**Mixed Mode Effects of Web and Telephone Surveys on Measuring Employment Status**

Joachim Schork, STATEC Luxembourg, joachim.schork@statec.etat.lu

Cesare A. F. Riillo, STATEC Luxembourg, cesare.riillo@statec.etat.lu

Johann Neumayr, STATEC Luxembourg, johann.neumayr@statec.etat.lu

**Abstract**

*Web questionnaires are increasingly used to complement traditional data collection, leading to different combinations of survey modes. The flexibility of mixed modes provides many advantages such as less nonresponse issues, lowered expenditures, and compensation for the decreasing availability of other data sources, i.e. fixed-line telephone numbers. However, the increased usage of web data raises concerns whether web questionnaires lead to systematic measurement errors, since responses given to web questionnaires may be significantly different compared to other survey modes.*

*We argue that the size of mixed mode effects strongly depends on the content of a variable. We investigate differences in web and telephone data in terms of objective and subjective variables. The study is based on the Luxembourgish Labour Force Survey that collects both objective (employment status) and subjective employment variables (wage adequacy and job satisfaction). Analysis of the raw data reveals significant differences in sample composition (e.g. respondents' personal characteristics such as age or nationality) as well as in objective and subjective employment variables.*

*In order to investigate whether differences in employment variables are caused by sample composition or survey mode effects, we match web and telephone samples according to variables that lead to dissimilarities in sample composition. We identify these variables by a combination of automatic variable selection via random forest and a theory driven selection. Based on the selected variables, we then apply a Coarsened Exact Matching that approximates randomized experiments by reducing dissimilarities between web and telephone samples.*

*After matching, we show that employment status is not affected by systematic measurement bias, but web respondents report lower levels of wage adequacy and job satisfaction. Even though further research on subjective variables is advisable, our results support the implementation of mixed survey modes in official statistics such as the Labour Force Survey.*

**Keywords:** web survey, telephone survey, mode effects, Coarsened Exact Matching, Labour Force Survey

**1. Introduction**

Statistical surveys provide an important basis for evaluation and decision making across many areas. For many decades, survey data was mainly collected by interviewers (e.g. telephone interviews or face-to-face interviews) or via self-administered mail questionnaires. Traditional data collection increasingly faces challenges such as declining response rates, less availability of traditional data sources (e.g. fixed-line telephone numbers), or the demand of more timely data dissemination, leading to lowered data quality and higher data collection costs (Blanke and Luiten, 2014; Dillman, 2017).

At the same time, technological improvements and the rapidly grown number of households with internet access provide data collectors with web questionnaires as additional data source. Compared to traditional collection modes, internet surveys potentially have numerous advantages, e.g., cost-efficiency or better coverage of individuals that are difficult to access via traditional channels (Bianchi et al., 2017; de Leeuw, 2005).

The increased usage of web surveys raises concerns whether collecting data online leads to mode effects, i.e. systematic measurement bias as result of significantly different responses of web participants compared to other data collection modes. Mixed mode effects on employment variables were, for instance, investigated within the framework of a European Statistical System Network (ESSnet) project on data collection for social surveys using multiple modes. In a comparison of Labour Force Survey (LFS) data collected by web and other modes, differences in the variable employment status could mostly be explained by auxiliary variables such as sex and age (Luiten and Blanke, 2015). In contrast, other employment related variables such as working hours, status in employment, and education level revealed significant mode effects (Körner and Liersch, 2014; Schouten and van der Laan, 2014).

These varying findings suggest that the size of mixed mode effects strongly depends on the specific content of a variable. Hence, mixed mode effects need to be examined separately for different variables. By investigating mixed collection modes on the basis of Luxembourgish LFS data, we contribute to previous research about employment status and employment related variables. Our contribution is twofold:

First, the present study includes investigations on a battery of variables that is exclusively available in the LFS of Luxembourg. Contrary to typically collected employment variables (e.g. employment status or income), these variables contain subjective information (e.g. wage adequacy or job satisfaction). By investigating on these variables, we provide insights whether mode effects vary within one survey depending on the type of a variable (i.e. objective vs. subjective). Hereby, we extend previous research about employment variables that focuses mainly on objective contents.

Second, we use a different methodological approach compared to previous work about mixed mode effects on measuring employment status. Most research in this field is based on randomized experiments and re-weighting approaches (Körner and Liersch, 2014; Pohjanpää, 2014; Schouten and van der Laan, 2014). Instead, we use Coarsened Exact Matching in order to harmonize web and telephone samples of the original LFS data after collection (Iacus et al., 2012). This provides us with the possibility to investigate whether a different methodology leads to similar results compared to randomized experiments.

**2. Differences in Web and Telephone Samples**

Research repeatedly found differences in numerous variables of web and telephone data, i.e. data before correcting for unit nonresponse or coverage errors (Körner and Liersch, 2014; Sarracino et al., 2017; Schouten and van der Laan, 2014). Participants of web samples are more often males; younger; better educated; and singles. Notably, such differences were also observed in employment variables: web respondents are more often in employment; have on average a higher income; and report more often to work more than 40 hours.

