

Evaluating the utility of linked administrative data for nonresponse bias adjustment in a piggyback longitudinal survey

Büttner, Tobias J. M.; Sakshaug, Joseph W.; Vicari, Basha

DOI:
[10.2478/JOS-2021-0037](https://doi.org/10.2478/JOS-2021-0037)

License:
Creative Commons: Attribution-NoDerivs (CC BY-ND)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Büttner, TJM, Sakshaug, JW & Vicari, B 2021, 'Evaluating the utility of linked administrative data for nonresponse bias adjustment in a piggyback longitudinal survey', *Journal of Official Statistics*, vol. 37, no. 4, pp. 837-864. <https://doi.org/10.2478/JOS-2021-0037>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Evaluating the Utility of Linked Administrative Data for Nonresponse Bias Adjustment in a Piggyback Longitudinal Survey

Tobias J.M. Büttner¹, Joseph W. Sakshaug¹, and Basha Vicari¹

Nearly all panel surveys suffer from unit nonresponse and the risk of nonresponse bias. Just as the analytic value of panel surveys increase with their length, so does cumulative attrition, which can adversely affect the representativeness of the resulting survey estimates. Auxiliary data can be useful for monitoring and adjusting for attrition bias, but traditional auxiliary sources have known limitations. We investigate the utility of linked-administrative data to adjust for attrition bias in a standard piggyback longitudinal design, where respondents from a preceding general population cross-sectional survey, which included a data linkage request, were recruited for a subsequent longitudinal survey. Using the linked-administrative data from the preceding survey, we estimate attrition biases for the first eight study waves of the longitudinal survey and investigate whether an augmented weighting scheme that incorporates the linked-administrative data reduces attrition biases. We find that adding the administrative information to the weighting scheme generally leads to a modest reduction in attrition bias compared to a standard weighting procedure and, in some cases, reduces variation in the point estimates. We conclude with a discussion of these results and remark on the practical implications of incorporating linked-administrative data in piggyback longitudinal designs.

Key words: Attrition; auxiliary data; between-wave events; panel survey; weighting.

1. Introduction

Unit nonresponse is an important component of the Total Survey Error framework (e.g., Groves et al. 2009). If respondents are a non-random subgroup of the sample, then population estimates can be biased. For panel surveys, the risk of nonresponse bias increases with every subsequent wave as attrition occurs. To better understand mechanisms of attrition, monitor, and possibly mitigate the effects of bias, survey methodologists are reliant on auxiliary data. Yet, rich individual-level auxiliary data are rarely available for

¹ Federal Employment Agency/Institute for Employment Research, 104 Regensburger Straße, Nuremberg 90478, Germany. Emails: tobias.buettner@arbeitsagentur.de, joe.sakshaug@iab.de and basha.vicari@iab.de

Acknowledgments: This study uses data from the National Educational Panel Study (NEPS): Starting Cohort 6 – Adults, DOI:10.5157/NEPS:SC6:9.0.1. From 2008 to 2013, NEPS data were collected as part of the Framework Programme for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, the NEPS survey is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network. This study also uses the weakly anonymous ALWA survey data linked to administrative data of the IAB (ALWA-ADIAB 7509 v1). Instead of the original administrative data of ALWA-ADIAB, we have used updated data from the Integrated Employment Biographies (IEB v13.1). Data access was provided by the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB). We thank FDZ and the department for Data and IT-Management (DIM) at IAB for their help.

both respondents and nonrespondents in telephone and address-based samples (Smith 2011). In panel surveys, potentially relevant information can be drawn from previous study waves to describe both groups. However, more up-to-date data covering the period in-between study waves might be more informative of the underlying response mechanisms. For example, losing one's job in-between waves might have a detrimental effect on subsequent wave response, resulting in the underestimation of labor market transitions. A promising area of research is to investigate the effects of between-wave events on attrition.

In this article, we investigate the utility of federal administrative records as an auxiliary data source for assessing and adjusting for panel nonresponse bias. As discussed later, administrative data have several advantages over other auxiliary data sources. However, rarely are such data available for nonrespondents in general population surveys. Rather, such data are typically available only for respondents who give consent and can be successfully linked to the target administrative database, as is routinely done in large-scale surveys for substantive purposes (Calderwood and Lessof 2009). To overcome this limitation, we propose a methodological framework aimed at a very specific type of longitudinal design that exploits the use of existing data linkages to study attrition bias. Specifically, we focus on so-called "piggyback" longitudinal designs, where the longitudinal sample is recruited from a stand-alone, cross-sectional survey (Cohen 2005; Edwards et al. 2011). Such designs are used in several longitudinal surveys, including the US Medical Expenditure Panel Survey-Household Component (MEPS-HC), which is subsampled from participants of the US National Health Interview Survey (Ezzati-Rice et al. 2008); the English Longitudinal Study of Ageing (ELSA), which is sampled from the Health Surveys for England (Taylor et al. 2007); and the GESIS Panel, which recruits refreshment samples from the German General Social Survey (Schaurer and Weyandt 2018).

Next to cost saving opportunities, the piggyback design can be useful for informing nonresponse adjustments in subsequent longitudinal surveys, as any information collected from the piggyback survey – including any performed data linkages – can describe the sample more comprehensively than many variable-poor sampling frames (Cohen 2005). Of course, this is predicated on the representativeness of the piggyback survey, which is typically assumed by applying nonresponse adjustment weights, which are carried over and further adjusted in the subsequent longitudinal survey. We examine whether administrative data linked to respondents in a piggyback survey are useful for assessing and adjusting for nonresponse bias over eight waves of a subsequent longitudinal survey of the general population. By doing so, we contribute to the relatively sparse literature on administrative auxiliary data and current- and between-wave events affecting nonresponse in longitudinal surveys.

2. Background

2.1. Attrition in Longitudinal Surveys

Experience shows that many longitudinal surveys suffer from a significant drop in participation rates over time, particularly in the initial waves (Watson and Wooden 2009; Sakshaug and Huber 2016). For example, the German Socio-Economic Panel (GSOEP) and the British Household Panel Survey (BHPS) experienced a loss of about one third of

their initial samples after eight annual waves (Spieß and Kroh 2004; Taylor et al. 2010). Such loss threatens the precision and accuracy of survey estimates, especially if the attrition is selective. Lepkowski and Couper (2002) suggest a theoretical framework distinguishing three conditional outcomes of obtaining a response in a longitudinal survey: location, contact, and cooperation of a sample unit. All three are assumed to be influenced by survey design features (e.g., mode, time between waves, number of contact attempts) and characteristics of the sample unit (e.g., propensity to move, at-home-patterns). Couper and Ofstedal (2009) focus on moving as an important determinant of location propensity, which may be influenced by societal-level factors (e.g., urbanization) and person-level factors (e.g., housing situation). Further, researchers discuss timing issues, emphasizing the relationship of a subject's *current* circumstances and their propensity of response. The likelihoods of moving, being located, and cooperating with the survey request tend to correlate with current job status, educational paths, and family or health circumstances, among others (Couper and Ofstedal 2009; Lemay 2009). Lemay (2009) distinguishes two ways in which life events might influence attrition: a *sociodemographic* explanation, which suggests that individual characteristics are inherently associated with attrition, and a *psychosocial* explanation that focuses on the “shock” caused by a disruptive event (e.g., moving after having lost a job, a change in household size) that affects later participation. To investigate these relationships and adjust for selective attrition, it is necessary to have auxiliary data that go beyond survey data.

2.2. Auxiliary Data Sources for Nonresponse and Attrition Adjustment

Auxiliary data are most effective for nonresponse bias adjustment when they are associated with both the propensity to respond and the substantive survey variables. Little and Vartivarian (2005) suggest that the association with the survey variables is more important, as it can also reduce the variances of (weighted) survey estimates. One often-used auxiliary data source is paradata – data about the survey process (Kreuter 2013), such as call record data or interviewer observations of the household/neighborhood. While these data are moderately associated with response propensities, their associations with survey variables are rather weak (Lin and Schaeffer 1995; Kreuter et al. 2010b; Kreuter and Kohler 2009; Sakshaug and Antoni 2019; West et al. 2014). Another possibility is to link commercial data to the sampling frame. For example, West et al. (2015) evaluate the utility of two commercial databases for nonresponse adjustment in the US National Survey of Family Growth and report only minor improvements compared to a paradata-only adjustment scheme (see also, Sinibaldi et al. 2014). Smith and Kim (2013) and Smith (2011) note several limitations of commercial databases, including outdated or inconsistent information and scarce documentation about the curation and quality of the data.

In the context of panel attrition, previous-wave survey data, in addition to paradata, are commonly used for adjustment (Kroh 2010; Taylor et al. 2010), as they are generally available for both respondents and nonrespondents at the current wave. A major advantage of these data is their strong correlation with the survey variables in the attrition-affected wave. However, a limitation is that they do not measure between-wave events or individuals' current circumstances that affect response propensity nor are they expected to

be strongly correlated with variables measuring change. Attrition models underlying these weighting schemes then assume that nonresponse is independent of these (unobserved) events and circumstances (Hoonhout and Ridder 2019).

2.3. *Linked Administrative Data as an Auxiliary Data Source*

Individual-level administrative data exhibit some promising features that might overcome some of the aforementioned limitations of other auxiliary sources. Although administrative data are not designed for research purposes, researchers use them extensively because they often contain detailed substantive information (e.g., welfare receipt, employment status, program participation, healthcare utilization). Typically, administrative data are generated longitudinally, which makes them a viable source for studying biographical changes, program evaluation, or simply as a complement to surveys to lower the burden of data collection (Olson 1999; Scholz et al. 2006; Antoni and Bethmann 2019). Besides substantive research, linked-administrative data are also used for methodological purposes, such as the assessment of nonresponse and measurement errors (Kreuter et al. 2010a; Meyer and Mittag 2019) and for improving survey data that are affected by these errors (Davern et al. 2019).

