

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Reinikainen, Jaakko; Tolonen, Hanna; Borodulin, Katja; Härkänen, Tommi; Jousilahti, Pekka; Karvanen, Juha; Koskinen, Seppo; Kuulasmaa, Kari; Männistö, Satu; Rissanen, Harri; Vartiainen, Erkki

Title: Participation rates by educational levels have diverged during 25 years in Finnish health examination surveys

Year: 2018

Version: Accepted version (Final draft)

Copyright: © The Authors 2017. Published by Oxford University Press on behalf of the Europe

Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Reinikainen, J., Tolonen, H., Borodulin, K., Härkänen, T., Jousilahti, P., Karvanen, J., Koskinen, S., Kuulasmaa, K., Männistö, S., Rissanen, H., & Vartiainen, E. (2018). Participation rates by educational levels have diverged during 25 years in Finnish health examination surveys. *European Journal of Public Health*, 28(2), 237-243. <https://doi.org/10.1093/eurpub/ckx151>

Participation rates by educational levels have diverged during 25 years in Finnish health examination surveys

Jaakko Reinikainen¹, Hanna Tolonen¹, Katja Borodulin¹, Tommi Härkänen¹, Pekka Jousilahti¹, Juha Karvanen², Seppo Koskinen¹, Kari Kuulasmaa¹, Satu Männistö¹, Harri Rissanen¹, Erkki Vartiainen¹

¹Department of Public Health Solutions, National Institute for Health and Welfare, P.O. Box 30, FI-00271, Helsinki, Finland

² Department of Mathematics and Statistics, University of Jyväskylä, P.O.Box 35 (MaD), FI-40014 University of Jyväskylä, Finland

Corresponding author: Jaakko Reinikainen, Email: jaakko.reinikainen@thl.fi, Address: P.O. Box 30, FI-00271 Helsinki, Finland, Tel.: +358 29 524 7420

Abstract

Background: Declining participation rates in health examination surveys may impair the representativeness of surveys and introduce bias into the comparison of results between population groups if participation rates differ between them. Changes in the characteristics of non-participants over time may also limit comparability with earlier surveys.

Methods: We studied the association of socio-economic position with participation, and its changes over the past 25 years. Occupational class and educational level are used as indicators of socio-economic position. Data from six cross-sectional FINRISK surveys conducted between 1987 and 2012 in Finland were linked to national administrative registers, which allowed investigation of the differences between survey participants and non-participants.

Results: Our results show that individuals with low occupational class or low level of education were less likely to participate than individuals with high occupational class or high level of education. Participation rates decreased in all subgroups of the population but the decline was fastest among those with low level of education.

Conclusions: The differences in participation rates must be taken into account to avoid biased estimates because socio-economic position has also been shown to be strongly related to health, health behaviour and biological risk factors. Particular attention should be paid to the recruitment of the less-educated population groups.

Keywords: non-participation, health examination survey, socio-economic position, educational level, occupational class

Introduction

Health examination surveys (HESs) are conducted to obtain information about health and health related factors in the population and their trends over time. Therefore it is important that the survey results represent well the population ¹. Potential sources of bias may relate to the sampling frame, sampling, participation rates and the survey measurements ². In this study we will consider participation rates.

Over the past decades, the participation rates in HESs have decreased ³. This leads to declining representativeness if participants and non-participants differ from each other. Thus, the interest towards the reasons and consequences of non-participation has grown ⁴⁻⁶. Non-participation increases uncertainty of estimates and, more seriously, may introduce selection bias if the set of participants is not representative. For example, health behaviour and health status have been seen to differ substantially between participants and non-participants ⁷⁻⁹.

Previous research has helped to assess the impact of survey non-participation on the results and to take it into account in estimation ¹⁰⁻¹². Information about the effect of recruitment methods on participation may be used to improve the implementation of future surveys ^{13,14}. Although the predictors of non-participation have varied across studies and populations, low occupational class and educational level have usually been associated with a low participation rate ^{4,15,16}. Knowledge on the characteristics of non-participation can help us to develop more effective, subgroup-specific recruitment methods and statistical methods to assess and correct the effect of non-participation in the estimation of health indicators.

