European Journal for Research on the Education and Learning of Adults, Vol.10, No.3, 2019, pp.291-306

# The potential of statistical matching for the analysis of wider benefits of learning in later life

*Maja Wiest*
Freie Universität Berlin, Germany (maja.wiest@fu-berlin.de)

*Tanja Kutscher*
Freie Universität Berlin, Germany; Leibniz Institute for Educational Trajectories (LIfBi), Germany (tanja.kutscher@lifbi.de)

*Janek Willeke*
Humboldt Universität zu Berlin, Germany (willekej@hu-berlin.de)

*Julie Merkel*
Freie Universität Berlin, Germany (julie.merkel@fu-Berlin.de.)

*Madlain Hoffmann*
Freie Universität Berlin, Germany (Madlain.Hoffmann@fu-berlin.de)

*Katrin Kaufmann-Kuchta*
German Institute for Adult Education; Leibniz Centre for Lifelong Learning (DIE), Germany (Kaufmann-Kuchta@die-bonn.de)

*Sarah Widany*
German Institute for Adult Education; Leibniz Centre for Lifelong Learning (DIE), Germany (Widany@die-bonn.de)

## Abstract

*It is challenging to investigate wider benefits of adult learning, especially in later life, due to limited data on educational activities and non-monetary returns in large, longitudinal surveys. Statistical matching provides an approach to exploit the potential of existing data by combining data sources with complementary features based on shared information. The paper describes the matching of two data sources (German Ageing Survey and Study of Educational Attainment and Interests of Older People) in order to examine the effects of educational participation on well-being in later life. We emphasize the matching procedure and how to identify the best-matched dataset. Based on matched data, effects of educational activities on life satisfaction are examined in later life. The*

ISSN 2000-7426

DOI 10.3384/rela.2000-7426.rela201910
www.rela.ep.liu.se

*discussion focuses on future demands on data and methods for investigating wider benefits of adult learning in quantitative research.*

**Keywords:** Educational activities; lifelong learning; non-monetary returns; statistical matching; well-being

## Introduction

Education and lifelong learning are one of the keystones of modern life and societies. Benefits of adult learning beyond social status, employability, and monetary returns, intertwine our lives as for example our health and our happiness are affected by our biographical educational background and current educational activities (e.g., Field, 2012; Schuller, 2017). However, it is challenging to support the increasing awareness for wider benefits of adult learning with profound empirical evidence (Field, 2011).

How can empirical research contribute to a better understanding of the wider benefits of lifelong learning? In general, qualitative approaches can facilitate valuable insight into the effects of individual learning experiences and their social and biographical context (e.g., Manninen et al., 2014; Sloane-Seale & Kops, 2008; Tam, 2013). However, other research objectives call for quantification of causal statements on the wider benefits of learning on the population level. In order to link learning or educational participation to a specific outcome, large samples, good measures, and longitudinal data are needed to assess positive (and negative) effects of adult learning. Whereas most of the large open access panel surveys (e.g. British Household Study, German Socioeconomic Panel, US Health and Retirement Study) provide sufficient data for causal analysis, they do not assess in-depth information about adults' educational activities over and above formal education. Additionally, samples often do not represent individuals in old age adequately. Therefore, researchers have limited access to representative, longitudinal data on participation in educational activities, motives or barriers of participation, or else up to old age.

Are there alternatives to funding cost-intensive large-scale studies on wider benefits of learning and waiting for exploitable results a considerable period of time? We think there are. In this paper, we want to propose the method of statistical matching (other terms used are data fusion or matching), which makes use of available data despite its shortcomings. Following a short illustration of the dependencies of survey data and research on non-monetary returns of learning in later life, statistical matching, and its underlying rationale are introduced. To illustrate the theoretical argumentation, an example of a successful statistical matching of two data sources, with the goal to examine wider benefits of learning, is elaborated. To conclude the paper, the results and quality of the matched data are discussed with regard to the challenges of future research on wider benefits of learning in later life. This paper aims at giving an idea of what statistical matching can do and how it might be useful rather than providing the methodological and technical background for it. For those who are interested, the work by Rässler (2012) and D'Orazio, Di Zio, & Scanu (2006) provide excellent methodological overviews.

## Research with secondary data on non-monetary returns of education (in later life)

With qualitative methods traditionally dominating the field of adult education research, calls for methodological plurality encourage quantitative or mixed method approaches (Boeren, 2018; Fejes & Nylander, 2015). So far, the literature featuring quantitative studies on wider benefits of adult learning is modest but increasing and is accompanied by a constructive discussion on the underlying theoretical framework, (secondary) data, and analysis techniques in this field (Field, 2011; Rüber, Rees & Schmidt-Hertha, 2018). Studies frequently find (mostly) positive associations between adult learning and outcomes such as physical and mental health, employment, social and civic engagement (for an overview see Field, 2012; Schuller, 2017), mostly based on cross-sectional analysis, but progressively based on advanced statistical modelling techniques, such as randomized, quasi-experimental or longitudinal designs. The former show positive correlations (e.g. adult learners are happier *than* adult non-learners), whereas the latter enable conclusions about the causal links between adult learning and respective outcomes (e.g. adults are happier *because* of learning). Even if the number of studies with a quasi-experimental or longitudinal design is still small, their findings validate benefits of adult learning for a range of non-monetary outcomes and consolidate the evidence from cross-sectional analyses.

However, only a small number of studies so far has investigated benefits of lifelong learning in later life and old age. These studies are also quite heterogeneous with regard to age ranges, nations, definitions of learning and outcomes. Previous research indicates positive effects but likewise no straightforward relationship (e.g. Mestheneos & Whitnall, 2016). Studies on older adults in non-formal general interest programmes find that engagement in learning facilitates knowledge-related, psychological, and social resources that foster the well-being of participants, even or especially for vulnerable groups (Åberg, 2016; Hammond, 2004). High vulnerability, which is more likely in old age, might be compensated by continuously engaging in lifelong learning (Leung & Liu, 2011; Narushima, Liu & Diestelkamp, 2018).