For administrative data to be useful, however, researchers must establish a case-by-case link with the survey sample. This step is technically straightforward when a unique identifier (e.g., Social Security number) is available, such as in countries that use population registers as sampling frames (UNECE 2007). If no unique identifier is available, then indirect linkage techniques (e.g., probabilistic linkage) are an alternative (Christen 2012; Sakshaug et al. 2017). The utility of the linked-data additionally depends on at least two further criteria. First, legal and ethical regulations may require that respondents provide informed consent to link their responses to the administrative data (Calderwood and Lessof 2009). However, without a 100% consent rate linked-data estimates are at risk of bias. Consent biases have been identified in several studies (e.g., Young et al. 2001; Knies et al. 2012); though, Sakshaug and Kreuter (2012) show that non-consent biases are generally small compared to nonresponse and measurement biases.

The second criterion is that all sample units are present in the target administrative database. If this is not the case, then the linked sample cases are only representative of the subpopulation that overlaps with the administrative population. For example, Meyer and Mittag (2019) report an imperfect overlap in a linkage of the New York State sample of the 2008–2013 Current Population Survey Annual Social and Economic Supplement (CPS-ASEC) to administrative records from the Office of Temporary and Disability Assistance and the Department of Housing and Urban Development. The authors correct for this shortcoming using an inverse probability weighting procedure. Sakshaug et al. (2017) investigate various linkage procedures between the German “Labor Market and Social Security” (PASS) survey and an administrative employment database of the German Federal Employment Agency (BA). They report varying linkage rates depending on self-reported employment status and age or sex, which is plausible since some subgroups (e.g., self-employed, civil servants, retired persons) are beyond the responsibility of the BA. In short, both non-consent and non-overlapping populations can lead to linkage bias if differences in linkage rates correlate with the linked-outcomes of interest.

2.4. *Linked Administrative Data for Attrition Adjustment in a Longitudinal Setting*

Given the increasing trend of linking surveys with administrative data (Couper 2017), there is a potential to leverage these data for nonresponse adjustment. Exploration of this topic, however, is mainly limited to cross-sectional studies. Sakshaug and Antoni (2019) found promising correlation patterns between linked administrative variables and the response and survey variables, but adding these auxiliary variables into the nonresponse weighting scheme had only minor impact on survey estimates. Bee et al. (2015) report mixed evidence for nonresponse bias adjustment depending on the variable of interest using CPS data enriched by tax records. The present study builds on these previous studies by evaluating the utility of linked-administrative data to assess and correct for nonresponse bias in a longitudinal setting. As previously discussed, it is plausible that events occurring between study waves or situational factors at the time of the current wave relate to attrition. This information is generally unavailable from standard paradata or prior-wave survey data, but utilizing linked administrative data that follow the life course of all (or most) sample members, including attriters, might be a viable source of this information.

Only few studies have considered auxiliary administrative data to evaluate the effects of between-wave events and current-wave status on attrition (Neukirch 2002; Trappmann et al. 2015). For instance, Trappmann et al. (2015) study the effect of changes in employment status, basic income support, moving, and household composition on dropping out within the first three waves of the PASS survey. They report positive effects of between-wave moving and ending benefit receipt on attrition propensities. Further, they showed that the existing weighting scheme eliminated biases in estimates of change in the linked variables and reduced attrition biases reasonably well.

2.5. *Research Questions*

Previous studies using linked-administrative data to investigate attrition are mainly based on register samples often representing special populations (e.g., benefit recipients). In contrast, non-register samples of the general population are rarely linked to administrative data for the purpose of studying nonresponse bias in surveys (for exceptions, see Sakshaug et al. 2017; Bee et al. 2015) as they usually lack unique identifiers to facilitate direct linkage. Nevertheless, many general population surveys, including cross-sectional surveys used for piggybacking longitudinal surveys, attempt to (indirectly) link administrative data to respondents for substantive research purposes (Antoni and Seth 2011; Freedman et al. 2014; Knies and Burton 2014; Korbmacher and Czaplicki 2013). We propose and evaluate the potential for exploiting these existing linkages specifically for the purpose of monitoring and adjusting for attrition in piggyback longitudinal designs.

We illustrate this new framework using the first eight waves of the National Educational Panel Study (NEPS) – Adult Cohort, a preceding cross-sectional forerunner survey, and federal administrative data in Germany. Since this particular piggyback design involves a selection step between the preceding cross-sectional survey and the subsequent longitudinal survey in the form of panel willingness consent, we consider the magnitude of this bias source and the utility of administrative data for its adjustment. In addition, we investigate whether incorporating the administrative data into the standard NEPS

weighting scheme improves estimation and reduces attrition bias. Specifically, we address the following research questions (RQ1–4):

1. To what extent do linked administrative variables describing a sample member's current status and/or changes in one's status before the attempted interview correlate with the response outcome and substantive survey variables? How do these correlations compare to standard weighting variables used for attrition adjustment, that is, paradata or previous-wave survey information?,
2. Does the inclusion of these administrative variables improve model fit in the panel willingness model and the selection models used for attrition adjustment?,
3. Does the inclusion of these administrative variables in a standard weighting scheme reduce panel willingness and attrition biases compared to the conventional weighting variables?, and
4. To what extent, if any, do weighted survey estimates differ depending on whether administrative variables are used in the weighting scheme? Is there evidence of a reduction in attrition bias for substantive estimates?

3. Data and Methods

3.1. *The National Educational Panel Study (NEPS) – Adult Cohort*

The NEPS is a multicohort longitudinal study that follows six cohorts, each representing different stages of educational and professional pathways in Germany (Blossfeld et al. 2011). The NEPS collects detailed information on education, competence development and contextual factors (e.g., family and peers, educational institutes attended, the workplace). We use data from the NEPS – Adult Cohort (SC6). SC6 consists of three subsamples: a cross-sectional sample from the preceding forerunner survey “Working and Learning in a Changing World” (ALWA), an augmentation sample, and a refreshment sample (Hammon et al. 2016). We use the ALWA sample only. This sample comprises individuals living in private households in Germany born between 1956 and 1986. The ALWA study was designed as an independent cross-sectional study to be used as a piggyback for the NEPS. The ALWA shows some substantive overlap with the NEPS since the investigators anticipated that it would lead to a longitudinal survey before funding was secured. Only ALWA participants who explicitly provided “panel willingness” consent were later recruited for the NEPS (Antoni et al. 2010). For the exact wording of the consent question, see Online supplemental data. The ALWA sample was drawn in 2005 in a two-stage process, with 250 municipalities drawn proportional to their size and a systematic sampling of 152 persons drawn from selected municipality records. Telephone data collection took place between 2007 and 2008. Among 9,649 eligible ALWA respondents, 8,997 (93.2%) expressed willingness to participate in a follow-up interview. This piggybacking step forms the entire eligible sample of the ALWA-portion of the NEPS.

For our case study, the NEPS piggyback sample diminishes for two reasons. First, due to technical reasons, 187 panel-willing individuals who requested Turkish or Russian language interviews could not be linked. For this reason, all 227 (187 plus 40 panel refusers) individuals with Turkish or Russian CATI were discarded in advance resulting in

9,422 eligible ALWA respondents. Since studies have shown that proficiency in the host country’s language (Burkam and Lee 1998) and being a member of an ethnic minority (Lepkowski and Couper 2002; Hammon et al. 2016) have a negative effect on the propensity to respond, this might be a shortcoming. However, we assume its potential effect on the results is small as the loss amounts to only 2% of the sample. Second, to address the research questions and fully exploit the piggyback design, only survey units that can be successfully linked to administrative records in the ALWA are used. Among the 9,422 eligible respondents, 8,635 (91.7%) consented to record linkage, which is slightly higher than other consent rates reported in Germany (Sakshaug and Kreuter 2012). Among the panel-willing eligible respondents – the subgroup that is relevant for piggybacking – the rate is a little higher (8,201 out of 8,810, or 93.1%). For 7,460 units (86.4% of all consenters) linkage was achieved using a combination of deterministic and probabilistic methods (for details, see Antoni and Seth 2011). This second drop can be partially attributed to the issue of nonoverlapping populations as described earlier. After these exclusions, 7,085 of the 8,997 panel-willing respondents remained. Figure 1 summarizes all selection steps in the order they are modeled later.

By using the ALWA for piggybacking, there are eight NEPS waves available for attrition analysis. These waves stem from annual mixed-mode (telephone and face-to-face) interviewing conducted between 2009–2017. The response rates (RR1 following AAPOR 2016) drop from 73.3% in wave 1 to 63.4% in wave 2 and a decreasing decline to 43.1% in wave 8 (see Figure 2). The larger dropout in the initial waves is consistent with other longitudinal surveys, which reflects a “pruning out of the uncooperative” cases (Olsen 2018, 513). In a longitudinal piggyback design, this pruning might however take place in earlier steps such as initial nonresponse and screening for panel willingness in the preceding survey. Rather high response rates in other piggyback studies (e.g., Cheshire et al. 2012) seem to support this notion.

