights, there is no variation across the subsamples in these variables, i.e. standard errors are zero.

** Percentage of significant differences (p<0.05) of the total number of variable categories.

*** Reduction of significant differences compared to significant differences found in unknown eligibility weighted data in percentage points.

Table 2. Comparison of secondary study estimates from PIAAC and Microcensus 2011

	PIAAC respondent data – Unknown eligibility weight				PIAAC respondent data – Nonresponse weight				PIAAC respondent data- Final weight				MC 11
	%	*SE*	*Relative bias*	*p-value t-test*	*%*	*SE*	*Relative bias*	*p-value t-test*	*%*	*SE*	*Relative bias*	*p-value t-test*	*%*
ISCED97													
Low education	15.8	0.51	12,2	0.000	14.7	0.50	18,3	0.000	17.3	0.48	3,9	0.412	18.0
Medium education	52.2	0.71	9,5	0.000	52.2	0.73	9,5	0.000	53.2	0.71	7,8	0.000	57.7
High education	32.0	0.75	31,7	0.000	33.1	0.76	36,2	0.000	29.5	0.55	21,4	0.000	24.3
Employment status													
Employed	76.1	0.63	4.2	0.000	76.6	0.63	4.9	0.000	75.4	0.57	3.3	0.000	73.0
Unemployed	4.0	0.31	13.0	0.063	4.0	0.31	13.0	0.063	4.2	0.33	8.7	0.220	4.6
Out of the labor force	19.9	0.59	11.2	0.000	19.4	0.59	13.4	0.000	20.4	0.56	8.9	0.000	22.4
Occupation*													
Managers	4.5	0.35	10.0	0.135	4.5	0.33	10.0	0.153	4.3	0.34	14.0	0.037	5.0
Professionals	19.1	0.70	9.1	0.031	19.7	0.70	12.6	0.003	17.2	0.48	1.7	0.400	17.5
Technicians and assoc. profess.	18.0	0.61	13.5	0.000	18.0	0.59	13.5	0.000	17.5	0.57	15.9	0.000	20.8
Clerical support workers	12.1	0.54	2.4	0.588	12.1	0.56	2.4	0.555	11.8	0.54	4.8	0.246	12.4
Service and sales workers	18.1	0.66	18.3	0.000	17.7	0.65	15.7	0.001	18.3	0.65	19.6	0.000	15.3
Skilled agricultural forestry and fishery workers	1.8	0.252	20.0	0.315	1.7	0.24	13.3	0.366	1.9	0.28	26.7	0.174	1.5
Craft and related trades	12.5	0.64	1.6	0.766	12.1	0.64	4.7	0.386	13.4	0.65	5.5	0.277	12.7
Plant and machine operators and assemblers	7.2	0.46	12.5	0.089	7.2	0.46	12.5	0.077	8.1	0.47	26.6	0.001	6.4
Elementary Occupations	6.7	0.41	20.2	0.000	6.9	0.43	17.9	0.001	7.5	0.46	10.7	0.061	8.4
Economic sector													
Agriculture. Forestry and Fishing (Nace A)	1.7	0.25	6.2	0.455	1.6	0.24	0.0	0.521	1.8	0.27	12.5	0.330	1.6
Industry (Nace B-F)	28.7	0.77	0.7	0.785	28.5	0.80	0.0	0.959	29.8	0.78	4.6	0.106	28.5
Services (Nace G-U)	69.6	0.82	0.4	0.720	69.9	0.84	0.0	0.987	68.4	0.82	2.1	0.082	69.9

	PIAAC respondent data – Unknown eligibility weight				PIAAC respondent data – Nonresponse weight				PIAAC respondent data- Final weight				MC 11
	%	*SE*	*Relative bias*	*p-value t-test*	*%*	*SE*	*Relative bias*	*p-value t-test*	*%*	*SE*	*Relative bias*	*p-value t-test*	*%*
Work contract*													
An indefinite contract	80.8	0.73	4.2	0.000	80.5	0.69	4.5	0.010	82.1	0.64	2.6	0.002	84.3
A fixed term contract	12.8	0.59	36.2	0.000	12.3	0.58	30.9	0.000	12.5	0.57	33.0	0.000	9.4
A temporary employment agency contract	1.1	0.17	50.0	0.000	1.1	0.17	50.0	0.000	1.1	0.19	50.0	0.000	2.2
An apprenticeship or other training scheme	5.3	0.47	29.3	0.016	4.3	0.42	4.9	0.523	4.3	0.40	4.9	0.191	4.1
Hours per week working (usually)													
1 to 18 hours	12.7	0.55	10.4	0.032	12.3	0.53	7.0	0.115	12.4	0.53	7.8	0.115	11.5
19 to 42 hours	60.8	0.89	17.4	0.000	60.8	0.87	17.4	0.000	60.7	0.91	17.5	0.000	73.6
43 to 60 hours	23.9	0.81	75.7	0.000	24.1	0.83	77.2	0.000	24.1	0.83	77.2	0.000	13.6
More than 60 hours	2.6	0.29	85.7	0.000	2.8	0.29	100.0	0.000	2.8	0.31	100.0	0.000	1.4
Size of business													
1 to 10	24.1	0.74	10.1	0.000	24.0	0.77	10.4	0.000	24.7	0.79	7.8	0.000	26.8
11 to 50	27.0	0.71	14.4	0.000	26.8	0.72	13.6	0.000	26.9	0.77	14.0	0.000	23.6
More than 50	48.9	0.88	1.4	0.492	49.2	0.91	0.8	0.712	48.4	0.94	2.4	0.250	49.6
Average relative bias			**18.3**				**17.7**				**17.8**		
Red. compared to UEW							***-0.6***				***-0.5***		
Number of significant differences(p<0.05)				**19**				**17**				**16**	
Share of sig. differences*				**65.5%**				**58.6%**				**55.2%**	
Reduction in pp. compared to UEW*								***-6.9 pp.***				***-10.3 pp.***	

* Employees who have a contract, without "other" and short-term contract

** Percentage of significant differences (p<0.05) of the total number of variable categories.

**** Reduction of percentage of significant differences compared to significant differences found in unknown eligibility weighted data.