Studies on representative samples can take into account systematic differences between learners and non-learners in older age and cover a broad range of learning activities. For example, Yamashita, Bardo, Liu and Yoo (2019) found in cross-sectional analysis that organised learning activities in later life mediate the effect of formal education on self-rated health. Longitudinal studies have looked at changes in well-being or health over time in relation to participation or non-participation in organised learning activities. Work by Jenkins and Mostafa found that benefits of learning, depending on the type and subject of courses and the educational background of the participants (Jenkins, 2011; Jenkins & Mostafa, 2015). The literature shows that the investigation of benefits of learning in later life requires differentiating between learning formats, learning contexts, learners and non-learners as well as between outcomes. However, data sources for quantitative analyses are extremely limited and progress on research evidence on the benefits of learning in later life relies on the utilisation of data (Jenkins, 2011; Jenkins & Mostafa, 2015).

As there is no "one size fits all" survey, secondary analysis quite naturally comes with strings attached. Where we find those strings, depends on the specific research question, but some features of secondary data are rather typical for research on the wider benefits of adult learning (in later life). Some aspects relate to *sample characteristics*: Surveys vary in their target populations and sample sizes. Especially for research on older adults a common bias is the focus on the employment age (e.g. the major international adult learning monitoring studies Adult Education Survey (AES) and Programme for the

International Assessment of Adults Competencies (PIAAC) do not sample respondents over 65 years). Also, institutionalized adults are virtually not covered. If older respondents are included, a low number of cases restricts analyses, as the subsample of older people is rather small. Whilst estimates on an overall impact of adult education are informative, often we are interested in benefits for specific (vulnerable) groups (e.g. people at risk of poverty in old age). These specific subsamples may have insufficient case numbers for sound analysis. Other aspects relate to the *scope of the survey*. As a prerequisite, there should be at least some information on participation in adult education and on the outcome of interest. Huge panel studies often have an interdisciplinary background and feature a framework from psychology, economics, and/ or sociology. Educational sciences, let alone adult education, however is rarely involved. You may find adults learning activities totally neglected, you may find questions on job-related training, only. If there is information on adult education in any form, *measurement quality* always is an issue (Felstead, Green & Mayhew, 1999). Depending on the theoretical framework and research question, specific characteristics of adults learning are necessary to assess, e.g. is a learning activity rather job- or leisure related, is it formal, non-formal or informal, who provided it, how time-intensive was it. Similar requirements apply to information on the outcome, of course. Given the limitations of existing data, research on non-monetary returns of learning benefits from the utilisation of different data sources.

## Statistical Matching

Statistical matching combines existing data from different data sources based on shared and unique information for new analysis. Typically, the combination of data sources (one being the donor of information, one being the recipient of information) supplements information relevant to a research objective, with both data sources having some mutual information. In educational sciences, statistical matching is not common, yet. Since more and more data is available from national statistics, research institutes, and commercial enterprises, statistical matching opens new possibilities for an efficient, resourceful use of existing information originating from different data sources.

In the following, the basic idea of statistical matching is illustrated. A new dataset on participation in non-formal education and its effect on well-being can be matched based on shared information (*matching variables*) which are associated with well-being and educational activities and are included both in the dataset 1 (*donor*) containing the data on educational activities and in the dataset 2 (*recipient*) providing data on subjective well-being. The key requirement is that both datasets describe the same target population. Therefore, the goal is to replicate the distribution of the recipient data in the new matched data by matching the information from the donor dataset using the matching variables. The quality of matching depends highly on several factors: (1) an accurate definition of the target population, (2) harmonization of data (e.g., adjustment of answer formats or classifications of categorical variables if identical information is differently assessed in the data sources), (3) selection of matching variables, and (4) selection of a suitable matching method. The amount and quality of shared variables as well as how strongly predicted variables depend on the shared variables (e.g., correlation of shared variables with well-being and educational activities) are crucial for a high-quality matching.

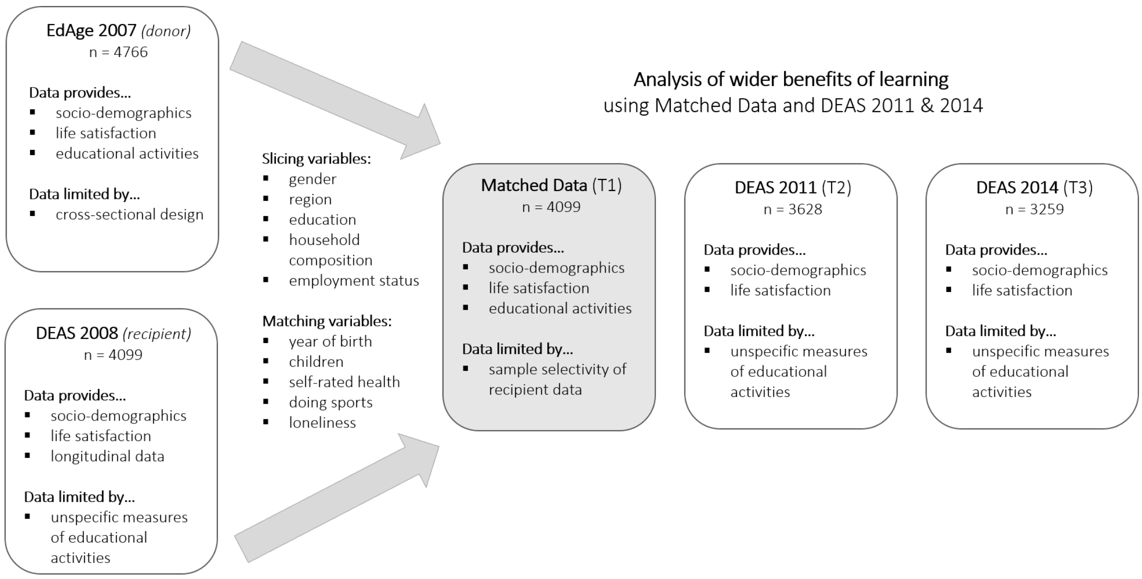
The most important step before matching is the harmonization of data sources (D'Orazio et al., 2006; Van der Laan, 2000). In real world applications of statistical matching, data sources differ on various aspects (e.g., reference periods, populations, variables, etc.). To apply statistical matching, the data sources need to be as similar as possible concerning specific characteristics of the target population and measures of shared information. Therefore, it may be necessary, for example, to exclude some individuals in one data source to have the same target population in both data sources. Or variables need to be harmonized before considering any matching as they measure the same construct on different scales or use different classifications.