To evaluate the utility of auxiliary administrative data for attrition adjustment, we compare the NEPS weighting approach to the same approach augmented by linked

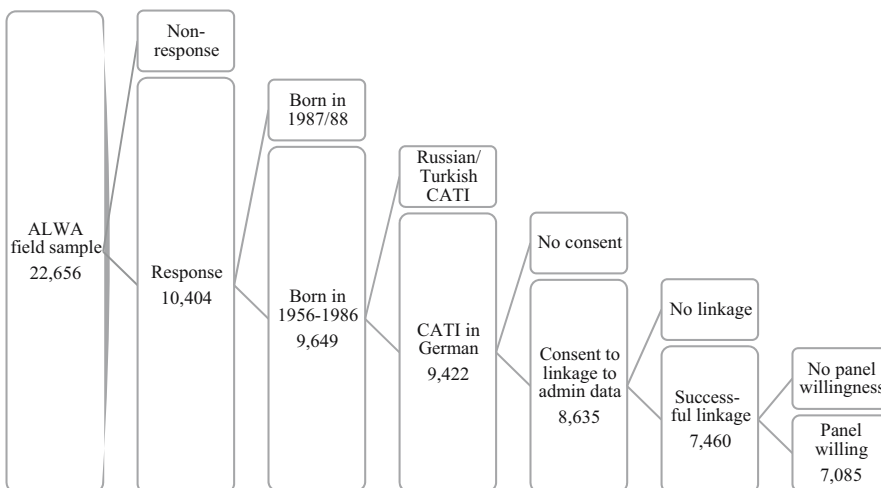


Fig. 1. Selection steps and numbers of cases from the ALWA sample to the NEPS piggyback sample.

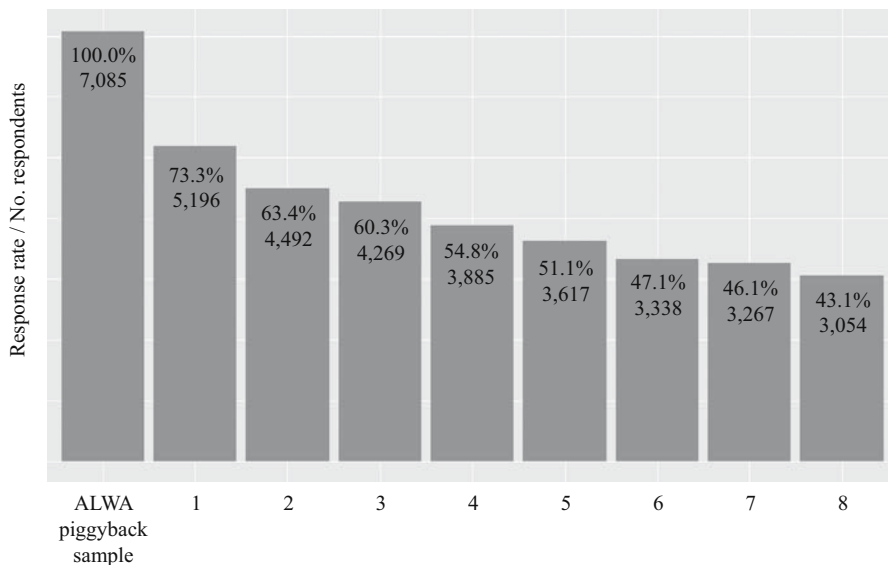


Fig. 2. Response rates by wave.

administrative variables. Therefore, the range of weighting variables are mostly predetermined. Hammon et al. (2016) and Hammon (2018) used time-constant sampling frame information of individuals' residence (federal state and BIK-10 municipality codes, which describe municipalities in Germany with regard to their size and regional structure (e.g., core, periphery, see Behrens and Wiese 2019) and survey variables (birth year, sex, marital status, mother tongue, household size, income and educational attainment). In addition, time-varying survey variables (marital status, household size and income) were used from the previous survey waves. The only paradata used for weighting is the number of contact attempts. See the Online supplemental data (Table S1) for descriptive statistics of all weighting variables. All missing values are replaced by single imputations using CART models (Burgette and Reiter 2010).

The substantive survey variables used to compare estimates under the different weighting procedures (see Subsubsection 3.3.3) are chosen based on literature research (e.g., Stöckinger et al. 2018) and discussions with data users. They include indicator variables for having children younger than six years-old in the household, for regular employment and self-employment, and indicators for working in one of three main economic sectors (WZ 2008): Manufacturing, Trade, and Health/Social Sector. Additional variables include gross monthly earned income, the number of further education courses taken in the last year, and assessment scores from the Cambridge Social Interaction and Stratification (CAMSIS) status scale. Finally, a key feature of the NEPS is the periodic collection of competence measurements (Weinert et al. 2019). We include metrics of reading, mathematics and scientific competence, vocabulary comprehension, literacy in information and communication technology (ICT), as well as reasoning skills and cognition speed (see Table S2 in the Online supplemental data for descriptive statistics of all substantive survey outcome variables).

3.2. IEB Administrative Data

The administrative data linked to the ALWA respondents comes from the Integrated Employment Biographies (IEB) of the BA. The data originate from compulsory notifications by employers to the BA and from administrative process data (Jacobebbinghaus and Seth 2007). The employer notifications are processed into individual spell data describing employment status, establishment characteristics, wages, and working time. The additional process data contain information about unemployment spells and benefit receipt, participation in active labor market programs, job search, as well as personal characteristics. As mentioned before, employment not liable to social security contributions (e.g., self-employed, civil servants) or people who never utilized one of the BA's services are not covered by the IEB. To retain all linked sample units, it is necessary to restrict the administrative variables to those generally observed for everybody, or create new variables where a missing spell in the administrative data at the time of the interview attempt clearly defines a category. For example, sample units without an employment spell in the IEB can be categorized as being "not regularly employed." Unfortunately, this rules out a few interesting variables like change of address or change in household composition, as this information is only reported to the BA for a small subgroup of current clients, mainly basic income recipients. Whenever possible, we tried to construct a proxy for these variables (e.g., started commuting to work in a different district).

Nevertheless, the IEB provides comprehensive longitudinal information on most sample units allowing to investigate their correlations with response and survey variables at, or prior to, the time of an attempted interview. For the forthcoming analyses, we use the following variables measured at the time of the current wave interview: currently employed, part-time employed, marginally employed, regular unemployment ("Unemployment Benefit I"; UBI), basic income support ("Unemployment Benefit II"; UBII), average daily wage, commuting to work in a different district, total number of employment and UBI spells in the last five years, and ever received UBI/UBII. Variables measuring changes in the individual's status between the previous and current waves include: change from not receiving UBI to receiving it (and vice versa), from non-employment to employment (and vice versa), from non-commuting to commuting to a different district, from part-time to full-time employment, and change of employer for persons who are employed in the current and previous wave. We refrained from controlling for the opposite directions of "to commuting" and "to full-time employment". The directions we control for are assumed to measure work-related changes that decrease a sample unit's availability and likelihood to respond. In this line, Trappmann et al. (2015) report that the effects of switching to full-time employment between waves on response are stronger than of switching to employment in general.

3.3. Statistical Analysis

3.3.1. Implemented Weighting Scheme

The ALWA respondents used as the sample basis for the NEPS-SC6 are those who expressed willingness to participate in a follow-up survey. In order to conserve the representativeness of the ALWA sample, the original ALWA weights are adjusted for

selection at the panel willingness step. Preceding this, adjustments for linkage consent and successful linkage to the administrative data are also performed to account for selection at these steps (the regression results modelling linkage consent and successful linkage (given consent) can be found in the Online supplemental data in Table S3). The attrition weights are constructed by replicating the weighting scheme of the NEPS, that is, using the same variables, the same approach to temporary dropout and so on (for details, see [Hammon et al. 2016](#) and [Hammon 2018](#)). As our focus is on longitudinal nonresponse, we omit the NEPS calibration steps. We also deviate from the original NEPS weighting scheme for waves 6 to 8, where response models are conditioned on being part of the active sample (i.e., the initial sample minus units that dropped out from the sample permanently) in wave 5. With an increasing number of permanent dropouts in the later waves, this approach departs from modelling the individual decision to respond for which we investigate the explanatory potential of up-to-date administrative information. We therefore condition each of the response models for waves 6–8 on being in the active sample of the respective waves. Each wave’s weights are calculated by multiplying the original ALWA weight with the inverse of the estimated propensities of the aforementioned outcomes and the previous waves’ response models. All models are fitted using logistic regression with positive outcomes (response, panel willing, linkage consent, successful linkage) coded as 1 (and 0 otherwise). Finally, as in the original NEPS weighting, the weights are trimmed above the 99th-percentile for each wave with trimmed probability mass distributed evenly over the non-trimmed cases ([Valliant et al. 2013](#)). To evaluate the impact of utilizing the linked-administrative variables, all selection models, beginning with the panel willingness model, are fitted twice: once using only the original NEPS weighting variables and once using these variables plus the linked-administrative variables. This produces two sets of adjustment weights, which we compare with regard to their impact on attrition bias.

3.3.2. Attrition Bias Reduction

The utility of the linked-administrative variables for attrition bias reduction is investigated in bivariate analyses by calculating absolute Pearson correlation coefficients (RQ1) and estimating multivariate logistic regression models on the response outcomes (RQ2). RQ3 is addressed by comparing mean estimates for a selection of linked-administrative variables with and without attrition adjustment ([Sakshaug and Antoni 2019](#)). Treating the administrative variables as outcome variables enables us to compare the wave t sample mean unaffected by attrition ($\bar{Y}_{n,t}$) (i.e., based on the respective waves’ linked respondents and nonrespondents) to the following three estimates: unadjusted, original NEPS adjustment, original-administrative adjustment, which are based on the linked-respondents only ($\bar{Y}_{r,t}$). Attrition bias is then calculated as:

$$\text{Attrition Bias}_t = \bar{Y}_{r,t} - \bar{Y}_{n,t} \quad (1)$$

We also report the absolute attrition bias (AAB), calculated as:

$$\text{Absolute Attrition Bias (AAB)}_t = |\bar{Y}_{r,t} - \bar{Y}_{n,t}| \quad (2)$$

We note that the unadjusted estimates $\bar{Y}_{r,t}$ and $\bar{Y}_{n,t}$ are unadjusted only with regard to later wave attrition. They are still weighted estimates with weights correcting for selectivity stemming from the piggyback design; that is, weights that adjust for the design

of the ALWA survey, the linkage steps, and panel willingness. In order to add to the generalizability of the results, the combined original-administrative weights for \bar{Y}_r are calculated without including the target administrative outcome variable in the selection models. For example, when evaluating the bias in an estimate of the administrative variable “average daily wage”, this variable is removed from the selection models used to generate the combined original-administrative weights.