Previous research already examined whether these differences between web and telephone data introduce bias in survey results. Hereby, three sources of mode-specific bias are recurrently discussed (Luiten and Blanke, 2015; Sarracino et al., 2017; Schouten and van der Laan, 2014): coverage bias (i.e. a mismatch between sampling frame and target population); nonresponse bias (i.e. differences in contactability and response behaviour); and measurement bias (i.e. the same person responds differently to different survey modes).

Different literature focusses on different types of bias, leading to inconsistent definitions of the term “mixed mode effect”. In the present study, we focus on mode-specific measurement bias and use measurement bias therefore as synonym for mode effect. Measurement bias can be caused by many different reasons such as social desirability; a tendency to respond similarly to questions of the same topic; a preference to give positive answers; and a difference in responses depending on the order of questions (see Sarracino et al. (2017) for a detailed overview).

**3. Data**

We use data of the Luxembourgish Labour Force Survey (LFS), which is designed to measure labour force participation for all civilian household members with an age of 15 years or older. Until 2015, the LFS was conducted via Random Digit Dialing. However, consistently dropping response rates led in 2015 to the introduction of a new sampling design, which is based on a mixed mode data collection consisting of Computer Assisted Telephone Interviewing (CATI) and Computer Assisted Web Interviewing (CAWI). Since 2015, the strategy for data collection is as follows: After drawing a sample from the Luxembourgish population register, it is verified if a phone number of the sampling unit can be found in the official white pages telephone directories online. Sampling units for which a telephone number is available are approached by telephone. Remaining units are approached via an invitation letter containing the internet address where the web questionnaire can be found. As result of this design, participants are not allocated randomly to web and telephone samples.

Figure 1 shows a cross tabulation of age, sex, and nationality. The share of the population of 2017 for each group is shown on the right column of the table. The two graphics illustrate weighted differences to the population of web and telephone samples of the LFS 2017, i.e. under-represented groups are displayed by a red bar and over-represented groups are displayed in green. The figure reveals mainly two systematic differences: First, age groups between 20-49 are more often collected by web and age groups 50+ are more often collected by telephone. Second, Luxembourgish people are collected more often by telephone. The impact of such differences on certain target variables is examined within the present study.

**Figure *1*. Cross tabulation of age, sex & nationality for population, web & telephone samples of the LFS 2017**



In order to base our results on a larger sample, we combine all Luxembourgish LFS data that was collected via mixed mode up to this point (i.e. 2015, 2016 and 2017). We use the reference person of each household and individuals with an age of 15-74 years, leading to a sample size of n = 57,566 with 60% web interviews.

We investigate on three target variables: employment status, wage adequacy, and job satisfaction. We consider employment status as objective variable (i.e. clear definition according to the ILO-classification of employment) and wage adequacy as well as job satisfaction as subjective variables (i.e. self-assessment of personal opinions).

Some auxiliary variables of the LFS contain missing values due to item nonresponse. We impute these missing values in the forefront of the mixed mode analysis via single imputation. We impute continuous variables by predictive mean matching and categorical variables via multinomial logistic regression (van Buuren, 2012).

**4. Method**

Differences in auxiliary variables (e.g. sex and age) may influence how respondents of the two groups (web and telephone) report target variables (e.g. employment status). To detect potential mixed mode measurement bias, we use Coarsened Exact Matching (CEM), a matching method that approximates randomized experiments by reducing dissimilarities in observed variables (Iacus et al., 2012). CEM creates strata using temporally coarsened variables (i.e. with values grouped into substantially meaningful categories) and retains observations so that strata have at least one unit of both groups. CEM provides a sub-set that has similar characteristics not only on average, but on the whole distribution of observable variables. Additionally, we adjust the CEM weights to accommodate survey data with sample weights according to Riillo (2017).

Constructing an appropriate matching model (i.e. a set of control variables that is used to match web and telephone samples) is an important requirement for the functionality of CEM (Iacus et al., 2012). To find the best matching model, we identify the most relevant variables with a combination of algorithm-based and theory-driven variable selections. First, we conduct an automatic variable selection via random forest in order to find strong predictors for the target variables and for the group assignment (Breiman, 2001). Subsequently, we check theory-based whether the automatically selected matching model includes all important variables and adjust our model accordingly. As result of this process, we use two different matching models for the CEM. For the target variable employment status we use age, sex, nationality, country of birth, ISCED, interview week, panel wave, questionnaire language, and collection year as matching variables. For the two variables wage adequacy and job satisfaction we apply the same matching model, consisting of age, sex, nationality, ISCED, income, NACE, ISCO, questionnaire language, and collection year.

**5. Results**

Figure 2 illustrates the differences for web and telephone samples before and after CEM for the target variable employment status. The figure displays proportions for the three employment status categories active, unemployed, and inactive. Weighted web and telephone samples before matching are represented by dark blue and dark green bars. Weighted web and telephone samples post matching are shown in light blue and light green.