All shared variables between data sources may be used as a matching variable (D'Orazio et al., 2006). Yet, the selection of suitable matching variables should primarily be driven by theoretical assumption about the relation of underlying constructs as well as by statistical aspects such as sufficient correlation between potential matching variables and the variables of interest that are only included either in the donor or the recipient dataset. However, computational complexity of the matching increases with the number of matching variables. Additionally, the selection of matching variables also depends on the method of matching (for an overview, see D'Orazio et al., 2006). Again, the method of matching should be chosen depending on the data sources, the research questions, and the follow-up analysis. In our example, we used a *nearest neighbour method.* Further, it is advisable to include so-called slicing variables to refine statistical matching (D'Orazio et al., 2006). Slicing variables are mostly nominal scaled variables (e.g., gender, employment status) which exactly match in the data sources. Using a set of slicing variables, it is possible to conduct statistical matching within each of a stratified subsample with a specific characteristic. For example, for a highly educated woman who is employed and has children in dataset 1, potential donors will be selected from the subsample of women with the same features in dataset 2. Thus, including a number of slicing variables can improve the quality of matching results by reducing the likelihood of a random assignment, as the number of potential donors gets smaller.

## *Example: Matching of EdAge and DEAS*

Our project aim was to investigate effects of educational activities on subjective well-being in the second half of life (40 years and older). Our objective requires continuous long-time observation of a representative sample and is beyond the scope of a typical research project limited in time and finances. Available data sources were limited either with regard to the population (e.g., only individuals in midlife are questioned), instruments (e.g., educational activities are measured too broad) or study design (e.g., cross-sectional). Those challenges led to the search for alternative solutions such as statistical matching. Two data sources were identified as limited on their own but suitable for data matching (see Figure 1): the German Aging Survey (DEAS) providing longitudinal data for individuals 40 years and older on various topics and questions. However, educational activities are measured too broadly to allow estimating the effect of educational activities on well-being ideally. The data collected within the 'Educational Attainment and Interests of Older People (EdAge)' project provides excellent data on educational activities in later life – but the data is cross-sectional only. In the following, we describe both datasets, why they are a good match and how matching enables the analysis of our research question (see section *scope of the matching*).

Figure 1: Statistical matching of DEAS 2008 and EdAge and follow-up analysis based on DEAS 2011 and DEAS 2014




## Description of DEAS (Recipient)

Data of the DEAS (Klaus, Engstler, Mahne, Wolff, Simonson, Wurm & Tesch-Römer, 2017) is used as the longitudinal data source. The DEAS is an ongoing nation-wide representative survey of German community-dwelling adults aged 40 years and older, which started in 1996. Every six years a new baseline sample stratified by age, gender, and place of residence (Eastern or Western Germany) is drawn. Since 2011, panel participants are not only interviewed with every new wave but also every third year. Participation in the survey comprises a computer-assisted personal interview as well as an additional self-administered questionnaire. The DEAS has a complex data structure with multiple longitudinal data points for panel participants. Further, the survey provides a wide range of information for social reporting and captures various topics (e.g. well-being). However, educational attainment is addressed only marginally. Employed participants are asked to report non-formal education by "How many courses or programs for occupational training or retraining have you attended in the past 6 years?" (giving the number in total) and everyone is asked "How often do you take classes or go to lectures, for example, for education and further training?" (on a scale ranging from daily to never). Thus, the DEAS does not differentiate between non-formal education and informal learning activities and applies an ambiguous scaling. The data is available to the scientific community. The DEAS was chosen as the recipient because of its longitudinal data structure. Statistical matching has already been done with this survey data. Simonson, Gordo and Kelle (2012) matched the DEAS with pension insurance data collected by public administration with the aim to analyse the employment and family biographies of the baby boomers and their general financial situation and attitudes.

## Description of EdAge (Donor)

Table 1: Descriptives of DEAS, EdAge, matched dataset and analysis dataset

| | | EdAge (donor) | DEAS (recipient) | Matched dataset | Analysis dataset* |
|---|---|---|---|---|---|
| n | | 4766 | 4099 | 4099 | 1920 |
| Year of birth | | 1946.21 (9.80) | 1946.89 (9.39) | 1944.88 (9.39) | 1945.10 (8.95) |
| Gender | *Female* | 52 | 49 | 49 | 51 |
| Region | *West* | 67 | 66 | 67 | 63 |
| Formal education | *Low* | 18 | 8 | 8 | 6 |
| | *Medium* | 55 | 51 | 51 | 50 |
| | *High* | 28 | 41 | 41 | 44 |
| Employment status | *Employed* | 42 | 36 | 36 | 36 |
| | *Retired* | 44 | 52 | 52 | 52 |
| | *Unemployed* | 7 | 3 | 3 | 4 |
| | *Other* | 7 | 8 | 8 | 8 |
| Children | *One or more* | 86 | 89 | 89 | 89 |
| Household composition | *Two persons and more* | 72 | 82 | 82 | 84 |
| Doing sports | *At least once a week* | 25 | 52 | 52 | 56 |
| | *At least once a month* | 6 | 7 | 7 | 8 |
| | *Less often* | 15 | 12 | 12 | 12 |
| | *Never* | 53 | 28 | 28 | 24 |
| Loneliness[1] | | 1.61 (0.85) | 1.16 (0.50) | 1.14 (0.48) | 1.13 (0.46) |
| Self-rated health[1] | | 3.68 (1.02) | 3.57 (0.81) | 3.57 (0.81) | 3.61 (0.79) |
| Non-formal education | *At least one reported* | 28 | - | 33 | 34 |
| Informal learning | *At least one reported* | 48 | - | 58 | 59 |
| Life satisfaction[2] | | 4.15 (0.67) | 3.94 (0.71) | 3.94 (0.71) | 3.96 (0.70) |

*Notes*: Mean (SD) or percentage reported; * Analysis dataset = Matched dataset (T1), DEAS 2011 (T2) and DEAS 2014 (T3) without missing in life satisfaction at T1, T2 or T3; [1] Range:1-4; [2] Range 1-5 (see also Table 2)

