Survey Research Methods (2017)
Vol. 11, No. 2, pp. 141-169
doi:10.18148/srm/2017.v11i2.6740

ISSN 1864-3361
http://www.surveymethods.org

Comparability of web and telephone surveys for the measurement of subjective well-being

Francesco Sarracino
Institut national de la statistique et des études économiques du Grand-Duché du Luxembourg (STATEC) and Agence pour la normalisation et l'économie de la connaissance (ANEC)
Luxembourg City, Luxembourg
and LCSR National Research University Higher School of Economics
Moscow, Russian Federation

Cesare A.F. Riillo
Institut national de la statistique et des études économiques du Grand-Duché du Luxembourg (STATEC) and Agence pour la normalisation et l'économie de la connaissance (ANEC)
Luxembourg City, Luxembourg

Małgorzata Mikucka
Mannheimer Zentrum für Europäische Sozialforschung
Mannheim, Germany
and Université Catholique de Louvain
Lauvain-la-Neuve, Belgium

We compare subjective well-being measures collected with web and telephone surveys to test whether survey mode affects people's evaluations of their well-being. We use unique, nationally representative data from Luxembourg which contain five measures of subjective well-being collected through web and telephone surveys. Oaxaca decomposition and multinomial logit with Coarsened Exact Matching indicate that the survey mode affects reported well-being scores. Web respondents are more likely to report low well-being and less likely to report the neutral category. However, the consequences for statistical inference are negligible. Our results support the view that web and telephone surveys are comparable tools for collecting subjective data, such as people's well-being.

Keywords: web survey; telephone survey; survey mode; subjective well-being; Oaxaca decomposition; Coarsened Exact Matching

1 Introduction

Survey data are an important source of information in modern societies, and numerous indicators rely on the availability of timely and reliable information, such as those related to quality of life.[1] The most commonly used survey administration modes include: online surveys, comprising e-mail and web surveys; telephone interviews, typically conducted as a computer assisted telephone interview (CATI), but sometimes with the use of automatized interactive voice recognition (IVR) technology; mail surveys; and face-to-face interviews, usually conducted as computer assisted personal interviews (CAPI) or as pen-and-paper interviews (PAPI). For a long time, telephone and face-to-face interviews have been the main modes of collecting micro data, yet their increasing costs and quality concerns (low response rates, missing data, presence of duplicates, etc.) pose a challenge to the future of survey research.

Web surveys, i.e. surveys administered via the Internet, have gained popularity in recent years as they have a number of benefits over conventional paper or face-to-face methods (Wyatt, 2000). In particular, web surveys potentially allow to reach a global audience; once set up they are cheap to carry out, making it easy to recruit large numbers of participants or to collect data repeatedly; the data collected are directly stored in electronic format, reducing the probability of encoding errors, thus speeding up data preparation and analysis. Web surveys also allow the use of multimedia and other software to administer experiments and to enhance the quality of the questions; participants may choose to respond

Contact information: Francesco Sarracino, Institut national de la statistique et des études économiques du Grand-Duché du Luxembourg (STATEC), 13, rue Erasme, L-2013 Luxembourg (e-mail: Francesco.Sarracino@statec.etat.lu)

[1] See, for instance, the Better Life Index of the OECD, the Eurostat's project Quality of life in Europe, the Italian Equitable and Sustainable Well-being, the British Measuring National Well-being, or the Luxembourgish PIBien-être.

at their convenience, with possible benefits for the quality of the answers. Moreover, Internet connections are widespread, with a penetration of more than 80% in Western Europe and North America, thus offering an ideal tool for administering surveys (We Are Social, 2016). For these reasons web surveys are an appealing tool for survey practitioners (Couper, 2013).

Previous research investigated comparability of well-being scores between face to face and telephone survey modes. Dolan and Kavetsos, 2012 showed that respondents of a telephone survey reported higher well-being than respondents of a face to face survey. The survey mode affected also the correlates of subjective well-being: life circumstances correlated with subjective well-being stronger in the face to face sample than in the telephone sample. Our work adds to this evidence by studying the comparability of subjective measures collected on nationally representative samples through telephone and web survey modes.

This issue is relevant for at least two reasons. The first one is the growing popularity of mixed data collection mode (Dillman et al., 2009; Jäckle, Lynn, & Burton, 2015; Nigg et al., 2009; Roster, Rogers, Albaum, & Klein, 2004; Schonlau, Asch, & Du, 2003). The second reason is the need to compare data collected with various modes across countries or time (as, e.g. in case of the World Values Survey and the European Values Study data). In such cases, the comparability of answers collected with various modes is a precondition to obtain reliable conclusions.

We contribute to this literature by testing the comparability of five measures of subjective well-being collected using telephone and web survey modes from a nationally representative sample of residents of Luxembourg (aged 18-64). Both for historical reasons (telephone surveys were introduced before web surveys) and for economic reasons (the former are more expensive than the latter) we consider telephone surveys as the conventional mode and we test whether web survey mode affects people's evaluations of their well-being compared to telephone surveys.

We focus on subjective well-being because in recent years it received considerable attention in the public, political, and academic debate. The reason is probably that subjective well-being allows a direct and reliable monitoring of how well people fare.[2] Nearly 40 years of psychological, economic and sociological research on what matters for people's well-being and what are its consequences for economic, social and political outcomes have confirmed that subjective well-being deserves close monitoring. Many surveys, covering virtually every country in the world, ask respondents to evaluate, among other things, their subjective well-being.[3]

We use Global Entrepreneurship Monitor data for Luxembourg, a country where Internet penetration is 96% (GEM, 2016). These data are unique because every year, from 2013 to 2015, nearly half of the sample has been interviewed via web survey and the other half via telephone survey. Additionally, respondents evaluated their subjective well-being using five different questions. We use the Blinder-Oaxaca decomposition to test whether the differences in well-being scores between the two samples may be attributed to different observable characteristics. Subsequently, we use multinomial logit with Coarsened Exact Matching to investigate the effect of survey mode on the differences in distributions of answers, after matching respondents on observable characteristics. Finally, we test how survey mode affects statistical inference.

Our results show that the use of web surveys slightly alters people's evaluations of their well-being. In three out of five measures web survey generates lower well-being scores than the telephone surveys. Moreover, for all five variables, respondents of web surveys are less likely to choose the neutral category ("neither agree nor disagree") than respondents of telephone survey. However, the effect of survey mode for statistical inference is negligible. These results support the view that, compared to telephone surveys, web surveys are convenient and reliable tools to collect data on subjective well-being.

The paper is organized as follows. In section 2 we present the main concepts and review previous literature on the differences between telephone and web survey modes. Section 3 and 4 describe, respectively, the data and the methods. Section 5 reports our results. We begin with descriptive statistics, then we present results of the Blinder-Oaxaca decomposition, the results from the Coarsened Exact Matching, and, finally, the results for statistical inference. The last section summarizes our findings and provides concluding remarks.

[2]The reliability of SWB measures has been corroborated by experimental evidence from several disciplines. For example, SWB correlates with objective measures of well-being such as the heart rate, blood pressure, duration of Duchenne smiles and neurological tests of brain activity (Blanchflower & Oswald, 2008; van Reekum et al., 2007). Moreover, SWB measures are correlated with other proxies of SWB (Schimmack, Krause, Wagner, & Schupp, 2010; Schwarz & Strack, 1999; Wanous & Hudy, 2001) and – more interestingly – they mirror the judgments about the respondent's happiness provided by acquaintances or clinical experts (Kahneman & Krueger, 2006; Layard, 2005; Schneider & Schimmack, 2009).