Planning the allocation of health care resources and evaluating health interventions or prevention programs require long-term health data. The decreasing participation rates may complicate the utilization of the collected data. If we do not know how well the population is represented by survey participants, we cannot know whether the observed changes in health indicators are due to real changes in the population or changes in the representativeness of the survey results.

We aim to provide new information on long-term changes in participation rates and their association with occupational class and educational level in Finland between 1987 and 2012. We will also explore the effects of sex and age and their interactions to deepen our understanding of non-participation in health examination

surveys and its relationship with socio-economic position. Also predictions of participation probabilities until the year 2022 will be presented.

Methods

Health examination surveys

Data from the National FINRISK Surveys ¹⁷ conducted every five years between 1987 and 2012, were used. These health examination surveys were conducted in selected geographical areas of Finland. In each study year an independent stratified random sample was taken. In 1987 and 1992, the study population covered the age range from 25 to 64 years. From 1997 onwards, also the age group 65-74 years was included. The sample sizes varied between 7,927 and 13,498 and the total sample size from the six cross-sectional surveys was 62,856. We restricted the data to 25-64-year-olds, the age range covered by each survey. After this restriction, we considered a total of 51,857 sampled individuals of whom 35,224 (68%) participated in the clinical measurements, in our analysis with combined data from all the survey years. The overall participation rate declined from 79% in year 1987 to 56% in year 2012.

The surveys included questionnaires, clinical measurements and collection of biological samples. The invitees were asked to come to the examination sites for the clinical examinations. Immigrants may be under-represented among participants as the questionnaires were only in Finnish in 1987 – 1997. Swedish was added in 2002 and English in 2012. The surveys received ethical approvals from the relevant ethics committees and were conducted in accordance with the Declaration of Helsinki.

Register data

Information on sex, age and place of residence was obtained from the sampling frame, the National Population Information System. Occupational class and educational level in the year of the survey were obtained for both survey participants and non-participants from Statistics Finland ^{18,19} through record linkage carried out using personal identification codes, which are unique for every resident of Finland.

The register data on occupational class included eight categories ¹⁸: entrepreneurs including farmers, upper white-collar workers, lower white-collar workers, blue-collar workers, students, retirees, others and unknowns. Household workers who worked at home were classified according to the head of the household (the person who is mainly responsible of the living of the household). Unemployed persons whose

unemployment had lasted less than six months, but had earlier been employed, were classified according to the previous occupation. Those who had been unemployed at least six months were classified into others. For our analyses, we reclassified occupational class into three categories: 1) entrepreneurs and white-collar workers, 2) blue-collar workers and 3) others. The main interest related to occupational class was to contrast entrepreneurs and white collar workers with blue collar workers. Because other occupational groups were comparably small, we categorized them into the same class. The proportion of unknown occupational class ranged from 2% (in 1997) to 9% (in 1987).

Educational levels were defined using the International Standard Classification of Education (ISCED 2011)²⁰. Levels of education were classified into three categories: 1) high level (tertiary education, ISCED levels 5-8), 2) middle level (secondary education, ISCED levels 3-4) and 3) low level (basic education or less or unknown, ISCED levels 0-2).

Those with primary school education or less could not be separated from those with unknown educational level in the register data. This means that we cannot know how many of those categorized into the lowest educational level have actually unknown education which should be in an upper category. However, the register data on education can be considered reliable when they are compared with interview data²¹.

Statistical methods

We defined participation rate as the number of participants divided by the eligible sample size (excluding those who had died or moved away from survey area between sample selection and survey fieldwork). Participation rates were calculated for sub-groups by occupational class, educational level, sex and survey year. Linear trends of participation rates over time were estimated using individual-level data and quasi-binomial model with identity link function.

Logistic regression model was applied to analyse the association of occupational class, educational level, sex, age and survey year with participation. Geographical area was also included in the model to adjust for the possible geographical variations. Survey year was treated as a continuous variable. The interaction terms that were negligible according to Akaike information criterion²² were omitted from the model. Non-linear

effects were taken into account by using restricted cubic splines ²³. Analyses were carried out using the R statistical software version 3.4.0 ²⁴.