The dataset providing in-depth information on educational activities and learning was collected in the project "EdAge – Bildungsverhalten und -interessen Älterer [Educational Attainment and Interests of Older People]" (Tippelt, Schmidt, Schnurr, Sinner, & Theisen, 2009). This project was a one-time add-on study to the recurring AES, which covers adults' participation (aged 25 to 64) in education and training and is one of the main data sources for EU lifelong learning statistics. The EdAge study has the same broad scope of adult learning as the AES, but it is a representative cross-sectional study of German community-dwelling adults aged 45 to 80. Computer-assisted personal interviews were carried out in 2007 (Schmidt, 2009). With our research question in mind, we focus on two types of educational activities within the wide range of adult learning in the EdAge. Non-formal education (NFE) is assessed by asking if participants did any kind of seminar, course, class or private lessons in the past 12 months. For the analysis, we use NFE as a dichotomous indicator (0 = no participation in NFE, 1 = at least one participation in NFE is reported). Supported by a list of different informal learning activities (INF), participants report if they have learned something in the past 12 months. The list covers: (A) reading of books and magazines, B) using the computer or the internet, C) using television, radio, or media, D) tours in museums, historical places, natural monuments, or industrial plants, E) visiting libraries or learning center, and F) learning through family members, friends or colleagues. For the analysis, INF is used as a dichotomous indicator (0 = no INF activity is reported, 1 = at least one INF activity is reported). Unfortunately, this dataset is not available to the public. The EdAge dataset was chosen as the donor because its cross-sectional data provides in-depth information on educational activities that are missing in the DEAS.

## Scope of the matching

Matching of the DEAS and EdAge data looked promising for our research scope as educational activities and well-being are measured in both data sources (see Figure 1), but with different degrees of differentiation (EdAge: various measures of educational activities, life satisfaction measured; DEAS: life satisfaction measured, educational activities measured very broad, on an imprecise scale). In addition, the cohort-sequential design of the DEAS provides a rich longitudinal data structure. The aim of the matching was to combine the data to analyse the effect of participation in non-formal and informal education on life satisfaction in later life. Missing data on educational activities in the DEAS dataset (recipient) is matched from the EdAge dataset (donor) using a set of matching and slicing variables assessed in both datasets. Furthermore, the population, as well as the time of assessment, is highly comparable. Both surveys were assessed in sequential years (EdAge in 2007; third wave of the DEAS in 2008) and study overlapping birth cohorts (EdAge: born between 1926-1961; DEAS born between 1911-1974). Thus, for the statistical matching, data of all survey participants were included if they: (a) were born between 1926 and 1961, (b) had German citizenship, and, additionally for DEAS, (c) were interviewed in 2008 and 2011 or 2008 and 2014, this ensures that at least two data points are available for each participant (this is important for the follow-up analysis). The samples represent the same target population (see Table 1). However, the datasets differ with regard to the proportion of individuals with a low education (lower in DEAS), the proportion of retired individuals (higher in DEAS), and the proportion of individuals reporting being regular physically active (higher in DEAS).

Table 2: Overview of variable harmonization between data sources

| Variable | EdAge (donor) | DEAS (recipient) | Harmonization |
|---|---|---|---|
| Doing sports | How often do you do each of the following activities: Active sport<br><br>1 = Weekly<br>2 = Monthly<br>3 = Less often<br>4 = Never | How often do you do sports such as hiking, soccer, gymnastics, or swimming?<br><br>1 = Daily<br>2 = Several times a week<br>3 = Once a week<br>4 = 1-3 times per month<br>5 = Less often<br>6 = Never | Recoding of the scale in the DEAS was necessary (value 1 includes value 1, 2, and 3 of scale used in DEAS):<br><br>1 = At least once a week<br>2 = At least once a month<br>3 = Less often<br>4 = Never<br><br>A low value represents a high level of doing sports |
| Loneliness | I often feel lonely<br><br>1 = Strongly agree<br>2 = Agree<br>3 = Disagree<br>4 = Strongly disagree | How often you have felt this way during the past week: I felt lonely<br><br>1 = Rarely or none of the time<br>2 = Some or a little bit of the time<br>3 = Occasionally or a moderate amount of time<br>4 = Most or all of the time | The scales in both datasets differ; recording of values in EdAge was necessary:<br><br>1= Rarely<br>2= A little<br>3= Occasionally<br>4= Most of the time<br><br>A low value represents a low level of loneliness |

Table 2. (Continued)

| Variable | EdAge (donor) | DEAS (recipient) | Harmonization |
|---|---|---|---|
| Self-rated health | How satisfied are you with your present state of health? Are you…<br><br>1 = Very satisfied<br>2 = Satisfied<br>3 = Depends<br>4 = Unsatisfied<br>5 = Very unsatisfied | How would you rate your present state of health?<br><br>1 = Very good<br>2 = Good<br>3 = Average<br>4 = Bad<br>5 = Very bad | The scales in both datasets differ; scale of DEAS was used, data was recoded:<br><br>5 = Very good<br>4 = Good<br>3 = Average<br>2 = Bad<br>1 = Very bad<br><br>A high value represents a better self-rated health |
| Life satisfaction | ▪ In most ways my life is close to my ideal<br>▪ I am satisfied with my life<br>▪ So far, I have gotten the important things I want in life<br><br>1 = Strongly Agree<br>2 = Agree<br>3 = Disagree<br>4 = Strongly disagree | ▪ In most ways my life is close to my ideal<br>▪ I am satisfied with my life<br>▪ So far, I have gotten the important things I want in life<br><br>1 = Strongly agree<br>2 = Agree<br>3 = Neither agree nor disagree<br>4 = Disagree<br>5 = Strongly disagree | Transforming of the answer format in EdAge dataset into a 5-point rating scale was done; individual scale means were calculated in both datasets:<br><br>5 = Strongly agree<br>…<br>1 = Strongly disagree<br><br>A high value represents a high level of life satisfaction;<br>Items are from the Satisfaction with Life Scale (Pavot & Diener, 1993) |

*Notes*: All other variables used as matching or slicing variables were assessed identically in the both datasets.