[3]Examples of cross-country surveys collecting information on people's well-being include, and are not limited to, the World Values Survey, the European Value Study, the European Social Survey, the General Social Survey, the International Social Survey Programme, the European Quality of Life Survey, besides a number of panel studies – such as the German Socio-Economic Panel, the British Household Panel Study, the Household, Income and Labour Dynamics in Australia, or the European Survey on Income and Living Conditions – and other more regional or sub-regional studies.

2 Comparability of data collected through telephone and web surveys

In general, comparability of surveys collected with various modes may be limited by two types of effects: sample and response effects. Sample effects (sometimes referred to as 'coverage errors') are differences in answers resulting from over- or under-representation in the sample of specific categories of respondents (Roster et al., 2004). In other words, sample effects result from the mismatch between the target population of a study (e.g., the adult population of a country) and the sampling frame (i.e. the population from which the study participants have been selected) (Couper, 2000).

Response effects (also called 'mode effects') are the response differences which result from the way respondents interact with the survey mode. Response effects include, among others, social desirability, satisficing (i.e. the tendency to provide similar answers to a battery of questions), acquiescence (i.e. the tendency to agree rather than to disagree), and order effects (in particular, recency and primacy effects). Response effects are typically larger for topics which are less crystallized in the minds of respondents, because the uncertainty of how to answer may push respondents to search for clues present in the situation (e.g. answering scales or other questions which has been asked before) to formulate their answers (Dillman & Christian, 2005). Responses to factual questions or to questions concerning issues on which respondents have crystallized opinions tend to be less affected by the survey mode. Both sample and response effects in the context of web and telephone surveys are discussed in detail below.

2.1 Sample effects

Telephone users typically differ from Internet users, and both groups differ from the general population, which may lower the representativeness of web and telephone samples (Yeager et al., 2011). Some people are systematically more likely to answer the phone, and others are more likely to fill in a web-based survey (Bethlehem, 2010; Dever, Rafferty, & Valliant, 2008; Dillman et al., 2009; Jäckle et al., 2015). Research showed that telephone samples over-represented the older respondents (Roster et al., 2004), whereas web surveys under-represented older respondents (Bech & Kristensen, 2009), over-represented younger people (Flemming & Sonner, 1999; Roster et al., 2004), those with higher incomes (Couper, 2000), more geographically mobile (Roster et al., 2004), and men (Flemming & Sonner, 1999). Results for education were mixed. Some studies showed that telephone surveys over-represented adults with a college degree (Berrens, Bohara, Jenkins-Smith, Silva, & Weimer, 2003; Roster et al., 2004), whereas other connected higher education with increased chances of being active online (Couper, 2000). In American studies, both telephone and web surveys seemed to under-represent ethnic minorities; however, the degree of under-representation seems more substantial for telephone than for web surveys (Berrens et al., 2003).

Socio-demographic differences between the samples may be problematic, but they can be partly corrected by weights, so that sample differences are taken into account in statistical analysis (Berrens et al., 2003; Dever et al., 2008). Differences on unobservables are more difficult to account for. Note that samples identical in terms of socioeconomic characteristics may lead to different conclusions about relationships among variables if they differ in terms of unobservable characteristics (Berrens et al., 2003; Couper, 2000).

Representativeness of samples strictly relates to the method of sample selection. Non-probability samples (typical for web surveys whose respondents are recruited through banners on websites or e-mail invitations sent to a large group) tend to be more variable in their accuracy than probability samples (Chang & Krosnick, 2009; Yeager et al., 2011). In some cases, self-selection (Bethlehem, 2010) may substantially affect data quality (Bethlehem, 2015; Hudson, Seah, Hite, & Haab, 2004). Research showed systematic differences of opinions and attitudes between respondents of telephone and web surveys, in particular those following the non-probability sampling (Flemming & Sonner, 1999). For example, non-probability web sample may be biased toward people engaged and knowledgeable about the survey's topic (Chang & Krosnick, 2009), or towards people using e-mail and Internet more often than others (Yun & Trumbo, 2000).

2.2 Response effects

Measurement error. Measurement error refers to the deviation of the answers from their true values (Couper, 2000). In self-administered surveys, such as web-based ones, it may result from low motivation of the respondents, problems with understanding a question, or problems related to the questionnaire itself. In telephone surveys an interviewer may motivate respondents, clarify questions, or probe incomplete responses. All these factors may increase the quality of collected data.

However, web surveys give respondents a possibility to check relevant information before answering a question (Braunsberger, Wybenga, & Gates, 2007), and they create less pressure than telephone surveys to provide an answer quickly. Moreover, respondents participating in web surveys have more freedom to choose when to answer the survey questions, may be therefore less distracted and more focused on the survey (Kiesler & Sproull, 1986; Yun & Trumbo, 2000). This explains why respondents participating to web surveys provided more accurate answers to questions testing their knowledge (Fricker, Galesic, Tourangeau, & Yan, 2005), and their answers were less affected by random measurement errors (Chang & Krosnick, 2009) than those of telephone survey respondents.

Satisficing. Satisficing occurs when respondents provide similar answers to a battery of questions because they do not devote effort to the responding process. Satisficing may occur especially in web surveys which use grids (so called 'straightlining') (Fricker et al., 2005; Tourangeau, Couper, & Conrad, 2004; Zhang & Conrad, 2014), but the evidence on this issue is mixed (Chang & Krosnick, 2009). Empirical studies show that satisficing is more frequent among lower educated respondents (Malhotra, 2008; Zhang & Conrad, 2014), and among respondents who prefer a different survey mode than the mode actually used (Smyth, Olson, & Kasabian, 2014).

Social desirability. The great advantage of telephone surveys, i.e. the use of interviewers who can support and guide respondents (Braunsberger et al., 2007), may also contribute to lowering the quality of obtained data by introducing social desirability bias (Braunsberger et al., 2007; Johnson, Fendrich, Shaligram, Garcy, & Gillespie, 2000). Self-administered web surveys provide privacy to the respondent and are more likely to report some socially undesirable opinions or behaviors. This may limit the comparability of interviewer- and self-administered modes (Klausch, Hox, & Schouten, 2013). Studies confirmed that web surveys typically yield less socially desirable answers than telephone surveys (Bason, 2000; Chang & Krosnick, 2009; Christian, Dillman, & Smyth, 2008; Kreuter, Presser, & Tourangeau, 2008; Nigg et al., 2009). Moreover, respondents participating in a telephone interview may perceive the same questions as more sensitive than respondents participating in a web survey (Kreuter et al., 2008). However, previous literature found only marginal differences in answering sensitive questions between telephone and web surveys (Bason, 2000; Parks, Pardi, & Bradizza, 2006).

Acquiescence: Positive and negative answers. Acquiescence, i.e. the tendency to give positive rather than negative answers is typically stronger in telephone than in web surveys. The evidence, however, is mixed: some studies found that web survey respondents tend to give more neutral and more negative answers, especially to questions referring to their attitudes or satisfaction, than in telephone surveys (Braunsberger et al., 2007; Christian et al., 2008; Dillman et al., 2009; Roster et al., 2004), whereas other studies did not find any significant difference (Fricker et al., 2005).

Primacy and recency. The recency effect is the tendency of respondents (primarily those participating in telephone interviews) to choose more often the last-offered answer. The primacy effect, on the contrary, refers to the tendency of respondents in self-administered surveys (such as mail or web surveys) to choose systematically more often the first-offered answer choices (Christian et al., 2008; Dillman & Christian, 2005).

2.3 Assessing comparability using regression analysis

Some studies have used regression analysis to assess the comparability of data collected with various survey modes (Chang & Krosnick, 2009; Dolan & Kavetsos, 2012; Roster et al., 2004; Yun & Trumbo, 2000). Results are mixed. For example, Yun and Trumbo, 2000 observed nearly identical R-squares in regression models estimated on data collected by traditional mail vs. electronic channel (web and e-mail surveys combined); they also noted few statistical differences between beta coefficients. On the contrary, Dolan and Kavetsos (2012) found differences in the size and significance of coefficients of the correlates of well-being collected with face to face and telephone surveys.