Results

Participation decreased in all occupational classes and on all levels of education for both men and women (Fig. 1). In each survey year, entrepreneurs and white-collar and highly educated individuals had the highest participation rates. This means that they were over-represented among the participants. The participation rates were lowest among 'others' who include, e.g. students and retirees, and among those with low educational level, so these groups were under-represented among the participants. In 1987, the differences between occupational classes and educational levels were rather small, but especially the differences between educational levels increased rapidly over the survey years. Overall, women had higher participation rates than men. Between 1987 and 2012, men's participation declined from 75% to 51% and women's participation from 82% to 60%, so the difference in participation rates between men and women increased.

The unadjusted linear time-trends in participation are presented in Supplementary table 1 as average annual percentage point changes in participation rates. Between the occupational classes, the differences in the trends were negligible. For educational levels, the decline in participation rate was fastest among those with low educational level among both men and women. The average annual changes in participation were -0.9 and -0.7 %-points among highly educated men and women, respectively, and -1.3 and -1.4 %-points among those with low education, respectively.

Distributions of occupational class and level of education changed notably over 1987-2012 (Supplementary table 2). The proportion of entrepreneurs and white-collar increased from 43% to 51%, whereas the proportion of others decreased from 36% to 28% in the cross-sectional samples. The proportion of blue-collar remained quite stable (change from 21% in 1987 to 20% in 2012).

A more prominent change occurred in the distribution of educational levels. The percentage of persons with a high level of education doubled from 18% to 38%. A large change was also observed in the proportion of middle educational level, which increased by half (from 31% to 46%). Correspondingly, the proportion of low educational level decreased from 51% to 16%.

The estimated associations of occupational class, level of education, sex, age and year and their interactions with participation are presented in Table 1. The model was also adjusted for geographical area and an

interaction between area and age as well as area and year. The results show that the association of occupational class with participation varied between the educational levels. High educational level was positively associated with participation. Women and older people were clearly more likely to participate. It is worth noting that the main effects cannot be interpreted to represent the overall effects due to interactions with other variables, so the main effects refer to the effects on the reference levels of the other variables.

Fig. 2 illustrates the interactions including occupational class or educational level to ease their interpretation. The lower the educational level, the greater is the difference in participation of blue-collar workers compared with others. The difference in participation of entrepreneurs and white-collar workers compared to blue-collar workers is greater among men than women. The educational level differences in participation are greater among younger people than older people. Differences in participation between others and the remaining two occupational classes are large in younger ages and there are virtually no differences among individuals aged 60 years or over. This may be explained by the fact that younger individuals are more often students and less often retirees than older individuals.

There are also some differences in the effect of survey year in different occupational classes and educational levels. For low educational level the effect of survey year is greater than for those with high education. The model did not include the interaction between year and occupational class, which means that the effect of survey year is similar in each occupational class.

We used the model also to predict the probability of participation for the next surveys to be conducted in 2017 and 2022. Table 2 shows the predicted participation probabilities for four example persons having different sex, age, occupational class and educational level combinations. If the observed trends will continue and other things are held constant, the probabilities will naturally decrease in future. There are likely to be notable differences in participation rates between population groups. For instance, the predicted probabilities for a 25 years old man with low occupational class and middle educational level are less than half of those for a 50 years old woman with high occupational class and educational level.

Discussion

This study shows that those with high educational level were over-represented and those with low educational level were under-represented among participants of Finnish health examination surveys between 1987 and 2012, and this imbalance has increased over time (Fig. 2). Also high occupational class was over-represented and low occupational class was under-represented among participants. The declining participation rates increase the risk of selection bias, but this study does not provide information about the magnitude of the potential bias in prevalence estimates.

Participation rates decreased in all occupational classes and levels of education over time for both men and women and the decline was fastest for those with low level of education. At the same time, the proportion of persons with a low level of education decreased and the proportion of persons with a high level of education increased considerably in the Finnish population (Supplementary table 2). Nevertheless, the overall participation rates declined. It is also worth noting that people with low education in 1987 and 2012 are quite different population groups, so the change in participation does not mean that the behaviour of those with low educational level in 1987 would have changed.