3.3.3. Impact on Weighted Survey Estimates

The approach depicted in Subsubsection 3.3.2 aims to remedy the problem of \bar{Y}_n typically being unobserved. However, it is clear that any weighting scheme should be evaluated with regard to estimates of the actual survey variables of interest (RQ4). Despite the lack of population benchmarks, we additionally compare unweighted and both weighted estimates of the aforementioned substantive survey variables (see Subsection 3.1). For each variable and wave, we calculate the difference between the weighted and unweighted estimates:

$$\text{Difference between weighted and unweighted estimates}_t = \bar{Y}_{r,wtd,t} - \bar{Y}_{r,unwtd,t} \quad (3)$$

and the Absolute Difference (AD):

$$\text{Absolute Difference (AD)}_t = |\bar{Y}_{r,wtd,t} - \bar{Y}_{r,unwtd,t}| \quad (4)$$

To assess the impact of the weighting schemes on the variability of the weighted estimates, we additionally report coefficients of variation (CV) for all three estimates and the difference between CVs for both weighting schemes' estimates:

$$\begin{aligned} \text{CV change using administrative variables}_t &= CV(\bar{Y}_{r,orig+admin,t}) - CV(\bar{Y}_{r,orig,t}) \\ &= \left(\frac{SE(\bar{Y}_{r,orig+admin,t})}{|\bar{Y}_{r,orig+admin,t}|} \times 100\% \right) - \left(\frac{SE(\bar{Y}_{r,orig,t})}{|\bar{Y}_{r,orig,t}|} \times 100\% \right) \end{aligned} \quad (5)$$

A negative value indicates reduced variation in the estimate due to the combined original-administrative weights relative to the estimate based on the original weights (i.e., without using the administrative data).

All analyses are conducted using the family of “svy” commands in Stata 15.1 (StataCorp 2017). Estimates of standard errors are based on Taylor-Series linearization.

4. Results

4.1. Correlation Between Linked Administrative Variables, Response, and Survey Variables

To evaluate the utility of the administrative variables, we first examine their correlation patterns with the response outcome and the NEPS substantive survey variables (RQ1). Figure 3 depicts the absolute Pearson correlation coefficients, calculated across all eight NEPS waves (for tabular versions of all correlation coefficients, see Tables S4 and S5 in the Online supplemental data). For comparison, the same correlations are shown for the original NEPS weighting variables. Starting with the response outcome, the correlations

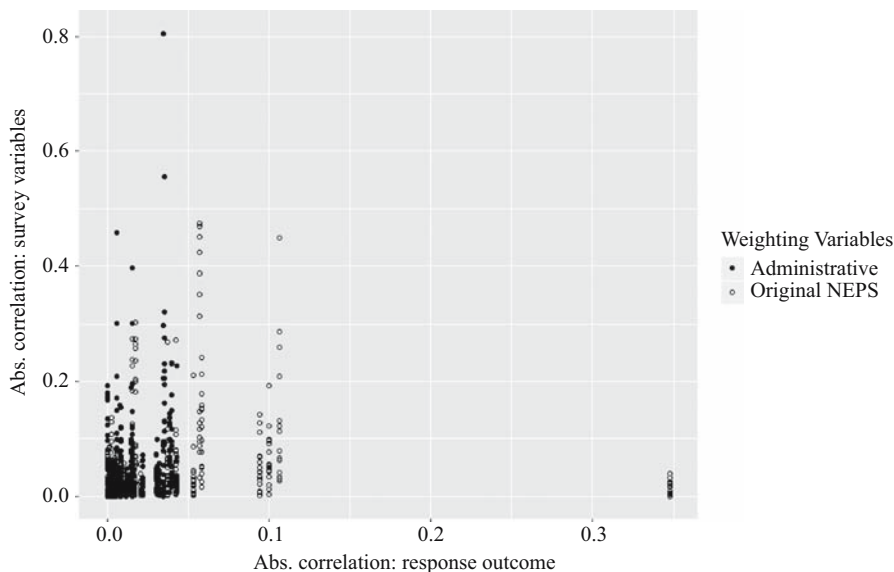


Fig. 3. Absolute Pearson correlation coefficients of the administrative and original NEPS weighting variables with the substantive survey variables and response outcome; Waves 1–8 pooled.

are generally low for both the administrative and NEPS weighting variables. The majority of correlations lie below 0.1 with all correlations involving administrative variables no larger than 0.04. The largest correlation (0.35) is observed for the paradata variable: number of contact attempts. The absolute correlations between both sets of variables and the panel willingness indicator are also very small, peaking at 0.05 for the original NEPS and 0.03 for the administrative variables (results not shown).

Around 83% (administrative variables) and 91% (original NEPS variables) of all possible correlations with the substantive survey variables are less than or equal to 0.1. Regarding the original NEPS weighting variables, the education indicators are moderately correlated (between 0.31 and 0.47) with the various competence measures and the CAMSIS score. In addition, sex, year of birth, born abroad, and mother tongue show higher correlations with the substantive survey variables. With respect to the administrative variables, there are moderate-to-high correlations between current wages (0.28) and employment (“Not employed”-category: 0.80) and their similarly-measured survey variables. The highest administrative variable correlation indicating a between-wave change is for leaving employment and the employment survey variable (0.23). The strong correlations with the substantive survey variables suggest that the administrative adjustment variables could have a decreasing effect on variances of the weighted estimates. We explore this possibility later.

4.2. Utility of Linked Administrative Variables in Panel Willingness and Response Models

Next, we compare the selection models with and without the linked-administrative variables (RQ2). The selection models are estimated separately for panel willingness and

each of the eight NEPS waves. The panel willingness model, shown in Table S6 in the Online supplemental data, shows a negative association with being born abroad and positive associations with parents' education and household income. Overall, including the administrative variables does not add much explanatory power to the already low fit of the original panel willingness model (Pseudo-R² of 0.03 in both models). The only statistically significant administrative variable is current UBII receipt, which is positively associated with panel willingness.

The response models for wave 1 indicate positive associations with the originally included variables age, gender (male), and higher education, whereas being born abroad, German mother tongue, and a high number of contact attempts have negative associations. These associations remain statistically significant after adding the administrative variables. Only one administrative variable, average daily wage, is statistically significant, having a positive association with response. All response models for waves 1 to 8 can be found in the Online supplemental data.

For brevity, Table 1 summarizes all significant administrative variables from the response models for waves 1–8. Some administrative variables are significant in multiple waves. Average daily wages, the number of UBI spells in the last five years, and having

Table 1. Summary of all statistically significant administrative predictors in NEPS waves 1–8 response models.

Administrative variable	NEPS-wave							
	1	2	3	4	5	6	7	8
Average daily wage (in EUR)	+				-			
Working hours								
Not employed	REF	REF	REF	REF	REF	REF	REF	REF
Part-time					+		+	
Full-time			+		+			
Receiving Unemployment Benefit II					+			
Commuting to work (five-digit municipality code)		+						
Number of employment spells last five years				+				
Number of UBI spells last five years			+	-				
Ever received UBII in lifetime					-			
<i>Variables of change since t-1</i>								
Became UBI recipient							+	
Became employed			-					
Left employment			-	-	-			
Stopped receiving UBII			-					
Different employer compared to last wave					+			-

Notes: +/- indicate the direction of statistically significant ($p < 0.05$) associations with survey response.

changed employers since the last interview are significant in two waves each. Their estimated coefficients, however, switch signs in their respective response models. For example, average daily wage is positively related to response in wave 1, but negatively in wave 5. Two other findings are possibly more generalizable. First, being employed (part-time or full-time) at the time of the interview attempt is positively related to response in three waves. Second, having left employment since the previous interview relates negatively to response in three waves. Note that individuals becoming UBI or UBII recipients are only a subset of all those who leave employment (about 23% of all linked sample units between waves 1 and 8). The IEB administrative data do not allow for further investigation of these changes in labor force participation.

As a sensitivity analysis, we dropped the wage variable, which is strongly correlated with the working hours indicators. This causes the indicator for working part-time in wave 5 to become insignificant, but at least one of the two working indicators is additionally significantly positive in waves 1, 3, and 6. This finding further underpins the predictive power of current employment on the response outcome. The results depicted in Subsections 4.3 and 4.4 are generally not affected from excluding the wage variable from the weighting scheme.

Table 2 summarizes two model fit statistics, Pseudo R^2 (higher is better) and Akaike's Information Criterion (AIC; lower is better), for the panel willingness model and each of the eight NEPS wave response models with and without the linked-administrative variables. Both statistics generally indicate no substantial improvements in model fit when including the linked-administrative variables.