3 Data

To evaluate how the web survey mode affects people's reports about their well-being, we use the pooled Adult Population Survey of the Global Entrepreneurship Monitor (GEM, 2016), a survey of the Statistical Office of Luxembourg (STATEC) conducted in 2013, 2014, and 2015 on a nationally representative sample of residents of Luxembourg aged between 18 and 64 years. GEM is designed as a source of harmonized, individual-level information about people's motives and aspirations for entrepreneurship. The survey is administered in more than 100 countries worldwide, covering over 75% of the world's population. Among the participating countries, Luxembourg is the only country that consistently monitored respondents' subjective well-being using a battery of five questions, including an item about life satisfaction. In addition, nearly 50% of the 6000 sampled respondents has been interviewed using computer assisted telephone interviews (CATI), while the other 50% has responded to a web survey. These proportions have changed in 2015 when 40% of the sample was interviewed by phone and the remaining 60% by web survey.

Telephone survey respondents have been randomly drawn from the household, fixed line phone registry, while participants of the web survey have been randomly selected from a database containing over 14,000 e-mail addresses. The database contains the e-mail addresses of people who have been previously contacted in various ways – either in previous interviews, or through advertisement and local newspapers campaigns – and who agreed to participate to web surveys in the future.

In the analysis, we limited the sample to respondents who potentially could have been selected for both the telephone and the web survey: we excluded people who did not have either the telephone line or the internet connection (i.e. 451 out of the 6095 observations).

Respondents were asked to evaluate their life satisfaction, by stating how strongly they agreed with the statement *"I am satisfied of my life."* The possible answers ranged from

1 – *strongly disagree* to 5 – *strongly agree*. The survey included also the following four alternative measures of subjective well-being: *"So far I have obtained the important things I want in life"*, *"If I could live my life again, I would not change anything"*, *"The conditions of my life are excellent"*, and *"In most ways my life is close to my ideal"*. Answers are coded on a scale from 1 to 5, consistently with the item about life satisfaction.

Table 1 shows the bivariate correlations among the measures of subjective well-being. The correlations between life satisfaction and having obtained the important things in life, having excellent living conditions, and having a life close to an ideal exceed 50%, whereas the correlation of life satisfaction with the unwillingness to change anything in one's life is 36%. The correlations between the four alternative measures of subjective well-being range between 36 and 56%.

The survey mode used to collect the data is encoded with a dummy variable set to zero if the respondent answered a telephone survey, and 1 if the respondent answered a web survey.

To account for individual socio-economic differences between the two groups of respondents, we include the following controls: age and age squared (continuous variables), gender (0 – men, 1 – women), education (dummies for: [a] elementary education; [b] secondary education; [c] master craftsman; [d] bachelor; and [e] master), employment status (dummies for: [a] employed full-time; [b] employed part-time; [c] self-employed; [d] retired or disabled; [e] home-maker; [f] student; and [g] other not working), and income of the respondents. The total annual income of all the household's members is observed through respondent's self-declaration of belonging to one of the following classes : [a] 0-20,000; [b] 20,001-40,000; [c] 40,001-60,000; [d] 60,001-80,000; [e] 80,001-100,000; [f] more than 100,000 euro. We assigned to the respondents the middle value in their income bracket, after limiting the upper category to 120,000 euro. Subsequently, we deflated the incomes to the real euro of 2005 using national consumer price index, and converted the income into logarithm. Finally, to account for time effect, we included year fixed effects. Descriptive statistics are presented in Table 2, whereas Figures A1 - A1 in Appendix A show the distribution of gender, education, age, income and occupation in the telephone and web samples. Figures indicate that the sample of people who participated in the web survey included a higher share of men, people with a bachelor or master degree, young, and full-time employed.

4 Methods

To test whether survey mode affects the reported well-being we use two complementary techniques: the Blinder-Oaxaca decomposition and the multinomial logit with Coarsened Exact Matching (CEM).

4.1 Blinder-Oaxaca decomposition

The Blinder-Oaxaca decomposition has been developed in the early '70s by Blinder (1973) and R. Oaxaca (1973) to study wage discrimination of women in the labor market. Recently, it has been applied also in other fields, including the literature on subjective well-being (e.g., Bartolini & Sarracino, 2015; Helliwell & Barrington-Leigh, 2010; Sarracino, 2013).

The Blinder-Oaxaca decomposition allows us to decompose the difference (gap) between subjective well-being declared in telephone and web surveys into two components. The first component is the explained part of the gap, which tells which part of the gap stems from the different observable characteristics of the samples, such as income, education, etc. The second component, known as the unexplained part of the gap, tells which part of the gap is due to differences in the estimated coefficients, i.e. how people interact with the specific survey mode. If the two groups of respondents are randomly chosen from the overall population, we expect the socio-demographic and economic variables to play a negligible role, i.e. the explained part of the gap should be non significant, and unexplained part, i.e. the survey mode, should be significant. Thus, the Oaxaca-Blinder decomposition, informs whether the two samples are comparable and whether the survey mode plays a significant role in generating the subjective well-being gap.

Formally, the Blinder-Oaxaca decomposition can be written as:

$$\Delta \text{SWB} = \underbrace{(\overline{\mathbf{X}}_A - \overline{\mathbf{X}}_B) \cdot \beta^*}_{\text{explained}} + \underbrace{\overline{\mathbf{X}}_A \cdot (\beta_A - \beta^*) + \overline{\mathbf{X}}_B \cdot (\beta^* - \beta_B)}_{\text{unexplained}} \tag{1}$$

where ΔSWB is the subjective well-being gap; A and B are the two groups (g); $\overline{\mathbf{X}}_g$ is a vector of group averages of explanatory variables; β_A and β_B are the coefficients estimated for group A and B, respectively; and β^* is a vector of *non-discriminatory* coefficients to assess how much different characteristics of individuals belonging to the two groups explain the overall difference. The coefficient β^* is estimated based on Ω, which is defined as follows:

$$\beta^* = \Omega \cdot \beta_A + (I - \Omega) \cdot \beta_B \tag{2}$$

where Ω is estimated with a pooled model to derive β^* (Neumark, 1988; R. L. Oaxaca & Ransom, 1994). Computations are based on the package *nldecompose* which allows to decompose the gap for a categorical variable using Stata (Sinning, Hahn, & Bauer, 2008).

In the vector of explanatory variables (the **X** in Equation 1), we include the following control variables: gender, age and age squared, employment status, education, income, and dummy for each year of observation.

Table 1
Correlation matrix of the measures of subjective well-being.

	life satisfaction	important things in life	not change anything	excellent life conditions	life close to ideal
life satisfaction	1				
important things in life	0.60***	1			
not change anything	0.36***	0.39***	1		
excellent life conditions	0.56***	0.51***	0.36***	1	
life close to ideal	0.60***	0.50***	0.37***	0.56 ***	1

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015 (N = 5644)
*** $p < 0.01$

Table 2
Descriptive statistics of variables used in analysis

variable	mean	sd	min	max	obs
Women	0.52	-	0	1	5,644
Age (in years)	42.69	12.66	18	64	5,644
Age squared / 100	19.82	10.58	3.24	40.96	5,644
Elementary education	0.04	-	0	1	5,502
Secondary education	0.48	-	0	1	5,502
Master craftsman	0.06	-	0	1	5,502
Bachelor	0.25	-	0	1	5,502
Master	0.15	-	0	1	5,502
Full-time employed	0.55	-	0	1	5,545
Part-time employed	0.13	-	0	1	5,545
Self-employed	0.05	-	0	1	5,545
Retired or disabled	0.11	-	0	1	5,545
Homemaker	0.05	-	0	1	5,545
Student	0.08	-	0	1	5,545
Other, not working	0.03	-	0	1	5,545
Income (log)	10.82	0.55	9.01	11.42	4,363
Year	-	-	2013	2015	5,644

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Standard deviations of dichotomous variables have been omitted. Sample limited to respondents with Internet access and fixed phone line at home.