Old age was also seen to predict participation (Fig. 2). We did not study the effect of birth cohorts here, but it may explain at least a part of the observed age effect. Within each survey, the young and the old individuals have had specific experiences in different birth cohorts, which may affect their participation in HESs.

The observed positive associations of occupational class and educational level with the participation rate are in concordance with many previous studies. Low socio-economic position has been found to associate with low participation rates in different countries in HESs and health questionnaire surveys. These studies include, for instance, a questionnaire on psychosocial health of the Finnish working-aged population ²⁵, a health survey among elderly Australians ²⁶, a US survey on drinking and intimate partner violence ²⁷ and Danish ²⁸ and Dutch ²⁹ HESs.

Wide socio-economic health inequalities have been reported in many countries, regardless of whether socio-economic position is measured by educational level, occupational class or income ^{30,31}. Thus, socio-economic differences in participation can lead to bias in the population-level estimates of health status and health

behaviours. For example, occupational class has been shown to be strongly related to the risks of total and cardiovascular disease mortality³². In addition, according to previous studies, high level of education and high income are positively associated with healthy dietary habits³³ and low socio-economic status is related to greater cigarette smoking³⁴. A Dutch study showed moreover that non-response leads to bias in prevalence estimates of health indicators such that the prevalence of harmful health behaviour will be under-estimated³⁵. Consequently, poor health and harmful health behaviour among the Finnish population are likely to be underestimated if the socio-economic differences in participation are not taken into account in analysing and reporting survey results.

The strengths of this paper are the inclusion of the data from a series of surveys covering 25 years and the availability of register data for both participants and non-participants. This made it possible to assess how well individuals with different occupational classes and educational levels were represented among survey participants. The unknown and missing information on occupational class and educational level in the register data are a limitation of this work. Misclassification of high education may have increased since the degrees completed abroad are not automatically updated to Finnish registers.

Also the determination of occupational class may have become more complicated as temporary and part-time employment has become more common. In future research, the questionnaire responses about socio-economic status could perhaps be utilized to decrease the misclassification. The differences in participation rates between occupational classes and educational levels might be even more prominent if the individuals could be classified perfectly. Students and farmers were so small groups that students were categorized in the same class with retirees and farmers with entrepreneurs, which made the class heterogeneous. The proportion of students varied between 1% (in 1987) and 5% (in 1997) and the proportion of farmers decreased from 8% (in 1987) to 2% (in 2012) in the survey years. The proportions of occupational classes varied a lot between the first surveys (Supplementary table 2), which may reflect the lower quality of the older register data and the fact that the capital area has been included in the FINRISK study only since 1992 and Oulu Province since 1997.

To prevent the increase of bias in estimates and the deterioration of the representativeness of health surveys, we should pay particular attention to the recruitment of those who are less willing to participate or hard to reach. For example, telephone interview has been found out to reach people with low education better than postal survey ³⁶. Previous studies have shown that even small monetary incentives can increase the response rates ³⁷ and that paying an incentive may be especially useful among the groups that would otherwise be under-represented among respondents ³⁸. Thus, participation could be increased by tailoring the recruitment separately for different socio-economic groups. However, it is worth remembering that selection bias is just one of the challenges in survey research. For example, the risk of measurement error would increase if different survey modes were used for the different population groups.

It is also important to consider using methods for missing data handling, such as multiple imputation or weighting methods ³⁹ and to include possible socio-economic factors in calculations. A Belgian study showed that using education in the calculation of post-stratification weights reduces bias in health-related indicators ⁴⁰. More research on non-response bias on health status and health behaviours estimated from survey data is needed to provide better understanding about the magnitude of the bias and to improve the methods aiming at minimizing non-response and its effects on the results.

Acknowledgements This work was supported by the Academy of Finland (grant number 266251).

Conflicts of interest: None declared.