4.3. Utility of Administrative Data for Attrition Bias Reduction

Next, we turn to the question of whether adding linked administrative variables to the original NEPS weighting procedure affects nonresponse biases over the course of the eight panel waves (RQ3). To this end, we compare estimated means and proportions of administrative variables using both sets of weights with the corresponding estimates based on the whole sample of linked sample units. The latter is therefore an estimate unaffected by initial nonresponse and attrition. Moreover, we report unweighted estimates using each waves' linked respondents and coefficients of variation of all estimates.

Before turning to attrition biases, the results of the estimated panel willingness biases and the impact of using administrative variables for their adjustment are shown. Table 3 depicts the absolute panel willingness biases for a subset of administrative and ALWA survey outcome variables (see Table S2 in the Online supplemental data for descriptive statistics for these variables). The column with the sample mean/percentage is based on all panel-willing sample units, adjusted for the design and selectivity in the consent and linkage steps. This information puts the absolute biases into perspective. They are generally very small, with the indicator variables not exceeding 0.4 percentage points. Still, both weighting schemes have a decreasing effect on the bias. Across all categorical variables, using the original NEPS weights results in an average absolute panel willingness bias of 0.2 percentage points, whereas the average bias using the original-administrative weights is slightly smaller with 0.1 percentage points. The difference between the pairs of weighted estimates is very small, which is why the original-administrative panel willingness weights are used for the remainder of the

Table 2. Model fit statistics for panel willingness models and each NEPS wave response models on original NEPS and administrative weighting variables.

	Model fit statistics	Panel willingness							
		1	2	3	4	5	6	7	8
Original NEPS	Pseudo R ²	0.07	0.14	0.12	0.12	0.11	0.14	0.12	0.10
	AIC	7714.5	4349.9	3132.7	3282.3	3760.9	3612.3	2940.1	2541.6
Original NEPS + administrative data	Pseudo R ²	0.08	0.14	0.14	0.13	0.12	0.15	0.13	0.11
	AIC	7718.4	4362.7	3119.3	3292.7	3756.6	3627.7	2943.4	2558.6

Table 3. Absolute panel willingness biases under both weighting schemes.

Data	Variable	Sample mean/ percentage	Absolute panel willingness bias		
			No weights	Weighted (orig)	Weighted (orig + admin)
Administrative data	Currently employed (%)	73.96	0.3	0.2	0.1
	Currently marginally employed (%)	12.24	0.1	0.1	0.1
	Receiving UB II (%)	5.75	0.0	0.1	0.0
	Ever received UB I in lifetime (%)	51.78	0.4	0.3	0.1
	Ever received UB II in lifetime (%)	9.71	0.1	0.1	0.1
	Commuting to work (%)	28.70	0.3	0.3	0.1
	Average daily wage (in EUR)	56.28	0.3	0.1	0.0
	Number of employment spells last five years	4.73	0.0	0.0	0.0
ALWA survey	Currently employed (%)	67.73	0.4	0.3	0.2
	Currently self-employed (%)	12.45	0.4	0.4	0.4
	Children younger than six years in household (%)	16.63	0.0	0.1	0.1
	Net monthly income (in EUR)	1672.26	6.2	0.6	1.2

Notes: Changes in the coefficient of variation after including administrative variables in the weighting procedure are all close to zero and not shown; biases for indicator variables in percentage points.

attrition analysis – a larger difference already at this point would otherwise question whether differences in later waves’ estimates are driven by attrition bias.

Table 4 summarizes the effects of using no weights, the original and the original-administrative weights on attrition bias and variance across all waves for selected variables (see Tables S15 to S26 in the Online supplemental data for results in greater detail). We skipped the analysis for the variable “Number of UBI Spells last five Years”, since already one other variable relates to the receipt of UBI in the past. Moreover, we dropped all variables with estimated proportions lower than 0.05 in the fully linked sample. The absolute biases for these were quite small and differences between both weighting schemes negligible. Averaged over all waves, estimates for most (7 out of 11) variables are less biased using the combined original-administrative adjustment. In two cases, the absolute bias increases. Estimated mean variances slightly decrease in all but one case.

Figure 4 depicts the development of absolute biases from wave 1 to wave 8 with the unweighted scenario giving an impression of attrition bias present in the panel. It highlights some considerable biases from attrition, often in the later waves. For example, estimates of average daily wage in wave 8 based on the full or the realized sample differ in absolute terms by almost EUR 5 (with a full sample mean of EUR 78.30). Other variables with larger biases are current employment and having ever received UBI or UBII. Using the additional administrative data clearly lowers AABs for the two variables employment

Table 4. Change in absolute attrition biases (AAB) and coefficients of variation (CV) under both weighting schemes; averages over eight waves.

Variable	Sample mean/ percentage All wave average	Absolute attrition bias (AAB) All wave average			CV change using administrative variables All wave average
		No weights	Weighted (orig)	Weighted (orig + admin)	
Currently employed (%)	76.38	1.5	1.4	0.8	-0.3
Currently marginally employed (%)	11.28	0.3	0.6	0.5	-0.1
Receiving UB II (%)	4.51	0.2	0.5	0.9	+0.0
Ever received UB I in lifetime (%)	56.61	1.3	1.2	0.9	-0.1
Ever received UB II in lifetime (%)	14.08	1.7	0.7	0.7	-0.1
Commuting to work (%)	30.97	0.7	0.9	0.6	-0.1
Average daily wage (in EUR)	69.88	3.2	3.5	1.9	-0.0
Number of employment spells last five years	4.75	0.1	0.1	0.1	-0.0
Became employed (%)	4.87	0.3	0.4	0.4	-0.0
Left employment (%)	4.58	0.5	0.6	0.5	-0.0
Started commuting to workplace (%)	3.93	0.3	0.5	0.4	-0.0
Different employer com- pared to last wave (%)	8.09	0.6	0.7	0.8	-0.1

Notes: CV change in percentage points, negative value means CV is reduced under the original-administrative weighting scheme; AAB for indicator variables in percentage points.

and daily wages, but for the remaining variables there is little difference between both weighting schemes.

Table 5 summarizes the precision effects of introducing administrative variables differentiating between waves. The results suggest, on average, small decreases in the estimates' sampling variance in later waves compared to the original weighting scheme.

4.4. Impact of Administrative Data on Weighted Survey Estimates

Finally, we assess the impact of the original-administrative weighting scheme on estimates of actual survey variables (RQ4). Altogether, we estimate proportions and means of 16 survey variables, nine of them over all eight waves. The other seven variables stem from competence measurements that were carried out in selected waves only. As an example, Table 6 shows the estimated proportion of regularly employed persons in the NEPS population. As a first result, each wave's estimated proportions based on no attrition adjustment, the original NEPS weighting scheme, and the combined original-administrative weights are rather similar. In every wave, all three 95% confidence

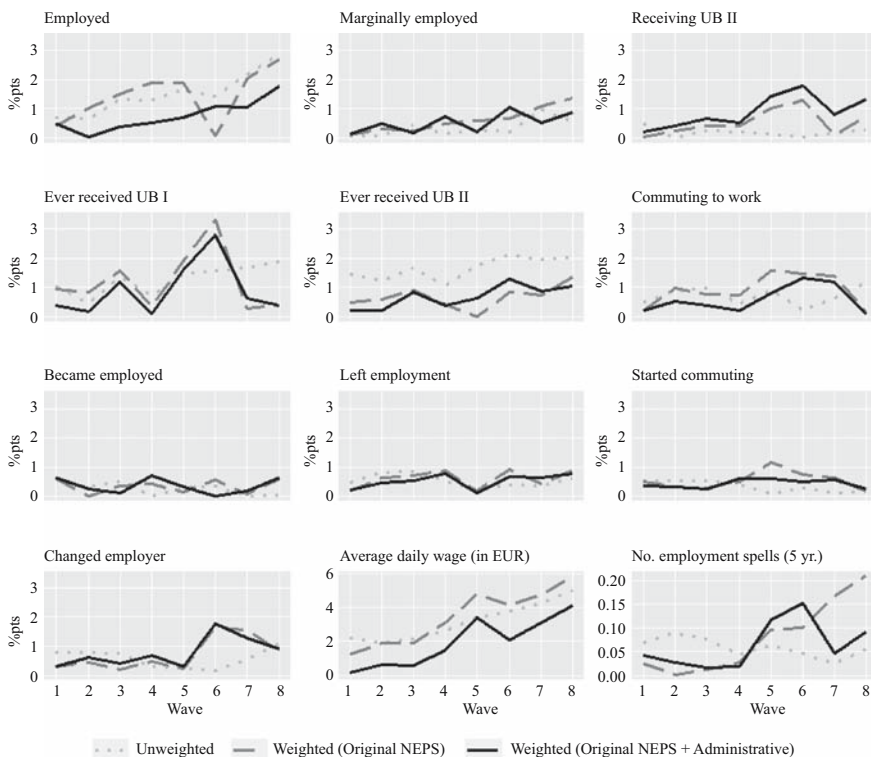


Fig. 4. Absolute attrition biases (AAB) by wave and weighting scheme.

Table 5. Average changes in coefficients of variation (CV) using the original-administrative weighting adjustment compared to the original NEPS adjustment by wave; averages calculated over 11 variables.

NEPS-Wave	1	2	3	4	5	6	7	8
CV change using administrative variables	+0.0	+0.0	+0.0	-0.0	-0.1	-0.1	-0.1	-0.2

intervals of the estimates overlap. This overlap can also be found for the remaining 15 variables (see Tables S27 to S35 in the Online supplemental data). From this we conclude, that including administrative variables in the weighting scheme does not seem to produce substantially different results.