4.2 Multinomial logit with Coarsened Exact Matching (CEM)

The Oaxaca-Blinder decomposition informs about how the survey mode affects the average well-being, but it does not tell how it affects people's choice of specific categories of well-being. To address this point we use a multinomial logit after implementing a Coarsened Exact Matching (CEM) procedure. CEM is a matching method recently introduced in the literature (Iacus, King, & Porro, 2011, 2012) that approximates randomized experiments by reducing dissimilarities between two groups of respondents. CEM ensures that the "treated" group (the web survey respondents) and the "control" group (the telephone survey respondents) are balanced, i.e. the treated and control groups are similar on the distribution of the covariates. CEM is a monotonic imbalance reducing matching method, i.e. a technique that allows the user to choose ex-ante the balance between the treated and control groups with the aim of controlling for some or all of the potentially confounding influence of pre-treatment control variables (Iacus et al., 2012). CEM creates strata using temporally coarsened (i.e. with values grouped into substantially meaningful categories) variables and retains observations so that strata have at least one treatment and one control unit. CEM may remove both treated and control units, and the identified sub-set has similar characteristics not only on average, but on the whole distribution of observable variables. Subsequent estimations are performed only on the matched data using the original uncoarsened variables (Blackwell, 2009). This allows to compare the outcomes of matched treated and control groups approximating an experimental setting. CEM-based estimates have powerful statistical properties and can outperform others matching procedures (Iacus et al., 2011). For instance, compared to Mahalanobis and propensity score matching methods, CEM generally provides the best trade-off between covariates' imbalance and sample size of matched observations (Iacus et al., 2012). CEM is successfully implement in innovation (Aggarwal & Hsu, 2014), well-being (Sidney, Coberley, Pope, & Wells, 2015) and environmental researches (Riillo, 2017).

In present study, individuals are exactly matched on gender, employment status, education level, and year of the survey, while they are coarsely matched on continuous variables, i.e. age and income. The CEM procedure is performed only for individuals with Internet access and fixed phone line at home, and CEM weights are adjusted to accommodate survey data with sample weights according to Riillo (2017). The CEM matched observations are subsequently analyzed using multinomial logit, which is suited to deal with categorical variables, such as our measures of well-being.[4]

Multinomial logit model estimates simultaneously the association between survey mode and all possible outcomes of subjective well-being, accounting for the set of control variables. Implementing the multinomial logit model on CEM retained data has the advantage to estimate the survey mode effect at different levels of subjective well-being, on a sample of observations that resemble experimental data.

5 Results

Figure 1 shows the distribution of life satisfaction by survey mode. Both distributions are right skewed, as it is often the case with subjective well-being data. Yet, while the share of those who are strongly satisfied with their lives is identical between the two groups, respondents of the web survey indicate more often the lower categories, and on average they report lower life satisfaction than respondents of the telephone survey. Also missing data are more frequent among respondents of the web survey. To test whether the two samples follow different distributions, we run a Kolmogorov-Smirnov test. This is a non parametric technique to compute the maximum absolute value of the differences between two cumulative distributions functions (CDFs). The null hypothesis is that two samples have identical distributions. The result of the test indicates that the CDFs of the two samples are significantly different from each other, i.e. that life satisfaction does not have the same distribution across survey modes: the maximum absolute difference between the two CDFs is 0.092 which is significant at 99%.

On average, respondents of the telephone survey report a life satisfaction score of 4.15, whereas the average in the web survey is 4. The difference is small, yet statistically significant ($p < 0.01$). While the share of people strongly satisfied with their lives is similar across survey modes, differences become more evident at the lower end of the scale: in the web survey 4.04% of respondents *"strongly disagreed"*, and 9.28% *"disagreed"* that they are satisfied with their lives; in the telephone survey the respective shares are 1.48% and 3.26%. The probabilities to choose categories *"neither agree nor disagree"* and *"agree"* are higher among respondents using the telephone survey (12.10% and 44.67%) than the web survey (7.61% and 39.80%).

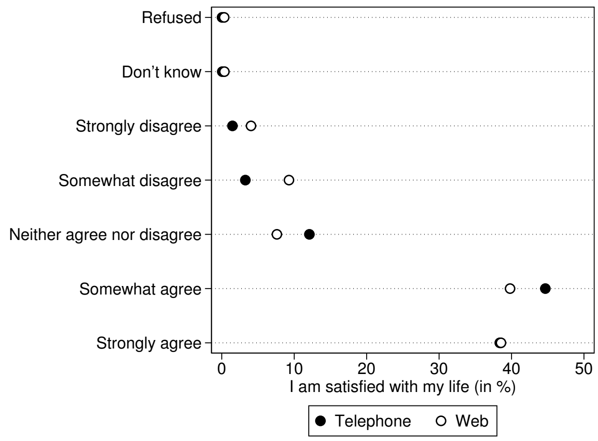

Figure 1. Distribution of life satisfaction by survey mode.

This suggests that the survey mode affects people's responses to the life satisfaction question: although small, the difference in the average life satisfaction between web and telephone respondents is significant. Additionally, respondents of the web survey have a higher probability to choose low categories compared to those who used the telephone survey.

The possible explanation of these differences is that the two groups of respondents differ in socio-demographic characteristics which also predict well-being. This is plausible because, as discussed in section 3 and reported in figures A1 - A1 in Appendix A, respondents of the web survey are on average younger, more educated, and a greater share of them are employed. Table 3 shows the marginal effects of the socio-demographic controls on the probability of participating in the web rather than in the telephone survey. Variables are incrementally included in the model to test their relative importance.

The probability of participating in the web survey is lower for women, retired people, homemakers, students, and people with lower incomes; the relationship with age is U-shaped: the probability of participating in a web survey de-

[4]An alternative method used with ordered variables, i.e. the ordered logit model, imposes the same coefficient of the variable of interest for all outcome categories. This assumption, known in the literature as parallel regression assumption (Long & Freese, 2006), is violated in our data.

Table 3
Predictors of participation in the web rather than in the telephone survey. Marginal effects after ordered probit regression.

	Model 1		Model 2		Model 3	
	dy/dx	t	dy/dx	t	dy/dx	t
Women	-0.08^{***}	−5.12	-0.07^{***}	−4.00	-0.06^{***}	−3.76
Age (in years)	-0.01^{**}	−2.79	-0.03^{***}	−5.48	-0.03^{***}	−5.52
Age squared / 100	0.01	1.79	0.03^{***}	4.62	0.03^{***}	4.60
Education (ref: Elementary)						
Secondary	-	-	0.03	0.84	0.02	0.78
Master craftsman	-	-	0.06	1.32	0.05	1.29
Bachelor	-	-	0.06^{*}	1.98	0.05	1.63
Master	-	-	0.03	1.00	0.02	0.54
Employment status (ref: Full-time)						
Part-time employed	-	-	−0.02	−0.65	−0.01	−0.51
Self-employed	-	-	0.03	0.93	0.04	1.08
Retired or disabled	-	-	-0.09^{**}	−3.06	-0.08^{**}	−2.66
Homemaker	-	-	-0.10^{*}	−2.40	-0.09^{*}	−2.23
Student	-	-	-0.22^{***}	−5.01	-0.21^{***}	−4.84
Other, not working	-	-	−0.07	−1.52	−0.05	−1.06
Household income (log)	-	-	-	-	0.04^{**}	2.74
Year (ref: 2013)						
year 2014	−0.01	−0.39	−0.00	−0.26	−0.01	−0.29
year 2015	0.07^{***}	3.51	0.07^{***}	3.60	0.07^{***}	3.59
Observations	4,238		4,238		4,238	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home. dy/dx: marginal effect; t: t-statistics. * $p < .10$ ** $p < .05$ *** $p < .01$

clines with age, but it reverts after a given threshold. Finally, the coefficient for year 2015 reflects the increased size of the web sample in that year. As some of these variables are also standard predictors of life satisfaction, it is plausible that the observed differences in life satisfaction are due to the different socio-demographic composition of the two groups. To address this issue we resort to the Blinder-Oaxaca decomposition.