## Harmonization of data and selection of matching variables

In the next step, we had to identify the best combination of matching variables. Matching variables could be selected from the following set of shared variables in DEAS and EdAge: socio-demographic variables (year of birth, gender, region, employment status, level of formal education, household size, having children) and other variables (doing sport, loneliness, self-rated health, life satisfaction).

The carried-out harmonization can be found in Table 2. Due to a high amount of missing values in the DEAS data, life satisfaction could not be used as a matching variable but is suitable as an external criterion for validating the quality of matching[1]. To identify possible matching variables, we calculated pairwise correlation and association values to identify which potential matching variables predict both educational activities and life satisfaction best. As a result, we identify that level of formal education, employment status, and year of birth are most predictive of both life satisfaction and educational activities, and suitable matching variables. Level of formal education and employment status were included as slicing variables, since those are among the strongest predictors of educational activities. This approach impeded that e.g. employed individuals are matched with unemployed individuals. We further included additional slicing variables (gender, household size, region). Because we were sceptical about a parsimonious statistical matching model including only year of birth as a matching variable, we tested additional models that included more matching variables. However, these additional matching variables, except having children, were weakly associated with both life satisfaction and educational activities. All matching models are present in Table 3.

Table 3: Overview of different combination of matching and slicing variables

| # | Matching variables | Slicing variables |
|---|---|---|
| 1 | year of birth | employment status, education |
| 2 | year of birth | employment status, education, gender |
| 3 | year of birth | employment status, education, gender, household composition |
| 4 | year of birth | employment status, education, gender, household composition, region |
| 5 | year of birth, children | employment status, education, gender, household composition, region |
| 6 | year of birth, children, self-rated health | employment status, education, gender, household composition, region |
| 7 | year of birth, children, self-rated health, doing sports | employment status, education, gender, household composition, region |
| 8 | year of birth, children, self-rated health, doing sports, loneliness | employment status, education, gender, household composition, region |

*Notes*: # Combination number

## Matching procedure and results

We used the nearest neighbour method, because the matching was based on several differently scaled matching variables (e.g., year of birth and loneliness). The method aims to match each record in the recipient to the closest record in the donor, according to a distance measure computed using the matching variables. There are different distance measures (e.g., Mahalanobis distance, Euclidean distance, etc.). As we use differently scaled matching variables, it is recommended to use the Gower distance (D'Orazio et al., 2006)[2]. This distance measure considers differences between categorical and metric variables and computes an average of distances for each variable. Data preparation and data harmonization were done using R (R Core Team 2018), the matching was done using the package StatMatch (Version 1.2.5; D'Orazio, 2017).

Statistical matching is a complex procedure, therefore evaluating the quality of the matching is indispensable (cf. Van Hattum & Hoijtink, 2008). We identified the best matching by considering the following aspects (see Table 4). First, the mean and the standard deviation of absolute distances within matching variables, a good matching is indicated by smaller values. All eight matching models indicated low distance values. Note that with increasing numbers of matching variables, the values inevitably increase. Therefore, these indices can only be compared directly within the same combination of matching variables. Second, we looked at the number of potential donors within a stratified subsample. Again, a smaller number is aspired, as it indicates a lower likelihood of a random assignment of identified donors to a particular case in the recipient dataset. Here, the average number of donors is the lowest for the combination 8 (on average 1.50 donors) compared with the other combinations (range: 39.84 to 1.72). In addition, we considered dissimilarity in the distribution of variables (here non-formal education, informal learning, and life satisfaction) using the total variation in distance. This indicates whether a particular variable is identically distributed in both the original and matched datasets. A small value means a small discrepancy. We can see that dissimilarity for educational activities and life satisfaction is generally low for all matching models. Due to a larger number of matching variables, it is slightly larger for combination 8 (0.11). Finally, we looked at life satisfaction[3] as an external criterion. The indices based on absolute differences in an external evaluation criterion between the original and matched datasets are central to evaluate the quality of the matching. We found that with an increasing number of considered matching variables, the differences in life satisfaction between the original and the matched data decreases. The combination 8 had the lowest mean absolute difference (0.69) and the lowest root mean square error representing the sample standard deviation of differences (0.89) compared to the other combinations. To sum up, we decided to use the matched data based on the matching variables: year of birth, having children, self-rated health, doing sports, and loneliness and the slicing variables: employment status, level of formal education, gender, household composition, and region.

Table 4: Results of matching using the Gower distance for different combinations of matching variables

| # | Mean.abs.dist (SD) | Mean.donors (SD) | TVD.NFE | TVD.INF | TVD.LS | LS.Mean.abs.diff | LS.RMSE |
|---|---|---|---|---|---|---|---|
| 1 | 0.00 (0.02) | 39.84 (24.62) | 0.01 | 0.04 | 0.02 | 0.80 | 1.01 |
| 2 | 0.00 (0.02) | 20.77 (12.75) | 0.02 | 0.04 | 0.03 | 0.82 | 1.04 |
| 3 | 0.00 (0.03) | 14.71 (10.77) | 0.02 | 0.05 | 0.06 | 0.79 | 1.00 |
| 4 | 0.01 (0.05) | 8.52 (7.16) | 0.01 | 0.05 | 0.05 | 0.80 | 1.00 |
| 5 | 0.01 (0.04) | 7.22 (6.59) | 0.01 | 0.05 | 0.05 | 0.77 | 0.98 |
| 6 | 0.02 (0.07) | 2.95 (2.83) | 0.01 | 0.03 | 0.05 | 0.73 | 0.94 |
| 7 | 0.05 (0.09) | 1.72 (1.24) | 0.06 | 0.09 | 0.08 | 0.71 | 0.92 |
| 8 | 0.06 (0.10) | 1.50 (0.96) | 0.05 | 0.10 | 0.11 | 0.69 | 0.89 |

*Notes*: # = Matching model with a specific set of matching and slicing variables (see Table 3); Mean.abs.dist = mean absolute distance in matching variables; SD = standard deviation; Mean.donors = the mean number of potential donors within slicing variables; TVD = total variation distance in non-formal education (NFE), informal education (INF), and life satisfaction (LS); LS.Mean.abs.diff = mean absolute difference in individuals' means in life satisfaction originated from the DEAS dataset and the EdAge dataset; LS.RMSE = the root mean square error for individuals' means in life satisfaction.