Key points

- Non-participation in Finnish health examination surveys was most common in low occupational classes and among those with a low educational level
- Participation rates decreased over time in all occupational classes and educational levels during 25 years
- The decline in participation was fastest for those with low level of education, so the differences in participation between educational levels increased
- Our results implicate that it is important to consider methods for handling missing data and to plan carefully subgroup-specific recruitment of individuals in future surveys in order to obtain reliable health information.

References

1. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol.* 2013;42(4):1012-1014. doi: 10.1093/ije/dys223.
2. Biemer PP, Lyberg LE. *Introduction to Survey Quality.* Hoboken, NJ. ; Great Britain: Wiley; 2003.
3. Mindell JS, Giampaoli S, Goesswald A, et al. Sample selection, recruitment and participation rates in health examination surveys in Europe -- experience from seven national surveys. *BMC Med Res Methodol.* 2015;15:78-015-0072-4. doi: 10.1186/s12874-015-0072-4.
4. Tolonen H, Helakorpi S, Talala K, Helasoja V, Martelin T, Prattala R. 25-year trends and socio-demographic differences in response rates: Finnish adult health behaviour survey. *Eur J Epidemiol.* 2006;21(6):409-415. doi: 10.1007/s10654-006-9019-8.
5. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol.* 2007;17(9):643-653. doi: S1047-2797(07)00147-0.

6. Gao L, Green E, Barnes LE, et al. Changing non-participation in epidemiological studies of older people: Evidence from the cognitive function and ageing study I and II. *Age Ageing*. 2015;44(5):867-873. doi: 10.1093/ageing/afv101.
7. Jousilahti P, Salomaa V, Kuulasmaa K, Niemela M, Vartiainen E. Total and cause specific mortality among participants and non-participants of population based health surveys: A comprehensive follow up of 54 372 Finnish men and women. *J Epidemiol Community Health*. 2005;59(4):310-315. doi: 59/4/310.
8. Chinn DJ, White M, Howel D, Harland JO, Drinkwater CK. Factors associated with non-participation in a physical activity promotion trial. *Public Health*. 2006;120(4):309-319. doi: S0033-3506(05)00288-X.
9. Haapea M, Miettunen J, Veijola J, Lauronen E, Tanskanen P, Isohanni M. Non-participation may bias the results of a psychiatric survey: An analysis from the survey including magnetic resonance imaging within the Northern Finland 1966 birth cohort. *Soc Psychiatry Psychiatr Epidemiol*. 2007;42(5):403-409. doi: 10.1007/s00127-007-0178-z.
10. Volken T. Second-stage non-response in the Swiss health survey: Determinants and bias in outcomes. *BMC Public Health*. 2013;13:167-2458-13-167. doi: 10.1186/1471-2458-13-167.
11. Hansen E, Fonager K, Freund KS, Lous J. The impact of non-responders on health and lifestyle outcomes in an intervention study. *BMC Res Notes*. 2014;7:632-0500-7-632. doi: 10.1186/1756-0500-7-632.
12. Karvanen J, Tolonen H, Harkanen T, Jousilahti P, Kuulasmaa K. Selection bias was reduced by recontacting nonparticipants. *J Clin Epidemiol*. 2016;76:209-217. doi: 10.1016/j.jclinepi.2016.02.026.
13. Christensen AI, Ekholm O, Glumer C, Juel K. Effect of survey mode on response patterns: Comparison of face-to-face and self-administered modes in health surveys. *Eur J Public Health*. 2014;24(2):327-332. doi: 10.1093/eurpub/ckt067.
14. Tolonen H, Ahonen S, Jentoft S, Kuulasmaa K, Heldal J, European Health Examination Pilot Project. Differences in participation rates and lessons learned about recruitment of participants -- the European health