5.1 Blinder-Oaxaca decomposition

Table 4 shows the decomposition results. The difference ('gap') in average life satisfaction between the two groups of respondents is 0.15 ($p < .01$). A positive gap means that on average telephone survey respondents report higher well-being than web survey respondents. The explained part of the gap is not significantly different from zero which means that the different socio-demographic characteristics of the two groups do not predict any significant well-being gap. The decomposition shows that all the difference is unexplained, which suggests that the gap is the result of the way people interact with survey modes. Specifically, the use of telephone survey reduces the gap, i.e. if respondents of the web survey had participated in the telephone survey, then the well-being gap between these two groups would be -0.246 points smaller, on a scale from 1 to 5. In other words, if the survey had been administered only via telephone, then the overall well-being would have been higher. On the contrary, the coefficients of web survey respondents predict an increase of the gap by 0.399 on a scale from 1 to 5. This confirms that the use of the web mode affects people's responses independently from their socio-economic characteristics.

In sum, the main result from the decomposition is that the gap in life satisfaction between web and telephone respondents is not due to the different socio-demographic characteristics of the two groups.

5.2 Multinomial logit with CEM

As explained in the methodological section, coarsened exact matching (CEM) aims to closely align the sample of respondents of web (treated) and telephone (controls). Table 5 shows the distribution of socio-demographic characteristics before and after CEM procedure. Second and fourth columns report the averages of variables in the web sample, while the third and fifth columns report the averages of variables in

Table 4
Blinder-Oaxaca decomposition of the life satisfaction gap by survey mode.

	coefficient	z-stat	p-values
Difference:	0.15	3.36	0.00
Decomposition:			
Explained	0.00	−0.61	0.54
Telephone	−0.25	−5.80	0.00
Web-survey	0.40	19.42	0.00
Observations	4,230	-	-

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home.

the telephone sample. The stars indicate that the differences between the averages of the web and telephone samples are statistically significant. Table 5 suggests that before CEM, the two samples of respondents have different characteristics, consistently with the results from Table 3. After CEM, with the exception of age, the differences between the web and telephone samples are not statistically significant. Iacus et al. (2012) introduced a comprehensive measure of global imbalance *L1* which helps evaluate how successfully CEM balances treated and control samples (*L1* ranges from 0 to 1 with 0 indicating perfect balance and 1 maximum imbalance). In our case, *L1* distance is .74 before applying CEM and .28 after CEM, suggesting that CEM procedure considerably reduces unbalance. The residual imbalance is statistically modelled using multinomial logit regression on the post-CEM sample.

Coefficients in Table 6 report the effect of web survey on the probability to choose each of the five responses to the life satisfaction question. Web survey respondents have a lower probability to choose the middle category (*"neither agree nor disagree"*) and a higher probability to choose the negative (*"disagree"* and *"strongly disagree"*) categories. The probability of choosing the upper two categories is not affected. The result is the slight downward shift of the distribution of the answers noted above, and a consequent lower average subjective well-being for people who used a web survey.

5.3 Alternative measures of subjective well-being

GEM data provide four other measures of subjective well-being. The distribution of the answers provided in web and telephone surveys, and the t-test of the difference of averages are reported in Appendix B. The telephone survey respondents declared, on average, significantly more positive scores of well-being for the statements *"The conditions of my life are excellent"* (gap of 0.149, $p < 0.01$), and *"In most ways my life is close to my ideal"* (gap of 0.062, $p < 0.05$).

Table 7
Blinder-Oaxaca decomposition of subjective well-being gap by survey mode. Alternative measures of well-being.

	coef.	z-stat	p-values
The conditions of my life are excellent			
Difference	0.15	4.70	0.00
Decomposition:			
Explained	−0.01	−1.36	0.18
Telephone	−0.33	−3.84	0.00
Web-survey	0.49	5.31	0.00
Observations	4,229		
In most ways my life is close to my ideal			
Difference	0.06	2.37	0.02
Decomposition:			
Explained	−0.00	−1.20	0.23
Telephone	−0.48	−10.96	0.00
Web-survey	0.55	13.80	0.00
Observations	4,220		

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home.

We run the Blinder-Oaxaca decomposition for these two variables (see Table 7). The results indicate that in both cases the gap is entirely unexplained, i.e. it is associated with the way respondents interact with the survey mode, and not with their socio-demographic characteristics. In both cases, the use of the web survey has a positive sign, i.e. it increases the well-being gap between the two groups.

The counterfactual analysis using multinomial logit and CEM (Table 8) confirms that also for the alternative measures web respondents have a lower probability to choose the middle category *"neither agree nor disagree"* than telephone respondents, and a higher probability to choose a negative answer (*"strongly disagree"* or *"disagree"*). This is consistent with the results obtained for life satisfaction.

However, for three alternative measures of well-being, web respondents have also a higher probability to choose a positive answer category (*"strongly agree"* or *"agree"*). This suggests that the tendency to report lower well-being is only part of the story. Web respondents tend to choose more often either positive or negative categories instead of the neutral category *"neither agree nor disagree"*.

Summing up, in three out of five measures of subjective well-being our results show a systematic downward bias associated with participation in web rather than telephone survey. This effect is small but statistically significant, and is not explained by socio-demographic characteristics of respon-

Table 5
Socio-demographic characteristics before and after CEM procedure.

	Pre-CEM		Post-CEM	
Variable	Web	Telephone	Web	Telephone
Woman	0.43	0.50***	0.38	0.40
	(40.09)	(44.38)	(25.61)	(21.59)
Age	40.98	43.52***	42.01	43.23**
	(153.73)	(163.93)	(121.89)	(107.34)
Education				
Primary	0.03	0.04*	0.01	0.01
	(8.89)	(9.85)	(3.85)	(3.94)
Secondary	0.45	0.48**	0.54	0.55
	(41.10)	(42.74)	(34.91)	(28.22)
Master	0.06	0.06	0.03	0.033
	(11.64)	(11.09)	(5.96)	(5.26)
Bachelor	0.27	0.24**	0.25	0.25
	(27.97)	(24.98)	(18.45)	(14.77)
Master	0.16	0.15	0.16	0.15
	(20.01)	(18.67)	(13.67)	(9.97)
Employment status				
Full-time	0.64	0.58***	0.77	0.76
	(62.23)	(52.80)	(61.06)	(50.21)
Part-time	0.11	0.12	0.09	0.09
	(16.63)	(17.40)	(10.32)	(9.03)
Retired	0.08	0.12***	0.07	0.08
	(14.42)	(17.44)	(9.31)	(9.27)
Homemaker	0.03	0.05***	0.01	0.02
	(7.98)	(10.66)	(4.32)	(3.96)
Student	0.05	0.06	0.05	0.04
	(10.12)	(10.19)	(6.83)	(5.89)
Other, not working	0.03	0.03	0.00	0.00
	(7.95)	(7.95)	(1.99)	(2.14)
Self-employed	0.06	0.05	0.00	0.00
	(11.22)	(9.78)	(2.23)	(2.36)
Income	10.85	10.79***	10.92	10.92
	(940.78)	(850.12)	(792.58)	(657.21)
Year				
year 2014	0.33	0.37**	0.33	0.32
	(32.22)	(33.91)	(22.17)	(16.84)
year 2015	0.36	0.30***	0.35	0.36
	(34.88)	(29.67)	(23.83)	(19.18)
Observations	2,181	2,016	1,070	1,064

Weighed means (or proportions) are reported. Robust standard errors in parentheses. Stars indicate that the difference between telephone and web sample is significant.
* $p < .10$ ** $p < .05$ *** $p < .01$

Table 6
Web survey mode as a predictor of life satisfaction scores. Marginal effects from multinomial logit with CEM.