In Table 1, the descriptive statistics for the recipient, the donor, and the matched dataset are shown. As mentioned, the matched data aims at reflecting the distribution of the recipient dataset (DEAS). This was mainly accomplished. Given the good quality of the matching, we used the matched data for examining our research question by predicting life satisfaction three (T2), respectively, six years later (T3) by participation in educational activities at T1. Here, original longitudinal data of the DEAS (T2, T3) is used with the matched data (T1).

### *Analysis of matched data: Example*

Using matched data and the two sequential waves from 2011 and 2014 of the DEAS, we examined wider benefits of learning in later life. The effect of educational activities on life satisfaction was investigated by generalized linear models using R. Four separate models were analysed as participation in non-formal education (NFE) and informal learning (INF) at T1 were differentiated as well as life satisfaction three years (T2) and six years (T3) later. The sample in the analysis is smaller due to missing data on life satisfaction at T1, T2 or T3 (n = 1920, $M_{age\ T1}$ = 62.5 (SD = 8.95), 49% female, 43.9% high educated, see Table 1). All analyses were controlled for age, gender, region, level of formal education, employment status, having children, household composition, self-rated health, doing sports, loneliness and life satisfaction at T1. Participating in NFE was associated with higher life satisfaction three years later ($B$ = 0.06, SE = 0.03, p = .02, CI95 [0.01, 0.12]). However, NFE did not predict life satisfaction at T3. For INF, there was no significant association with life satisfaction neither at T2 nor at T3. Life satisfaction at T2 was significantly predicted by life satisfaction at T1, self-rated health and household composition (see Table 5). Life satisfaction at T3 was significantly predicted only by life satisfaction at T1 and T2 and self-rated health (see Table 5). Participants who did any kind of course, seminar, class or private lesson at T1 reported a higher life satisfaction three years later[4]. Since the analyses are based on matched data, they only provide an estimate of the effect within the sample population.

Table 5: Results of regression analysis predicting life satisfaction at T2 and T3 using DEAS data of 2011 and 2014 by educational activities at T1 based on matched data

| Predictor | NFE → life satisfaction at T2 | NFE → life satisfaction at T3 | INF → life satisfaction at T2 | INF → life satisfaction at T3 |
|---|---|---|---|---|
| NFE/INF at T1 | **0.06** | 0.00 | 0.02 | -0.02 |
| Year of birth | -0.00 | 0.00 | -0.00 | 0.00 |
| Gender | -0.01 | 0.03 | -0.01 | 0.03 |
| Region | 0.02 | 0.04 | 0.02 | 0.04 |
| Formal education[1] | 0.01 | 0.00 | 0.01 | 0.01 |
| Employment status[2] | -0.02 | 0.02 | -0.00 | 0.02 |
| Children | -0.01 | 0.04 | -0.01 | 0.04 |
| Household composition | **0.09** | -0.06 | **0.09** | -0.06 |
| Doing sports[3] | -0.01 | 0.00 | -0.01 | 0.00 |
| Loneliness | -0.01 | 0.02 | -0.01 | 0.02 |
| Self-rated health | **0.05** | **0.03** | **0.05** | **0.03** |
| Life satisfaction at T1 | **0.60** | **0.33** | **0.60** | **0.33** |
| Life satisfaction at T2 | - | **0.44** | - | **0.44** |

Note: Significant coefficients are printed bold ($p < .05$); $n_{\text{analysis dataset}} = 1920$; reference categories for predictors are displayed in Table 1; [1]used as 1 = high education, 0 = medium and low education; [2]used as 1 = employed, 0 = retired or otherwise not employed; [3]used as a continuous predictor

## Discussion

Introducing statistical matching, we aimed at giving an insight into the not yet disseminated method in educational sciences, which allows to utilize secondary data. Since adult education is rarely within the scope of interdisciplinary, long-running panel studies and learning activities are consequently most often measured inadequately or not at all, quantitative research on wider benefits of adult learning needs to find new ways to push the field forward. Statistical matching enabled us to utilise existing datasets and examine the effect of participation in different learning activities on life satisfaction three, respectively, six years later in a sample covering a broad age range. Within the original data, this analysis would not have been possible.

Analysis of the wider benefits of adult learning with secondary data brings the advantage of affordable access to data that would blow up the budget and time of any medium-sized project. From an ethical and legal point of view, it accounts for data minimization and the sustainable use of personal data. However, given the multi-purpose design of large-scale assessments, research with secondary data risks a data-driven approach. Research conducted in this manner can only ever produce findings on the benefits of learning (in later life) that are already embedded in the data source (Field, 2011). Secondary data has several restrictions. First, most likely, the data collected will always deviate from primary data a researcher would collect in his or her own study. Second, provision of these kind of data sources will vary. Large panel studies, for example, typically conglomerate in western countries with a well-funded science infrastructure. Third, results will always relate to some unique features of the data base (e.g. cohort, reference time, type of adult learning), therefore, generalisability of evidence on the benefits of adult learning will always be limited to some extent (Rüber et al., 2018). Thus, it is essential to balance trade-offs between the research interest and the data basis and to critically reflect findings against this backdrop. Statistical matching does not offer a solution to all of these restrictions, it does though provide technical means to overcome limitations, especially with regard to scarce information. In our example, we were able to utilise a cross-sectional special interest survey with a good measurement quality of educational activities in combination with data of a multi-purpose panel survey that provided the desired sample characteristics and design to analyse wider benefits of learning in later life. Looking ahead, this type of utilisation could be maximized if prospective cross-sectional surveys were already designed with potential matching data sources in mind.

Given that we often have to make concessions within secondary data analysis in terms of sample populations, assessed information, or study design, this approach allows to some extent to deal with those restrictions. Statistical matching can also guideline which indicators or populations might be worthwhile to study in-depth and therefore enable researchers to efficiently invest resources (e.g. focus on specific populations, focus on specific indicators and outcomes). With regard to the competitive nature of funding this is an advantage, but also from an ethical and legal point of view this ensures data minimization and the sustainable use of personal data.