Table 6 also shows the difference between the unweighted estimates and the original- and original-administrative-weighted estimates. For the regular employment variable, using the combined original-administrative adjustment results in slightly larger deviations from the unweighted estimate. As already mentioned before, we generally do not have benchmark estimates unaffected by attrition. However, for this variable correlation patterns shown in Subsection 4.1 imply similarity in measurement to the employment variable in the administrative data. In both cases, using the original-administrative weight shifts the estimated employment rate downwards (see Table S15 in the Online supplemental data). The smaller bias in the administrative counterpart lends support to the assumption that point estimates of the survey estimates are less biased when using the combined weighting scheme. Still, this reasoning is restricted to the employment variable, since we do not have

Table 6. Differences in estimated proportions of regular employment and coefficients of variation (CV) between both weighting schemes.

NEPS-Wave	Linked respondents	Linked respondents		Difference between weighted and unweighted estimates		% CV		CV change using admin
		No weights	Weighted (orig)	Weighted (orig)	Weighted (orig + admin)	No weights	Weighted (orig)	
1	5196	0.76 (0.75, 0.77)	0.76 (0.74, 0.77)	0.00	-0.01	2.9	3.0	-0.0
2	4492	0.77 (0.75, 0.78)	0.77 (0.76, 0.79)	0.00	-0.01	3.2	3.5	-0.0
3	4269	0.79 (0.78, 0.80)	0.80 (0.78, 0.81)	0.00	-0.01	3.5	3.9	-0.0
4	3882	0.79 (0.77, 0.80)	0.79 (0.77, 0.80)	0.00	-0.01	3.6	4.2	-0.1
5	3617	0.80 (0.78, 0.81)	0.80 (0.77, 0.82)	0.00	-0.01	3.8	6.4	-0.5
6	3338	0.79 (0.78, 0.81)	0.78 (0.75, 0.81)	-0.01	-0.02	4.0	6.7	-0.4
7	3267	0.80 (0.78, 0.81)	0.80 (0.77, 0.83)	0.01	-0.01	4.1	7.4	-0.5
8	3054	0.81 (0.79, 0.82)	0.79 (0.76, 0.82)	-0.01	-0.03	4.3	8.2	-1.2

Notes: 95% confidence intervals of the estimates in parentheses.

further similarly-measured administrative counterparts for the remaining survey variables. If the differences in the estimates between both weighting schemes were not generally small (see Table 7 for a summary of all 16 survey variables), then subject matter expertise could be a means to conclude whether the estimates might be affected by attrition bias and which weighting scheme is more effective at reducing the bias.

Finally, the results in Subsection 4.3 suggested small decreases in the estimates' CVs under the original-administrative weighting scheme. Table 8 summarizes the changes in coefficients of variation over the nine survey variables for which we have data from each wave. Across all waves, the largest mean change in percentage points is -0.3, which occurs in NEPS-wave 8. This suggests that time-varying and up-to-date variables derived from the administrative data might be more correlated with the later waves' survey variables than are the mostly time-constant original weighting variables. Looking at wave-specific correlation patterns as in Subsection 4.1 does not however yield clear-cut evidence for this. The larger change in average CV in wave 8 seems to be driven, for the most part, by the regular employment variable (see Table 6). In general, effects on variances are small.

Table 7. Means and percentages of survey outcomes, and absolute differences to unweighted estimates (AD) under both weighting schemes; all available wave averages.

Variable	Unweighted mean/percentage All available wave average	Absolute difference (AD) All available wave average	
		Weighted (orig)	Weighted (orig + admin)
Regular employment (%)	78.75	0.55	1.28
Self-employed (%)	14.09	1.39	0.80
Economic sector: manufacturing (%)	21.36	0.38	0.52
Economic sector: trade (%)	7.94	0.32	0.45
Economic sector: health/social (%)	11.75	0.64	0.66
Having children younger than six years (%)	13.68	1.20	1.24
Gross monthly earned income (in EUR)	3054.32	60.15	84.00
Number of further education courses last year	0.72	0.04	0.04
CAMSIS status score	52.57	0.50	0.35
Math score	-0.11	0.05	0.04
Reading score	-0.03	0.05	0.04
ICT literacy score	0.08	0.00	0.01
Scientific literacy score	0.02	0.03	0.04
Vocabulary comprehension	73.42	1.16	0.97
Reasoning	8.75	0.04	0.01
Cognition speed	34.70	0.24	0.35

Notes: AD for indicator variables in percentage points.

Table 8. Average changes in coefficients of variation (CV) using the original-administrative weighting adjustment compared to the original NEPS adjustment by wave; Averages calculated over nine survey variables.

NEPS-Wave	1	2	3	4	5	6	7	8
CV change using administrative variables	+0.0	-0.0	+0.0	+0.0	-0.1	+0.0	-0.1	-0.3

5. Discussion

The aim of this study was to evaluate whether a specific type of longitudinal survey design (a so-called piggyback design) has, besides various cost saving opportunities, potential for improving nonresponse and attrition adjustments through the use of existing linked-administrative data. We investigated four research questions using a combination of a general population panel survey (the NEPS) and federal administrative data. These administrative data include detailed longitudinal labor market and social security related biographies. If not already part of the sampling frame, such auxiliary administrative data are typically unavailable for initial nonrespondents and attriters of a longitudinal survey. Here, however, the sample was recruited from a preceding cross-sectional survey (the ALWA), where consenting respondents' interview data had already been linked to the administrative data. We deem this setting of using independent cross-sectional surveys as a basis to recruit subsequent longitudinal study participants, as well as the general linkage of survey and administrative data, to be relatively common practices, which are likely to become more frequent as survey budgets become more constrained. Combining both practices in order to improve the quality of longitudinal survey data is a new concept, which we evaluated in the present study.

Within our first research question (RQ1), we compared correlation patterns of the survey's original nonresponse weighting variables and a set of linked-administrative variables each with the response indicator and a selection of substantive survey outcome variables across all survey waves. With one exception, correlations of both variable sets with response were rather low. Regarding the substantive survey variables, both sets showed low-to-moderate correlations. Some of the administrative variables showed similar and even higher correlations with (construct-similar) survey variables.

RQ2 checked relationships of both variable sets with panel willingness and response in the wave-specific selection models. Adding the 16 administrative variables to each of the models showed only small gains in model fit. Although the panel willingness model yielded only few significant effects, several of the employment biography-related variables were significant in different response models across the eight study waves. Two associations proved to be relatively stable over the investigated timespan: Current employment and having left employment since the previous interview showed significant associations (positive and negative, respectively) with response in three out of eight waves. Proxy-variables for moving, such as change of employer or starting to commute to a different district, were not consistent significant predictors of response, in contrast to findings reported in related studies (e.g., [Trappmann et al. 2015](#); [Kroh 2010](#); [Short and McArthur 1986](#); [Watson and Wooden 2009](#)).

Using the administrative data as outcome variables to assess attrition bias, we compared adjustment effects of alternative weighting schemes – with and without adding administrative variables into the weighting procedure (RQ3). Although attrition biases were rather low for many of the investigated variables, we observed a clear reduction of bias for the variables most affected by attrition: either the combined original-administrative weights outperformed the original NEPS weighting procedure (for employment and wage variables) or both weighting schemes performed similarly well in reducing the bias (for Unemployment Benefit I/II receipt variables). The same applied to

the adjustment of panel willingness bias. An important result for the application of the piggyback design is that actual panel willingness biases are generally small, and both weighting schemes successfully reduce it even further. Concerning RQ4, we found that adding administrative variables into the weighting procedure had only minor effects on point estimates of substantive survey variables; again, this could be because there was little attrition bias in the survey to begin with. However, there was some evidence of a bias reduction for the survey-measured current employment status variable, which had one of the largest biases in its administrative variable counterpart. In terms of variance, we observed a tendency towards slightly smaller coefficients of variation in later waves when the administrative variables were included in the weighting scheme.

We note that the dropout in the NEPS sample was quite substantial, after eight waves only about 43% of the wave 1 sample continued to respond. With regard to the linked-administrative variables that were available for both respondents and nonrespondents, the evidence suggested substantial attrition bias for some variables. For some policy-relevant indicators (employment, marginal employment, wages), we saw an increasing trend in bias over time. For instance, the average daily wage in wave 8 was (unadjusted) estimated at EUR 83.2, which is roughly 6% more than the EUR 78.2 estimated from the full sample unaffected by attrition. As mentioned before, here the combined original-administrative weights reduce some of the bias (producing an estimate of EUR 82.3) compared to the original-weighted estimate (EUR 84.1). The remaining bias suggests, however, that the assumption of respondents and nonrespondents being equal given the auxiliary variables does not entirely hold, and that the procedure does not eliminate all bias.

With modest benefits, the utility of linked-administrative records for attrition adjustment is even more sensitive to its costs. This study simulated the application of a piggyback design, also with reference to its cost-saving potentials. In a piggyback design, with administrative data already linked to the preceding survey, the threshold to implement a weighting adjustment using the administrative data is possibly lower than if the data are not already linked. However, one must consider the potential loss of sample units due to unsuccessful linkage and panel non-willingness (e.g., [Sin 2006](#)), which could offset the benefits of augmented weighting if not properly accounted for. Moreover, it should be mentioned that administrative data can contain specific measurement error (e.g., [Pavlopoulos and Vermunt 2015](#); [Pankowska et al. 2018](#)) which might attenuate correlations that are crucial for the proposed procedures.

Assuming that these issues can be addressed, this approach is worth considering when setting up longitudinal surveys using a piggyback design that includes linked-administrative data from the preceding survey. Depending on the scope of the study, this approach seems likely to yield advantages from having rich information on initial sample units that can be used to compensate for their decreasing participation over time. As shown, augmented nonresponse weights can be useful for bias adjustment without increasing sampling variance. Moreover, we envision the suggested strategy to help with monitoring nonresponse (with regard to certain subgroups) over the course of the panel, even if the administrative variables are not included in any adjustment procedure. This could help to assess the sample's longitudinal representativeness and potentially substantiate design decisions regarding future waves or refreshment samples.