I am satisfied with my life:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Web survey mode	0.02**	0.05***	–0.04*	–0.06*	0.03
	(0.007)	(0.012)	(0.015)	(0.024)	(0.023)
Observations	2,134	2,134	2,134	2,134	2,134

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home. Coefficients for control variables are omitted for brevity. The complete list of results is available in Table **??** in Appendix D. CEM for survey data computed according to Riillo (2017). Robust standard errors in parentheses.
* $p < .10$ ** $p < .05$ *** $p < .01$

dents. Moreover, in all five measures of well-being, the web survey respondents have a lower probability of choosing the middle (neutral) category, higher probability of choosing the negative answers, and – in three out of five measures – higher probability of choosing positive answers.

5.4 Implications for statistical inference

A small effect on the average and on the distribution can, nonetheless, have a significant impact on statistical inference. To evaluate whether a survey mode affects the conclusions about the correlates of subjective well-being, we run a pooled subjective well-being regression on the survey mode dummy, socio-economic controls and the interaction of the survey mode with all socio-economic controls. The Wald test of joint significance of all interactions confirms that the coefficients from the telephone and the web survey are jointly statistically different from each other. For life satisfaction the Chi^2 of the Wald test is 41.46 with 21 degrees of freedom and a p-value of 0.0049.

Table 9 allows to identify the drivers of such difference. The table reports the marginal effects for the fifth category of life satisfaction (*"strongly agree"*) after a multinomial logit regression in which each independent variable is interacted with the survey mode. Columns two and three report the marginal effects for the respondents of the telephone and of the web survey respectively. Column 4 reports the difference between the previous two and informs whether such difference is statistically significant.

The results from the telephone and web sample are consistent with previous evidence from the literature for both samples. Women are on average more satisfied than men; life satisfaction is higher for people in the age bracket 56-64 compared to the age category 18-29; the category "Other, not working", which captures the unemployed, is a significant and negative predictor of being satisfied with one's life; home-making is positively associated with life satisfaction; rich people report on average a higher life satisfaction than poor people. The results from the two samples have the same sign and significance level, but their magnitude differs. This is particularly the case for income, in case of all five measures of well-being, and age for the proxy "If I could live my life again I would not change anything" (see Tables C1 - C4 in Appendix C). In sum, we find that survey mode affects regression coefficients, but it does not change their signs and significance. Explaining why income is more strongly correlated to well-being in web rather than in telephone surveys is beyond the scope of present paper. We speculate that the answers provided in web surveys may be more accurate.

6 Conclusions

Surveys are a precious source of information about many social, economic, and political issues, including people's subjective well-being. Assessing how is life, whether it has improved, and which factors matter for people's well-being is relevant for both academic and policy-making purposes. To these aims, the prompt availability of high quality data is pivotal. Yet, the increasing costs of surveys impose a trade-off between quality and quantity: it is expensive and increasingly difficult to run timely and high quality surveys. Web surveys, i.e. surveys administered via the Internet, provide a possible solution to such conundrum: the wide penetration of Internet in many Western countries, combined with lower administration costs, time-effectiveness, and automatisation, create a possibility to collect reliable and affordable nationally representative data. However, little is known about how the use of telephone and web surveys on nationally representative samples affects responses to subjective questions, such as those about people's well-being.

In particular, previous literature questioned the comparability of the answers to subjective data collected using different survey modes (Dolan & Kavetsos, 2012). This paper used unique data from the Luxembourgish Global Entrepreneur-

Table 8
Web survey mode as a predictor of subjective well-being scores. Marginal effects from multinomial logit with CEM.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
The conditions of my life are excellent (n=2,131)					
Web survey mode	0.03***	0.06***	−0.07***	−0.06*	0.04
	(0.008)	(0.013)	(0.019)	(0.025)	(0.021)
In most ways my life is close to my ideal (n=2124)					
Web survey mode	0.02**	0.04***	−0.14***	0.03	0.05*
	(0.009)	(0.013)	(0.020)	(0.024)	(0.019)
So far I have obtained the important things I want in life (n=2,133)					
Web survey mode	0.02**	0.03*	−0.12***	−0.02	0.09***
	(0.007)	(0.015)	(0.018)	(0.024)	(0.022)
If I could live my life again, I would not change anything (n=2101)					
survey mode	0.01	0.10***	−0.19***	0.09***	−0.01
	(0.015)	(0.018)	(0.020)	(0.023)	(0.016)

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home. Coefficients for control variables are omitted for brevity. The complete list of results is available from Table **??** - **??** in Appendix D. CEM for survey data computed according to Riillo (2017). Observations differ because of missing values in the dependent variable. Robust standard errors in parentheses.
* $p < .10$ ** $p < .05$ *** $p < .01$

ship Monitor from years 2013–2015 to test whether the survey mode affects people's evaluations of their subjective well-being by comparing web and telephone surveys. The data at hand are unique because Luxembourg is the only country that consistently collected data on well-being using five different measures, and administered a web survey to nearly half of the sample. We used Oaxaca-Blinder decomposition and multinomial logit with Coarsened Exact Matching to establish whether and in which direction the participation in the web survey alters people's responses about their well-being.

Our results showed a downward bias in web, compared to telephone survey. Respondents of the web survey reported on average slightly, yet significantly, lower well-being than respondents of the telephone survey. Our analysis showed that this difference is not driven by different socio-demographic characteristic of the two samples. Ceteris paribus, respondents of the web survey had a higher probability to choose the two lowest categories of well-being than telephone respondents. In three out of five measures web survey respondents also had a higher probability to choose the two highest categories. Finally, web survey respondents had a lower probability to choose the neutral ("neither agree nor disagree") category. Hence, our first conclusion is that, for the purpose of descriptive statistics, the choice of survey mode alters people's answers on well-being.

We also tested whether survey mode affects the statistical inference. To do this, we compared the results from a happiness regression estimated for telephone and web survey respondents. The five measures of well-being were regressed over a set of socio-demographic control variables including age, education, gender, income, employment status, and their interaction with survey mode. The results showed little differences between the coefficients estimated for the two samples. In sum, our second conclusion is that the effects of survey mode for inference are negligible, based on observable variables.

Previous studies suggest three mechanisms to explain these results. First, web respondents tend to provide less socially desirable answers than telephone respondents, which, in many social contexts may result in declaring lower life satisfaction. Second, web respondents tend to choose more often negative answers. Third, web respondents are more prone to the primacy effect, i.e. they tend to choose more often the first proposed answer. This could explain the higher share of negative answers in our web sample. Telephone respondents, on the other hand, are more prone to the recency effect, i.e. they choose more often the last category mentioned by the interviewer. In our case this could explain the higher share of positive answers among telephone respon-

Table 9
Average marginal effects of the correlates of life satisfaction from a multinomial logit with all variables interacted with survey mode.