Some limitations need to be addressed specifically to the matching of the DEAS and EdAge data. In general, an *in-depth checking of the data*, especially working with secondary data, is necessary before any matching. In the distribution of education in both data sources, we see a mismatch in low-educated individuals (see Table 1). A rather typical selective pattern of attrition in longitudinal surveys leads to an overrepresentation of older individuals with high formal qualification in the panel data of the DEAS in comparison to the EdAge data. We used data from all DEAS individuals interviewed in 2008 to ensure a sufficient sample size. However, in 2008, not only a new representative sample of people aged 40 to 85 years was drawn, but also participants from the first wave in 1996 and second wave in 2002 were re-interviewed, therefore highly educated individuals were overrepresented in the recipient data. It is crucial to keep the sample specifics in mind for follow-up analysis of the matched data as the distribution characteristics of the recipient data source are replicated in the matched data. Further, *identifying matching variables* is critical, considering all shared information on the other hand, is not recommended. Matching the EdAge data to the DEAS data was a challenge and a conservative treatment of considered matching variables led to an unsatisfying result (cf. Table 4), therefore additional variables were checked and self-rated health, doing sports, and loneliness were included which improved the matching. Given that the complexity of the matching increases with each additional matching variable, the process of including more matching variables needs to be done carefully. It is also challenging that evaluation indices may not agree with each other and support different matchings. Therefore, it is advisable to consider different indices for deciding what matching yield the best result.

Our analysis is in line with previous findings on the positive effect of educational activities on well-being in later life (Jenkins, 2011, Jenkins & Mostafa, 2015), even though we covered a greater age range and had a longer period between T1 and T2. So far, no other study has been using longitudinal data from older participants in Germany to examine wider benefits of learning. However, the positive effect is limited to participating in non-formal education only and to life satisfaction three years later only. Within our analysis, life satisfaction six years later was not related to non-formal education, this supports the notion that direct effects of educational activities are not translated over a long period and rather affect individual's well-being within shorter time frames (Hoffmann, Wiest, Widany, & Kaufmann-Kuchta, in press). Unfortunately, we were not able to differentiate between job- or leisure-related educational activities, nor to include how time-intense those were. The results show no contribution of informal learning activities to life satisfaction. Many of the listed activities feature learning formats embedded in the everyday life of participants. An analysis that refers to distinctive characteristics of these activities might provide a different picture on the impact of informal learning. It would be interesting to differentiate, for example, between more social activities (e.g. learning from friends and family) and more cultural activities (e.g. visiting a museum). An important limitation of our analyses is the sample. Although a broad age range is included (46-82 years at T1), the sample is highly educated and thus an above average educational participation can be assumed as well as an overall better life situation (better health, better financial resources), which may result in less interindividual differences in life satisfaction.

## Remarks on statistical matching in general

Statistical matching in general has its limitations. Among others, as stressed before, the quality of matching highly depends on the quality and scope of the data sources. If crucial information is lacking, no data harmonization is able to fill this gap. Second, all matching procedures do have their strengths and weaknesses. The nearest neighbour approach used to match data of the DEAS and the EdAge is more heuristic than statistical in nature. This means there are no pre-defined measures to evaluate the predicted values. However, using multiple imputation some of the uncertainty of the prediction can be quantified (D'Orazio et al., 2006). Yet it should be kept in mind that any prediction error committed during the matching also carries over into follow-up analysis. Third, as for all statistical methods a profound knowledge is needed to understand and interpret findings of different matching procedures.

## Conclusion

A real-world application of statistical matching does come with its challenges, but at the same time, it allows us to deal with restricted secondary data in an efficient, inventive, and resourceful way. With regard to benefits of learning in later life, statistical matching can guideline investment in future research by estimating effects of educational attainment on well-being in specific underrepresented populations such as old age and contribute to an informed debate on public spending in education. Most surveys that cover adults learning more comprehensive (AES, PIAAC), are cross-sectional in design and therefore limited to findings on associations. Statistical matching with panel studies significantly improves their analytic potential. Minor changes in measurements can significantly improve the fit of matching variables and therefore the overall quality of the statistical matching. Therefore, we propose cross-sectional studies keep potential recipient panel-studies in mind.

## Acknowledgements

We thank the EdAge project for the possibility to analysis the data within our project. This work was part of the project "Educational Activities and Quality of Life in the Second Half of Life (BiLeQua)" and was funded by a grant of the German Federal Ministry of Education and Research. The content is the sole responsibility of the authors.

## Endnotes

[1] In a perfect scenario, a third variable/construct is available in both datasets to evaluate the quality of the matching. Unfortunately, this is not the case in this matching scenario due to limited shared information in both datasets. Thus life satisfaction serves as the outcome we are investigating in the follow-up analysis as well as the variable to assess the quality of the matching process between recipient and donor data.

[2] For comparisons, we also matched the two datasets using the Manhatten distance, Mahalanobis distance and Euclidean distance with different combinations of matching and slicing variables. None of these measures yielded better matching results than using the recommended Gower distance.

[3] It is important to note that life satisfaction was assessed using different scales in the datasets (see Table 2), therefore a perfect match was only possible for the scores 1, 2.33, 3.66, and 5 as these are the transformed values of the 4-point scale of the EdAge data (1 = 1, 2 = 2.33, 3 = 3.66, 4 = 5).

[4] The same pattern of results are found when checking for different subsamples. NFE participation is related to higher life satisfaction at T2 only in analysis including a) people 65 years and older ($n = 872$, $M_{age} = 70.83$, $SD = 4.33$, 99% retired), b) retired individuals ($n = 988$, $M_{age} = 69.19$, $SD = 5.57$), and c) individuals reporting a low self-rated health ($n = 782$, $M_{age} = 63.16$, $SD = 8.94$, 31% employed, 57% retired, 12% otherwise not employed). We did not find any other significant association between educational activities and life satisfaction at T2 or T3 when analysing these subsamples.

## References

Åberg, P. (2016). Nonformal learning and well-being among older adults: Links between participation in Swedish study circles, feelings of well-being and social aspects of learning. *Educational Gerontology*, 42, 411–422. doi:10.1080/03601277.2016.1139972

Boeren, E. (2018). The methodological underdog: A review of quantitative research in the key adult education journals. *Adult Education Quarterly*, 68, 63–79. doi: 10.1177/0741713617739347

D'Orazio, M. (2017). Statistical matching and imputation of survey data with StatMatch: Retrieved November 06, 2018, from https://cran.r-project.org/ web/packages/StatMatch/ vignettes/Statistical_Matching_with_StatMatch.pdf.