6. References

- American Association for Public Opinion Research. 2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, (9th edition). Lanexa: AAPOR. Available at: https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf (accessed October 2021).
- Antoni, M., and A. Bethmann. 2019. "PASS-ADIAB – Linked Survey and Administrative Data for Research on Unemployment and Poverty." *Journal of Economics and Statistics* 239(4): 747–756. DOI: <https://doi.org/10.1515/jbnst-2018-0002>.
- Antoni, M., K. Drasch, C. Kleinert, B. Matthes, M. Ruland, and A. Trahms. 2010. *Working and Learning in a Changing World, part I: Overview of the study, FDZ Methodenreport No. 5/2010 (en)*. Nürnberg: Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (IAB). Available at: http://doku.iab.de/fdz/reporte/2010/MR_05-10_EN.pdf (accessed October 2021).
- Antoni, M., and S. Seth. 2011. *ALWA-ADIAB – Linked individual Survey and Administrative Data for Substantive and Methodological Research, FDZ Methodenreport No. 12/2011 (en)*. Nürnberg: Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (IAB). Available at: http://doku.iab.de/fdz/reporte/2011/MR_12-11_EN.pdf (accessed October 2021).
- Bee, C.A., G.M.R. Gathright, and B.D. Meyer. 2015. "Bias from Unit Non-Response in the Measurement of Income in Household Surveys." Paper presented at the Joint Statistical Meetings of the American Statistical Association, August 9, 2015, Seattle, USA. Available at: <http://www.solejole.org/16068.pdf> (accessed February 2020).
- Behrens, K., and K. Wiese. 2019. "BIK-Regionen." In *Regionale Standards: Ausgabe 2019*, edited by K. Beckmann, K. Behrens, H. Hoffmann, M. Pfister, K. Wiese, J.H.P. Hoffmeyer-Zlotnik, G. Rösch, P. Siegers, W. Sodeur, U. Hanefeld, R. Herter-Eschweiler, and E. Krack-Roberg: 114–126. Cologne, Germany: GESIS – Leibniz-Institut für Sozialwissenschaften.
- Blossfeld, H.P., H.G. Roßbach, and J. von Maurice (Eds.). 2011. *Education as a Lifelong Process – The German National Educational Panel Study (NEPS)*, Zeitschrift für Erziehungswissenschaft: Sonderheft 14. Wiesbaden: Springer VS.
- Burgette, L.F., and J.P. Reiter. 2010. "Multiple Imputation for Missing Data via Sequential Regression Trees." *American Journal of Epidemiology* 172(9): 1070–1076. DOI: <https://doi.org/10.1093/aje/kwq260>.
- Burkam, D.T., and V.E. Lee. 1998. "Effects of Monotone and Nonmonotone Attrition on Parameter Estimates in Regression Models with Educational Data." *The Journal of Human Resources* 33(2): 555–574. DOI: <https://doi.org/10.2307/146441>.
- Calderwood, L., and C. Lessof. 2009. "Enhancing Longitudinal Surveys by Linking to Administrative Data." In *Methodology of Longitudinal Surveys*, edited by P. Lynn: 55–72. West Sussex, England: John Wiley & Sons.
- Cheshire, H., D. Hussey, J. Medina, K. Pickering, N. Wood, K. Ward, K. Taylor, and C. Lessof. 2012. *Financial circumstances, health and well-being of the older population in England: The 2008 English Longitudinal Study of Ageing: Wave 4 Technical Report*. London: National Centre for Social Research. Available at: https://ifs.org.uk/elsa/report10/w4_tech.pdf (accessed October 2021).

- Christen. 2012. *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. Berlin, Germany: Springer. DOI: https://doi.org/10.1007/978-3-642-31164-2_2.
- Cohen, S.B. 2005. "Integrated survey designs: A framework for nonresponse bias reduction." *Journal of Economic and Social Measurement* 30(2–3): 101–114. DOI: <https://doi.org/10.3233/JEM-2005-0244>.
- Couper, M.P., and M.B. Ofstedal. 2009. "Keeping in Contact with Mobile Sample Members." In *Methodology of Longitudinal Surveys*, edited by P. Lynn: 183–203. West Sussex, England: John Wiley & Sons.
- Couper, M.P. 2017. "New Developments in Survey Data Collection." *Annual Review of Sociology* 43(1): 121–145. DOI: <https://doi.org/10.1146/annurev-soc-060116-053613>.
- Davern, M.E., B.D. Meyer, and N.K. Mittag. 2019. "Creating Improved Survey Data Products Using Linked Administrative-Survey Data." *Journal of Survey Statistics and Methodology* 7(3): 440–463. DOI: <https://doi.org/10.1093/jssam/smy017>.
- Edwards, B., L. Branden, and M. Stange. 2011. "Piggyback Survey Respondents and Mode: Lessons Learned from Design and Operations." In *JSM Proceedings*, Survey Research Methods Section, August 1, 2011: 5009–5021. Alexandria, VA: American Statistical Association. Available at: http://www.asasrms.org/Proceedings/y2011/Files/302992_69483.pdf (accessed October 2021).
- Ezzati-Rice, T.M., F. Rohde, and J. Greenblatt. 2008. *Sample Design of the Medical Expenditure Panel Survey Household Component, 1998–2007. Methodology Report No. 22. March 2008*. Rockville, MD, USA: Agency for Healthcare Research and Quality. Available at: http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr22/mr22.shtml (accessed February 2020).
- Freedman, V.A., K. McGonagle, and P. Andreski. 2014. *The Panel Study of Income Dynamics: Linked Medicare Claims Data*. PSID Technical Series Paper No. 14-01. University of Michigan. Available at: https://psidonline.isr.umich.edu/publications/Papers/tsp/2014-01_PSIDMedicare.pdf (accessed October 2021).
- Groves, R., F.J. Fowler, M.P. Couper, J.M. Lepkowski, E. Singer, and R. Tourangeau. 2009. *Survey Methodology*, (2nd ed.). Hoboken, NJ: John Wiley & Sons.
- Hammon, A., S. Zinn, C. Aßmann, and A. Würbach. 2016. *Samples, Weights, and Nonresponse: The Adult cohort of the National Educational Panel Study (Wave 2 to 6)*. Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study. NEPS Survey Paper No. 7. DOI: <https://doi.org/10.5157/NEPS:SP07:1.0>.
- Hammon, A. 2018. *Samples, Weights, and Nonresponse: The Adult Cohort of the National Educational Panel Study (Wave 7 to 9)*. Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study. Available at: https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC6/9-0-0/SC6_9-0-0_W.pdf (accessed October 2021).
- Hoonhout, P., and G. Ridder. 2019. "Nonignorable Attrition in Multi-Period Panels With Refreshment Samples." *Journal of Business & Economic Statistics* 37(3): 377–390. DOI: <https://doi.org/10.1080/07350015.2017.1345744>.
- Jacobebbinghaus, P., and S. Seth. 2007. "The German Integrated Employment Biographies Sample IEBS." *Schmollers Jahrbuch: Journal of Applied Social Science*

- Studies* 127: 335–342. Available at: https://www.ratswd.de/download/schmollers/2007_127/Schmollers_2007_2_S335.pdf (accessed October 2021).
- Korbmacher, J., and C. Czaplicki. 2013. “Linking SHARE Survey Data with Administrative Records: First Experiences from SHARE-Germany.” In *SHARE Wave 4: Innovations & Methodology*, edited by F. Malter and A. Börsch-Supan: 47–53. Munich: MEA, Max Planck Institute for Social Law and Social Policy.
- Knies, G., J. Burton, and E. Sala. 2012. “Consenting to Health Record Linkage: Evidence from a Multipurpose Longitudinal Survey of a General Population.” *BMC Health Services Research* 12(52). DOI: <https://doi.org/10.1186/1472-6963-12-52>.
- Knies, G., and J. Burton. 2014. “Analysis of four studies in a comparative framework reveals: health linkage consent rates on British cohort studies higher than on UK household panel surveys.” *BMC Medical Research Methodology* 14(125). DOI: <https://doi.org/10.1186/1471-2288-14-125>.
- Kreuter, F., and U. Kohler. 2009. “Analyzing Contact Sequences in Call Record Data: Potential and Limitations of Sequence Indicators for Nonresponse Adjustments in the European Social Survey.” *Journal of Official Statistics* 25(2): 203–226.
- Kreuter, F., G. Müller, and M. Trappmann. 2010a. “Nonresponse and Measurement Error in Employment Research: Making Use of Administrative Data.” *Public Opinion Quarterly* 74: 880–906. DOI: <https://doi.org/10.1093/poq/nfq060>.
- Kreuter, F., K. Olson, J. Wagner, T. Yan, T.M. Ezzati-Rice, C. Casas-Cordero, M. Lemay, A. Peytchev, R.M. Groves, and T.E. Raghunathan. 2010b. “Using Proxy Measures and Other Correlates of Survey Outcomes to Adjust for Nonresponse: Examples from Multiple Surveys.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173: 389–407. DOI: <https://doi.org/10.1111/j.1467-985X.2009.00621.x>.
- Kreuter, F. 2013. *Improving Surveys with Paradata: Analytic Uses of Process Information*. Hoboken, NJ: Wiley.
- Kroh, M. 2010. *DIW Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2009)*. DIW Data Documentation 50. Berlin: German Institute for Economic Research. Available at: https://www.diw.de/documents/publikationen/73/diw_01.c.359697.de/diw_datadoc_2010-050.pdf (accessed October 2021).
- Lemay, M. 2009. *Understanding the Mechanisms of Panel Attrition*. PhD Dissertation. University of Maryland, College Park. Available at: <http://hdl.handle.net/1903/9631> (accessed October 2021).
- Lepkowski, J.M., and M.C. Couper. 2002. “Nonresponse in the Second Wave of Longitudinal Household Surveys.” In *Survey Nonresponse*, edited by R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little: 259–272. New York, NY: John Wiley & Sons.
- Lin, I., and N.C. Schaeffer. 1995. “Using Survey Participants to Estimate the Impact of Nonparticipation.” *Public Opinion Quarterly* 59: 236–258. DOI: <https://doi.org/10.1086/269471>.
- Little, R.J.A., and S. Vartivarian. 2005. “Does Weighting for Nonresponse Increase the Variance of Survey Means?” *Survey Methodology* 31: 161–168. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2005002/article/9046-eng.pdf?st=3epbd-uy> (accessed October 2021).