	Telephone sample	Web sample	Difference Teleph.-Web
Women	0.03	0.02	−0.01
	(0.020)	(0.021)	(0.028)
Age: (ref: 18-29)			
30-39	−0.09***	−0.06*	0.04
	(0.035)	(0.031)	(0.047)
40-47	−0.13***	−0.07**	0.05
	(0.033)	(0.030)	(0.045)
48-55	−0.11***	−0.09***	0.02
	(0.034)	(0.031)	(0.046)
56-64	−0.05	−0.01	0.04
	(0.041)	(0.042)	(0.059)
Education: (ref: elementary)			
secondary	0.00	−0.01	−0.01
	(0.036)	(0.044)	(0.056)
master craftsman	0.07	−0.04	−0.10
	(0.055)	(0.059)	(0.080)
bachelor	−0.02	−0.05	−0.03
	(0.036)	(0.044)	(0.057)
master	−0.03	−0.04	−0.01
	(0.039)	(0.046)	(0.060)
Employment status: (ref: full-time)			
part-time employed	0.01	0.07**	0.06
	(0.030)	(0.033)	(0.045)
self-employed	0.01	−0.01	−0.03
	(0.040)	(0.040)	(0.056)
retired	−0.01	0.04	0.05
	(0.039)	(0.048)	(0.061)
home maker	0.05	0.09	0.04
	(0.055)	(0.067)	(0.087)
student	−0.01	0.05	0.06
	(0.052)	(0.052)	(0.073)
not working, other	−0.15***	−0.21***	−0.07
	(0.053)	(0.042)	(0.068)
Household income: (ref: first quintile)			
second quintile	0.09***	0.17***	0.09**
	(0.029)	(0.030)	(0.042)
third quintile	0.16***	0.24***	0.08*
	(0.029)	(0.030)	(0.042)
fourth quintile	0.14***	0.29***	0.15***
	(0.032)	(0.031)	(0.044)
fifth quintile	0.21***	0.40***	0.18***
	(0.030)	(0.027)	(0.040)
Year: (ref: 2013)			
2014	−0.02	−0.01	0.02
	(0.021)	(0.025)	(0.033)
2015	0.02	−0.03	−0.05*
	(0.023)	(0.023)	(0.032)
Obs.	2166	2394	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. The marginal effects are computed with reference to the category "strongly agree". Estimates are computed interacting all variables with the survey mode. Weighted estimations and robust standard error in parenthesis.
* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

dents. Available studies, however, do not provide any explanation for why web respondents choose the neutral category less often than telephone respondents, and why for three out of five measures they choose positive answers more often. We speculate that answering a survey through a computer allows the respondent to make a more informed choice. At any rate, explaining these results goes beyond the purpose of present study.

One of the limitations of present study is the unavailability of a sample made of respondents who responded to both web and telephone surveys. Laboratory experiments could help address this issue, yet we believe that our large and nationally representative sample adds value to our conclusions. In an ideal setting, the availability of large survey in which the same respondents use at least two different survey modes would allow to better disentangle the effect of survey mode from the effect of sample composition. Unfortunately, to the best of our knowledge, such data are still not available. Furthermore, future research should check whether the results change when using variables with different answering scales. Hopefully, the availability of new and more refined data will allow to address these points.

Present results support the view that, compared to telephone surveys, web surveys are a convenient and reliable mode to collect subjective data, such as those on people's well-being. They are also safe to run statistical inference. However, practitioners should be careful when using these data for descriptive purposes, especially when combining data collected with various survey modes.

Acknowledgements

This article reflects the view of the authors and do not engage in any way STATEC research, ANEC and funding partners. The authors gratefully acknowledge the support of the Observatoire de la Compétitivité, Ministère de l'Economie, DG Compétitivité, Luxembourg and STATEC. Possible errors or omissions are entirely the responsibility of the authors who contributed equally to this work.

References

Aggarwal, V. A. & Hsu, D. H. (2014). Entrepreneurial exits and innovation. *Management Science*, *60*(4), 867–887.

Bartolini, S. & Sarracino, F. (2015). The dark side of Chinese growth: declining social capital and well-being in times of economic boom. *World Development*, *74*, 333–351.

Bason, J. J. (2000). *Comparison of telephone, mail, web, and IVR surveys of drug and alcohol use among University of Georgia students.* Washington D.C.: American Association of Public Opinion Research.

Bech, M. & Kristensen, M. B. (2009). Differential response rates in postal and web-based surveys in older respondents. *Survey Research Methods*, *3*(1), 1–6.

Berrens, R. P., Bohara, A. K., Jenkins-Smith, H., Silva, C., & Weimer, D. L. (2003). The advent of Internet surveys for political research: a comparison of telephone and Internet samples. *Political Analysis*, *11*(1), 1–22.

Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, *78*(2), 161–188.

Bethlehem, J. (2015). Essay: sunday shopping-the case of three surveys. *Survey Research Methods*, *9*(3), 221–230.

Blackwell, M. (2009). Cem: coarsened exact matching in Stata. *Stata Journal*, *9*(4), 524–546.

Blanchflower, D. & Oswald, A. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine*, *66*, 1733–1749.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, *8*(4), 436–455.

Braunsberger, K., Wybenga, H., & Gates, R. (2007). A comparison of reliability between telephone and web-based surveys. *Journal of Business Research*, *60*(7), 758–764.

Chang, L. & Krosnick, J. A. (2009). National surveys via RDD telephone interviewing versus the internet comparing sample representativeness and response quality. *Public Opinion Quarterly*, *73*(4), 641–678.

Christian, L. M., Dillman, D. A., & Smyth, J. D. (2008). The effects of mode and format on answers to scalar questions in telephone and web surveys. *Advances in Telephone Survey Methodology*, *12*, 250–275.

Couper, M. P. (2000). Review: web surveys: a review of issues and approaches. *Public Opinion Quarterly*, *64*(4), 464–494.

Couper, M. P. (2013). Is the sky falling? New technology, changing media, and the future of surveys. *Survey Research Methods*, *7*(3), 145–156.

Dever, J. A., Rafferty, A., & Valliant, R. (2008). Internet surveys: can statistical adjustments eliminate coverage bias. *Survey Research Methods*, *2*(2), 47–62.

Dillman, D. A. & Christian, L. M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods*, *17*(1), 30–52.

Dillman, D. A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J., & Messer, B. L. (2009). Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Social Science Research*, *38*(1), 1–18.

Dolan, P. & Kavetsos, G. (2012). Happy talk: mode of administration effects on subjective well-being. *Journal of Happiness Studies*, 1–19.

Flemming, G. & Sonner, M. (1999). *Can internet polling work? Strategies for conducting public opinion surveys online.* Paper presented on the Annual Meeting of the American Association for Public Opinion Research, St. Petersburg Beach, FL.

Fricker, S., Galesic, M., Tourangeau, R., & Yan, T. (2005). An experimental comparison of web and telephone surveys. *Public Opinion Quarterly*, *69*(3), 370–392.

GEM. (2016). *Global Entrepreneurship Monitor: Adult Population Survey.* Data for Luxembourg 2013 - 2015, Global Entrepreneurship Research Association.

Helliwell, J. F. & Barrington-Leigh, C. P. (2010). Viewpoint: Measuring and understanding subjective well-being. *Canadian Journal of Economics*, *43*(3), 729–753.

Hudson, D., Seah, L.-H., Hite, D., & Haab, T. (2004). Telephone presurveys, self-selection, and non-response bias to mail and internet surveys in economic research. *Applied Economics Letters*, *11*(4), 237–240.

Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, *106*(493), 345–361.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened Exact Matching. *Political Analysis*, *20*(1), 1–24.

Jäckle, A., Lynn, P., & Burton, J. (2015). Going online with a face-to-face household panel: effects of a mixed mode design on item and unit non-response. *Survey Research Methods*, *9*(1), 57–70.