D'Orazio, M., Di Zio, M., & Scanu, M. (2006). *Statistical matching: Theory and practice*. *Wiley series in survey methodology*. Chichester England, Hoboken NJ: Wiley.

Fejes, A., & Nylander, E. (2015). How pluralistic is the research field on adult education? Dominating bibliometrical trends, 2005-2012. *European Journal for Research on the Education and Learning of Adults*, 6, 103–123. doi: 10.3384/rela.2000-7426.rela9063

Felstead, A., Green, F., & Mayhew, K. (1999). Britain's training statistics: A cautionary tale. *Work, Employment and Society*, 13, 107–115.

Field, J. (2011). Researching the benefits of learning: the persuasive power of longitudinal studies. *London Review of Education*, 9, 283–292. doi: 10.1080/14748460.2011.616320

Field, J. (2012). Is lifelong learning making a difference? Research-based evidence on the impact of adult learning. In D. Aspin, J. Chapman, K. Evans, & R. Bagnall (Eds.), *Second International Handbook of Lifelong Learning* (pp. 887–897). Dordrecht: Springer. doi: 10.1007/978-94-007-2360-3_54

Hammond, C. (2004). Impacts of lifelong learning upon emotional resilience, psychological and mental health: fieldwork evidence. *Oxford Review of Education*, 30, 551-568. doi: 10.1080/0305498042000303008

Hoffmann, M., Wiest, M., Widany, S., & Kaufmann-Kuchta, K. (in press). Der Beitrag non-formaler Bildungsbeteiligung für Lebenszufriedenheit älterer Erwerbstätiger [The role of non-formal education for life satisfaction of older workers]. *Zeitschrift für Erziehungswissenschaft*.

Jenkins, A. (2011). Participation in learning and wellbeing among older adults. *International Journal of Lifelong Education*, 30, 403-420. doi: 10.1080/02601370.2011.570876

Jenkins, A., & Mostafa, T. (2015). The effects of learning on wellbeing for older adults in England. *Ageing & Society*, 35, 2053-2070. doi: 10.1017/S0144686X14000762

Klaus, D., Engstler, H., Mahne, K., Wolff, J. K., Simonson, J., Wurm, S., & Tesch-Römer, C. (2017). Cohort profile: The German Ageing Survey (DEAS). *International Journal of Epidemiology*, 46, 1105-1105. doi: 10.1093/ije/dyw326.

Leung, D. S., & Liu, B. C. (2011). Lifelong education, quality of life and self-efficacy of Chinese older adults. *Educational Gerontology*, 37, 967-981. doi: 10.1080/03601277.2010.492732

Manninen, J., Sgier, I., Fleige, M., Thöne-Geyer, B., Kil, M., Možina, E., Operti, F. (2014). Benefits of lifelong learning in Europe: Main results of the BeLL-Project: Research report. Retrieved January 12, 2016, from http://www.bell-project.eu/cms/wp-content/uploads/2014/06/BeLL-Research-Report.pdf

Mestheneos, E., & Withnall, A. (2016). Ageing, learning and health: Making connections. *International Journal of Lifelong Education*, 35, 522-536. doi: 10.1080/02601370.2016.1224039

Narushima, M., Liu, J., & Diestelkamp, N. (2018). Lifelong learning in active aging discourse: Its conserving effect on wellbeing, health and vulnerability. *Ageing & Society*, 38, 651-675. doi: 10.1017/S0144686X16001136

Pavot, W., & Diener, E. (1993). The affective and cognitive context of self-reported measures of subjective well-being. *Social Indicators Research, 28*, 1-20, DOI: 10.1007/BF01086714

Rässler, S. (2012). *Statistical matching: A frequentist theory, practical applications, and alternative Bayesian approaches*. New York: Springer Science & Business Media.

Rüber, I. E., Rees, S.-L., & Schmidt-Hertha, B. (2018). Lifelong learning – lifelong returns? A new theoretical framework for the analysis of civic returns on adult learning. *International Review of Education*, 64, 543–562. doi: 10.1007/s11159-018-9738-6

Schmidt, B. (2009). *Weiterbildung und informelles Lernen älterer Arbeitnehmer: Bildungsverhalten. Bildungsinteressen. Bildungsmotive*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Schuller, T. (2017). What are the wider benefits of learning across the life course? Future of Skills & Lifelong Learning. *Government Office for Science, UK: London.* Retrieved April 20, 2018, from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/635837/Skills_and_lifelong_learning_-_the_benefits_of_adult_learning_-_schuller_-_final.pdf

Simonson, J., Gordo, R., & Kelle, N. (2012). Statistical matching of the German Aging Survey and the Sample of Active Pension Accounts as a source for analyzing life courses and old age incomes. *Historical Social Research*, 37, 185–210.

Sloane-Seale, A., & Kops, B. (2008). Participation and successful aging. *Canadian Journal of University Continuing Education*, 34, 37–62.

Tam, M. (2013). Perceptions of successful ageing and implications for late-life learning. *Studies in Continuing Education*, 35, 281–298. doi: 10.1080/0158037X.2012.746228

Tippelt, R., Schmidt, B., Schnurr, S., Sinner, S., & Theisen, C. (2009). *Bildung Älterer: Chancen im demografischen Wandel*. *DIE spezial*. Bielefeld: Bertelsmann. doi: 10.3278/85/0009w

Van der Laan, P. (2000). Integrating administrative registers and household surveys. *Netherlands Official Statistics,* 15*,* 7–15.

Van Hattum, P., & Hoijtink, H. (2008). The proof of the pudding is in the eating. Data fusion: An application in marketing. *Journal of Database Marketing & Customer Strategy Management*, 15, 267-284. doi: 10.1057/dbm.2008.24

Yamashita, T., Bardo, A. R., Liu, D., & Yoo, J. W. (2019). Education, lifelong learning and self-rated health in later life in the USA. *Health Education Journal*, 78, 328-339. doi: 10.1177%2F0017896918809500