Proportions of web and telephone samples before matching are very different within the categories. In comparison to the telephone sample, web participants have much more often an active employment status and are slightly more often unemployed. In contrast, web participants are less often inactive than participants of the telephone sample. After CEM, however, differences between web and telephone in the target variable employment status are not statistically significant, indicating that differences before matching are exclusively due to sample compositions of web and telephone data. Employment status is therefore not affected by systematic measurement errors due to mixed data collection modes.

**Figure *2*. Mixed mode effects on employment status**



Figure 3 visualizes differences before and after matching for the two target variables wage adequacy and job satisfaction. Both variables consist of the categories strongly agree, agree, disagree, and strongly disagree, with strongly agree for people with the highest wage adequacy or job satisfaction, respectively. As in Figure 2, web and telephone pre CEM are represented by the colors dark blue and dark green and web and telephone post CEM are represented by light blue and light green.

Before CEM, web participants perceive their salary less often as fair. Web participants select less often the categories strongly agree or agree and more often disagree or strongly disagree. Similar observations can be made for the variable job satisfaction. Unmatched web participants select the category strongly agree much less often and the categories agree, disagree, and strongly disagree more often than telephone participants. Notably, systematic differences between web and telephone samples remain after CEM. In the matched data, telephone participants still have a higher probability to select positive categories. The collection mode, hence, has a significant impact on the way how people respond to these two variables.

**Figure *3*. Mixed mode effects on wage adequacy & job satisfaction**



Interpreting the results, the question arises why wage adequacy and job satisfaction are affected by mode effects, but employment status is not. The main difference between those variables is the variable's objectiveness. Employment status is relatively stable and does not depend on personal opinions of the respondent – the variable is very objective. Wage adequacy and job satisfaction, in contrast, are very subjective. Objective variables such as employment status seem to be clearer in respondent's minds and are therefore less vulnerable to measurement bias. About subjective variables, respondents have to think more deeply, before they are able to reply. This thinking process might be disturbed by an interviewer on the telephone. Furthermore, we assume that social desirability effects have a larger impact on subjective variables. It might be unpleasant for respondents to reveal dissatisfaction with their employment situation to a real person and, hence, they might present their employment situation as too positive. The results therefore suggest that objective variables are less affected by mixed mode effects than subjective variables.

**6. Summary**

We contributed to mixed mode research by investigating on objective and subjective variables of the LFS. We harmonized differences in sample composition of web and telephone data via CEM, a methodological approach that – to our best knowledge – was not used for employment variables yet. With CEM, we approximated a random experiment and separated mode-specific measurement bias from mode-specific coverage and response rates. We applied CEM to the objective variable employment status and to the subjective variables wage adequacy and job satisfaction.

We found no significant mixed mode measurement effects for employment status, but substantial effects for the two subjective variables wage adequacy and job satisfaction. Results therefore suggest that mixed mode effects are less problematic for objective than for subjective variables. We assume that the presence of an interviewer affects participants' response behaviour, e.g., because of social desirability. Even though further research about mixed mode effects on subjective variables is advisable, our results overall support the view that mixed mode is an attractive way to collect survey data.

**References**

Bianchi, A., Biffignandi, S., and Lynn, P. (2017). Web-face-to-face mixed-mode design in a longitudinal survey: Effects on participation rates, sample composition, and costs. Journal of Official Statistics, 33(2):385-408.

Blanke, K. and Luiten, A. (2014). Query on data collection for social surveys. ESSnet Project "Data Collection for Social Surveys using Multiple Modes".

Breiman, L. (2001). Random forests. Machine Learning, 45(1):5-32.

de Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. Journal of Official Statistics, 21(2):233-255.

Dillman, D. A. (2017). The promise and challenge of pushing respondents to the web in mixed-mode surveys. Survey Methodology, 43(1):3-30.

Iacus, S. M., King, G., and Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. Political Analysis, 20(1):1-24.

Körner, T. and Liersch, A. (2014). Case study on mode effects in the German Labour Force Survey. Deliverable for work package III of the ESSnet on Data Collection for Social Surveys Using Multiple Modes.

Luiten, A. and Blanke, K. (2015). Conclusions of the ESSnet – DCSS on web and mixed mode data collection in official social surveys. Conference of New Techniques and Technologies for Statistics (NTTS).

Pohjanpää, K. (2014). The report of web pilot study of LFS (WPIII). Statistics Finland. The ESSnet project on Data Collection for Social Surveys using Multiple Modes.

Riillo, C. A. F. (2017). Beyond the question "Does it pay to be green?": How much green? and when? Journal of Cleaner Production, 141:626-640.

Sarracino, F., Riillo, C. A. F., and Mikucka, M. (2017). Comparability of web and telephone surveys for the measurement of subjective well-being. Survey Research Methods, 11(2):141-169.

Schouten, B. and van der Laan, J. (2014). ESSnet deliverable WPIII: Mode effect decompositions for the Dutch Labour Force Survey. Deliverable for work package III of the ESSnet on Data Collection for Social Surveys Using Multiple Modes.

van Buuren, S. (2012). Flexible Imputation of Missing Data. Taylor & Francis Ltd.