- examination survey pilot project. *Scand J Public Health*. 2015;43(2):212-219. doi: 10.1177/1403494814565692.
15. Ekholm O, Gundgaard J, Rasmussen NK, Hansen EH. The effect of health, socio-economic position, and mode of data collection on non-response in health interview surveys. *Scand J Public Health*. 2010;38(7):699-706. doi: 10.1177/1403494810382474.
16. Demarest S, Van der Heyden J, Charafeddine R, Tafforeau J, Van Oyen H, Van Hal G. Socio-economic differences in participation of households in a Belgian national health survey. *Eur J Public Health*. 2013;23(6):981-985. doi: 10.1093/eurpub/cks158.
17. Borodulin K, Vartiainen E, Peltonen M, et al. Forty-year trends in cardiovascular risk factors in Finland. *Eur J Public Health*. 2015;25(3):539-546. doi: 10.1093/eurpub/cku174.
18. Statistics Finland. Sosioekonomisen aseman luokitus (socio-economic classification). Handbook 17. statistics finland, helsinki. 1989.
19. Statistics Finland. Official statistics of Finland (OSF): Educational structure of population [referred: 5 october 2016]. URL: http://www.stat.fi/til/vkour/index_en.html. *Official Statistics of Finland (OSF): Educational structure of population*. 2015. http://www.stat.fi/til/vkour/index_en.html.
20. UNESCO. *International standard classification of education: ISCED 2011*. UNESCO Institute for Statistics, Montreal; 2012.
21. Ruuskanen T. Rekisteristä saadaan luotettava tieto koulutuksesta. *Hyvinvointikatsaus*. 2011(3).
22. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 1974;19(6):716-723.
23. Marrie RA, Dawson NV, Garland A. Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables. *J Clin Epidemiol*. 2009;62(5):511-7.e1. doi: 10.1016/j.jclinepi.2008.05.015.

24. R Core Team. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. 2017.
25. Korkeila K, Suominen S, Ahvenainen J, et al. Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol.* 2001;17(11):991-999.
26. Jacomb PA, Jorm AF, Korten AE, Christensen H, Henderson AS. Predictors of refusal to participate: A longitudinal health survey of the elderly in Australia. *BMC Public Health.* 2002;2:4.
27. Caetano R, Ramisetty-Mikler S, McGrath C. Characteristics of non-respondents in a US national longitudinal survey on drinking and intimate partner violence. *Addiction.* 2003;98(6):791-797. doi: 407.
28. Drivsholm T, Eplov LF, Davidsen M, et al. Representativeness in population-based studies: A detailed description of non-response in a Danish cohort study. *Scand J Public Health.* 2006;34(6):623-631. doi: M580776H71710240.
29. Boshuizen HC, Viet AL, Picavet HS, Botterweck A, van Loon AJ. Non-response in a survey of cardiovascular risk factors in the Dutch population: Determinants and resulting biases. *Public Health.* 2006;120(4):297-308. doi: S0033-3506(05)00251-9.
30. Mackenbach JP, Stirbu I, Roskam AJ, et al. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med.* 2008;358(23):2468-2481. doi: 10.1056/NEJMsa0707519.
31. Palosuo H, Koskinen S, Lahelma E, et al. *Health inequalities in Finland: Trends in socioeconomic health differences 1980-2005.* Ministry of Social Affairs and Health, Helsinki, Finland; 2009.
32. Harald K, Salomaa V, Jousilahti P, Koskinen S, Vartiainen E. Non-participation and mortality in different socioeconomic groups: The FINRISK population surveys in 1972-92. *J Epidemiol Community Health.* 2007;61(5):449-454. doi: 61/5/449.