- Meyer, B.D., and N. Mittag. 2019. "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety Net." *American Economic Journal: Applied Economics* 11(2): 176–204. DOI: <https://doi.org/10.1257/app.20170478>.
- Neukirch, T. 2002. "Nonignorable Attrition and Selectivity Biases in the Finnish Subsample of the ECHP: an Empirical Study Using Additional Register Information." *Chintex Working Paper 5*. Frankfurt a. M.: J.W. Goethe-Universität. Available at: <http://hdl.handle.net/10068/92455> (accessed October 2021).
- Olsen, R. 2018. "Panel Attrition." In *The Palgrave Handbook of Survey Research*, edited by D. Vannette and J. Krosnick: 509–517. DOI: https://doi.org/10.1007/978-3-319-54395-6_59.
- Olson, J.A. 1999. "Linkages with Data from Social Security Administrative Records in the Health and Retirement Study." *Social Security Bulletin* 62(2): 73–85. Available at: <https://www.ssa.gov/policy/docs/ssb/v62n2/v62n2p73.pdf> (accessed October 2021).
- Pankowska, P., B. Bakker, D.L. Oberski, and D. Pavlopoulos. 2018. "Reconciliation of two data sources by correction for measurement error: the feasibility of parameter reuse." *Statistical Journal of the IAOS* 34(3): 317–329. DOI: <https://doi.org/10.3233/SJI-170368>.
- Pavlopoulos, D., and J.K. Vermunt. 2015. "Measuring temporary employment. Do survey or register data tell the truth?" *Survey Methodology* 41(1): 197–214. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2015001/article/14151-eng.pdf?st=tAsC-b20> (accessed October 2021).
- Sakshaug, J.W., M. Antoni, and R. Sauckel. 2017. "The Quality and Selectivity of Linking Federal Administrative Records to Respondents and Nonrespondents in a General Population Survey in Germany." *Survey Research Methods* 1: 63–80. DOI: <https://doi.org/10.18148/srm/2017.v1i1.6718>.
- Sakshaug, J.W., and M. Antoni. 2019. "Evaluating the Utility of Indirectly Linked Federal Administrative Records for Nonresponse Bias Adjustment." *Journal of Survey Statistics and Methodology* 7: 227–249. DOI: <https://doi.org/10.1093/jssam/smy009>.
- Sakshaug, J.W., and M. Huber. 2016. "An Evaluation of Panel Nonresponse and Linkage Consent Bias in a Survey of Employees in Germany." *Journal of Survey Statistics and Methodology* 4(1): 71–93. DOI: <https://doi.org/10.1093/jssam/smv034>.
- Sakshaug, J.W., and F. Kreuter. 2012. "Assessing the Magnitude of Non-Consent Biases in Linked Survey and Administrative Data." *Survey Research Methods* 6(2): 113–122. DOI: <https://doi.org/10.18148/srm/2012.v6i2.5094>.
- Schaurer, I., and K. Weyandt. 2018. *GESIS Panel Technical Report: Recruitment 2016 (Wave d11 and d12)*. Cologne, Germany: GESIS – Leibniz-Institut für Sozialwissenschaften.
- Scholz, J., A. Seshadri, and S. Khitatrakun. 2006. "Are Americans Saving 'Optimally' for Retirement?" *Journal of Political Economy* 114: 607–643. DOI: <https://doi.org/10.1086/506335>.
- Short, K., and E. McArthur. 1986. "Life Events and Sample Attrition in the Survey of Income and Program Participation." In *Proceedings of the Survey Research Methods Section of the American Statistical Association*: 200–205. Available at: <http://www.asarms.org/Proceedings/y1986f.html> (accessed October 2021).

- Sin, C.H. 2006. "The feasibility of using national surveys to derive samples of older people from different ethnic groups in Britain: Lessons from 'piggy-backing' on the Family Resources Survey." *International Journal of Social Research Methodology* 9(1): 15–28. DOI: <https://doi.org/10.1080/13645570500435264>.
- Sinibaldi, J., M. Trappmann, and F. Kreuter. 2014. "Which is the better investment for nonresponse adjustment: Purchasing commercial auxiliary data or collecting interviewer observations?" *Public Opinion Quarterly* 78(2): 440–473. DOI: <https://doi.org/10.1093/poq/nfu003>.
- Smith, T.W. 2011. "The Report of the International Workshop on Using Multi-Level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys." *International Journal of Public Opinion Research* 23: 389–402. DOI: <https://doi.org/10.1093/ijpor/edr035>.
- Smith, T.W., and J. Kim. 2013. "An Assessment of the Multi-Level Integrated Database Approach." *The Annals of the American Academy of Political and Social Science* 645: 185–221. DOI: <https://doi.org/10.1177/0002716212463340>.
- Spieß, M., and M. Kroh. 2004. "Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (GSOEP)." *Research Note No. 28a*. Berlin, Germany: German Institute for Economic Research (DIW). Available at: https://www.diw.de/documents/publikationen/73/diw_01.c.43407.de/diw_datadoc_2004-001.pdf (accessed October 2021).
- StataCorp. 2017. *Stata 15 Base Reference Manual*. College Station, TX: Stata Press.
- Stöckinger, C., S. Kretschmer, and C. Kleinert. 2018. *Panel Attrition in NEPS Starting Cohort 6: A Description of Attrition Processes in Waves 2 to 7 with Regard to Nonresponse Bias, NEPS Survey Paper No. 35*. Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Taylor, M.F., Brice, J., Buck, N., and Prentice-Lane, E. (Eds.). 2010. *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Colchester: University of Essex.
- Taylor, R., L. Conway, L. Calderwood, C. Lessof, H. Cheshire, K. Cox, and S. Scholes. 2007. *Health, wealth and lifestyles of the older population in England: The 2002 English Longitudinal Study of Ageing: Technical Report*. London: Institute of Fiscal Studies. Available at: https://ifs.org.uk/elsa/report03/w1_tech.pdf (accessed October 2021).
- Trappmann, M., T. Gramlich, and A. Mosthaf. 2015. "The Effect of Events Between Waves on Panel Attrition." *Survey Research Methods* 9(1): 31–43. DOI: <https://doi.org/10.18148/srm/2015.v9i1.5849>.
- UNECE (United Nations Economic Commission for Europe). 2007. *Register-Based Statistics in the Nordic Countries: Review of Best Practices with Focus on Population and Social Statistics*. Technical Report E.07.II.E.11. New York/Geneva: United Nations. Available at: http://www.unece.org/fileadmin/DAM/stats/publications/Register_based_statistics_in_Nordic_countries.pdf (accessed February 2020).
- Valliant, R., J.A. Dever, and F. Kreuter. 2013. *Practical Tools for Designing and Weighting Survey Samples*. New York: Springer. DOI: https://doi.org/10.1007/978-3-319-93632-1_13.

- Watson, N., and M. Wooden. 2009. "Identifying Factors Affecting Longitudinal Survey Response." In *Methodology of Longitudinal Surveys*, edited by P. Lynn: 157–181. West Sussex, England: John Wiley & Sons.
- Weinert, S., C. Artelt, M. Prenzel, M. Senkbeil, T. Ehmke, C.H. Carstensen, and K. Lockl. 2019. "Development of Competencies Across the Life Course." In *Education as a Lifelong Process: The German National Educational Panel Study (NEPS), Edition ZfE Vol. 3*, edited by H.P. Blossfeld and H.G. Roßbach: 57–81. Wiesbaden: Springer VS. DOI: <https://doi.org/10.1007/978-3-658-23162-0>.
- West, B.T., F. Kreuter, and M. Trappmann. 2014. "Is the Collection of Interviewer Observations Worthwhile in an Economic Panel Survey? New Evidence from the German Labor Market and Social Security (PASS) Study." *Journal of Survey Statistics and Methodology* 2: 159–181. DOI: <https://doi.org/10.1093/jssam/smu002>.
- West, B.T., J. Wagner, F. Hubbard, and H. Gu. 2015. "The Utility of Alternative Commercial Data Sources for Survey Operations and Estimation: Evidence from the National Survey of Family Growth." *Journal of Survey Statistics and Methodology* 3: 240–264. DOI: <https://doi.org/10.1093/jssam/smv004>.
- Young, A.F., A.J. Dobson, and J.E. Byles. 2001. "Health Services Research Using Linked Records: Who Consents and What is the Gain?" *Australian and New Zealand Journal of Public Health* 25(5): 417–420. DOI: <https://doi.org/10.1111/j.1467-842X.2001.tb00284.x>.

Received October 2020

Revised March 2021

Accepted May 2021