33. Konttinen H, Sarlio-Lahteenkorva S, Silventoinen K, Mannisto S, Haukkala A. Socio-economic disparities in the consumption of vegetables, fruit and energy-dense foods: The role of motive priorities. *Public Health Nutr.* 2013;16(5):873-882. doi: 10.1017/S1368980012003540.
34. Hiscock R, Bauld L, Amos A, Fidler JA, Munafo M. Socioeconomic status and smoking: A review. *Ann N Y Acad Sci.* 2012;1248:107-123. doi: 10.1111/j.1749-6632.2011.06202.x.
35. Van Loon AJ, Tijhuis M, Picavet HS, Surtees PG, Ormel J. Survey non-response in the Netherlands: Effects on prevalence estimates and associations. *Ann Epidemiol.* 2003;13(2):105-110. doi: S1047279702002570.
36. Helakorpi S, Makela P, Holstila A, Uutela A, Vartiainen E. Can the accuracy of health behaviour surveys be improved by non-response follow-ups? *Eur J Public Health.* 2015;25(3):487-490. doi: 10.1093/eurpub/cku199.
37. Edwards P, Cooper R, Roberts I, Frost C. Meta-analysis of randomised trials of monetary incentives and response to mailed questionnaires. *J Epidemiol Community Health.* 2005;59(11):987-999. doi: 59/11/987.
38. Singer E, Kulka RA. Paying respondents for survey participation. In: Ver Ploeg M, Moffitt RA, Citro CF, eds. *Studies of welfare populations: Data collection and research issues.* National Academy Press Washington, DC; 2002:105-128.
39. Härkänen T, Karvanen J, Tolonen H, et al. Systematic handling of missing data in complex study designs – experiences from the Health 2000 and 2011 surveys. *Journal of Applied Statistics.* 2016;43(15):2772-2790.
40. Van der Heyden J, De Bacquer D, Gisle L, et al. Additional weighting for education affects estimates from a national health interview survey. *Eur J Public Health.* 2017. doi: 10.1093/eurpub/ckx005.

Table 1 Odds ratios (OR) with 95% confidence intervals (CI) for participation. Wald test measures the overall association of the predictor

Predictor	Term	OR	95% CI	Wald test	
				df	p-value
Occupational class	entrepreneur, white-collar	ref.		2	0.643
	blue-collar	0.89	(0.68, 1.15)		
	others	0.95	(0.77, 1.17)		
Educational level	high	ref.		2	< 0.001
	middle	0.71	(0.58, 0.86)***		
	low	0.38	(0.30, 0.48)***		
Sex	men	ref.		1	< 0.001
	women	1.55	(1.34, 1.80)***		
Age	spline(age)			2	< 0.001
Year	spline(year)			3	< 0.001
Occupational class *	blue-collar * middle	0.86	(0.71, 1.04)	4	< 0.001
Educational level	others * middle	0.77	(0.68, 0.87)***		
	blue-collar * low	0.93	(0.76, 1.14)		
	others * low	0.64	(0.56, 0.73)***		
Occupational class * Sex	blue-collar * women	1.13	(1.02, 1.26)*	2	0.023
	others * women	1.11	(1.01, 1.22)*		
Educational level * Age	educational level * spline(age)			4	< 0.001
Occupational class * Age	occupational class * spline(age)			4	< 0.001
Educational level * Year	educational level * spline(year)			6	< 0.001
Sex * Age	sex * spline(age)			2	< 0.001
Sex * Year	sex * spline(year)			3	0.072

Significance codes: *** $p < 0.001$, ** $0.001 \leq p < 0.01$, * $0.01 \leq p < 0.05$

Table 2 Examples of predicted participation probabilities for the years 2017 and 2022. The predictions are for the capital area of Finland

Sex	Age	Occupational class	Educational level	Probability of participation (95% CI)	
				Year 2017	Year 2022
Man	25	Others	Middle	0.29 (0.24, 0.36)	0.26 (0.19, 0.35)
Man	40	Blue-collar	Middle	0.41 (0.35, 0.48)	0.37 (0.29, 0.47)
Woman	65	Others	Low	0.45 (0.37, 0.54)	0.39 (0.28, 0.51)
Woman	50	Entr., & white-collar	High	0.61 (0.55, 0.67)	0.57 (0.48, 0.66)

Fig. 1 Participation rates by occupational class and educational level for men (left panels) and women (right panels) over the years 1987 – 2012.

Fig. 2 Predicted participation probabilities with 95% pointwise confidence intervals illustrating the interactions in the logistic regression model. The bottom panels illustrate also the model-based predictions of participation probability from the year 2012 to 2022. Predictor values are fixed to: year = 2002, sex = men, age = 45, occupational class = blue-collar and educational level = middle if the predictor is neither of the two variables in the interaction. The association between occupational class and survey year is also presented although there was not an interaction between them.