Johnson, T. P., Fendrich, M., Shaligram, C., Garcy, A., & Gillespie, S. (2000). An evaluation of the effects of interviewer characteristics in an RDD telephone survey of drug use. *Journal of Drug Issues*, *30*(1), 887–892.

Kahneman, D. & Krueger, A. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, *20*, 3–24.

Kiesler, S. & Sproull, L. S. (1986). Response effects in the electronic survey. *Public Opinion Quarterly*, *50*(3), 402–413.

Klausch, T., Hox, J. J., & Schouten, B. (2013). Measurement effects of survey mode on the equivalence of attitudinal rating scale questions. *Sociological Methods & Research*, *42*(3), 227–263.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys the effects of mode and question sensitivity. *Public Opinion Quarterly*, *72*(5), 847–865.

Layard, R. (2005). *Happiness: lessons from a new science.* New York: Penguin Press.

Long, J. S. & Freese, J. (2006). College Station: Stata Press.

Malhotra, N. (2008). Completion time and response order effects in web surveys. *Public Opinion Quarterly*, *72*(5), 914–934.

Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources*, *23*, 279–295.

Nigg, C. R., Motl, R. W., Wong, K. T., Yoda, L. U., McCurdy, D. K., Paxton, R., ... Dishman, R. K. (2009). The impact of mixed survey modes on physical activity and fruit/vegetable consumption: a longitudinal study. *Survey Research Methods*, *3*(2), 81–90.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, *14*(3), 693–709.

Oaxaca, R. L. & Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of econometrics*, *61*(1), 5–21.

Parks, K. A., Pardi, A. M., & Bradizza, C. M. (2006). Collecting data on alcohol use and alcohol-related victimization: a comparison of telephone and web-based survey methods. *Journal of Studies on Alcohol*, *67*(2), 318–323.

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.

Roster, C. A., Rogers, R. D., Albaum, G., & Klein, D. (2004). A comparison of response characteristics from web and telephone surveys. *International Journal of Market Research*, *46*, 359–374.

Sarracino, F. (2013). Determinants of subjective well-being in high and low income countries: do happiness equations differ across countries? *The Journal of Socio-Economics*, *42*, 51–66.

Schimmack, U., Krause, P., Wagner, G., & Schupp, J. (2010). Stability and change of well-being: An experimentally enhanced latent state-trait-error analysis. *Social Indicators Research*, *95*(1), 19–31.

Schneider, L. & Schimmack, U. (2009). Self-informant agreement in well-being ratings: a meta-analysis. *Social Indicators Research*, *94*(3), 363–376.

Schonlau, M., Asch, B. J., & Du, C. (2003). Web surveys as part of a mixed-mode strategy for populations that cannot be contacted by e-mail. *Social Science Computer Review*, *21*(2), 218–222.

Schwarz, N. & Strack, F. (1999). Reports of subjective well-being: judgmental processes and their methodological implications. In D. Kahneman & N. Schwarz (Eds.), *Well-being: the foundations of hedonist psychology*. New York: Russell Sage Foundation,

Sidney, J., Coberley, C., Pope, J., & Wells, A. (2015). Extending coarsened exact matching to multiple cohorts: an application to longitudinal well-being program evaluation within an employer population. *Health Services and Outcomes Research Methodology*, *15*(2), 136–156.

Sinning, M., Hahn, M., & Bauer, T. K. (2008). The Blinder-Oaxaca decomposition for nonlinear regression models. *The Stata Journal*, *8*(4), 480–492.

Smyth, J., Olson, K., & Kasabian, A. S. (2014). The effect of answering in a preferred versus a non-preferred survey mode on measurement. *Survey Research Methods*, *8*(3), 137–152.

Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position, and order interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly*, *68*(3), 368–393.

van Reekum, C., Urry, H., Johnstone, T., Thurow, M., Frye, C., Jackson, C., … Davidson, R. (2007). Individual differences in amygdala and ventromedial prefrontal cortex activity are associated with evaluation speed and psychological well-being. *Journal of Cognitive Neuroscience*, *19*(2), 237–248.

Wanous, J. & Hudy, M. (2001). Single-item reliability: a replication and extension. *Organizational Research Methods*, *4*, 361–375.

We Are Social. (2016). Digital in 2016. Technical report. Retrieved from http://wearesocial.com/uk/special-reports/digital-in-2016

Wyatt, J. C. (2000). When to use web-based surveys. *Journal of the American Medical Informatics Association*, *7*(4), 426–430.

Yeager, D. S., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., & Wang, R. (2011). Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, *75*(4), 709–747.

Yun, G. W. & Trumbo, C. W. (2000). Comparative response to a survey executed by post, e-mail, & web form. *Journal of Computer-Mediated Communication*, *6*(1).

Zhang, C. & Conrad, F. (2014). Speeding in web surveys: the tendency to answer very fast and its association with straightlining. *Survey Research Methods*, *8*(2), 127–135.

Appendix A
Distribution of socio-demographic variables

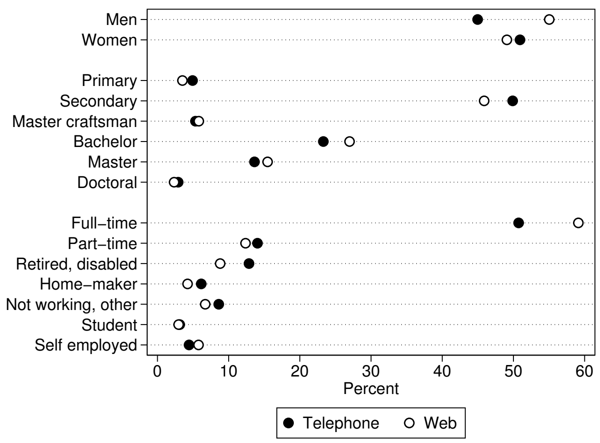

Figure A1. Distribution of gender, education and employment status by survey mode.

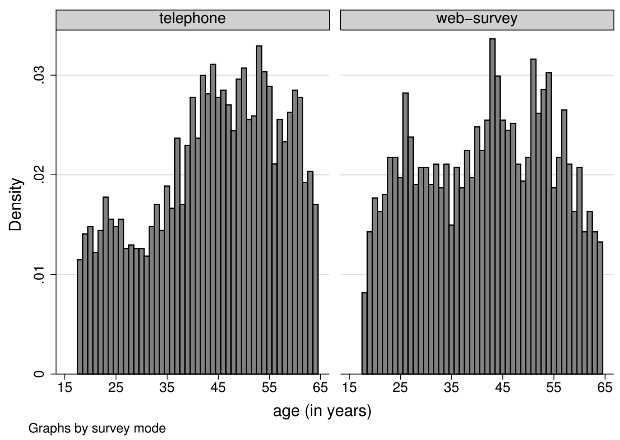

Figure A2. Distribution of age by survey mode.

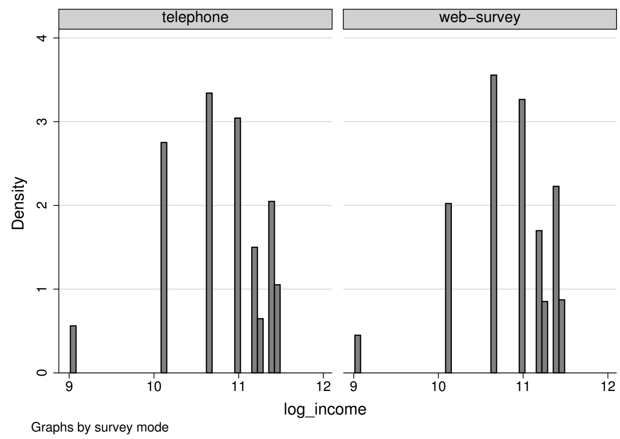

Figure A3. Distribution of income by survey mode.

Appendix B
Distribution of alternative measures of well-being

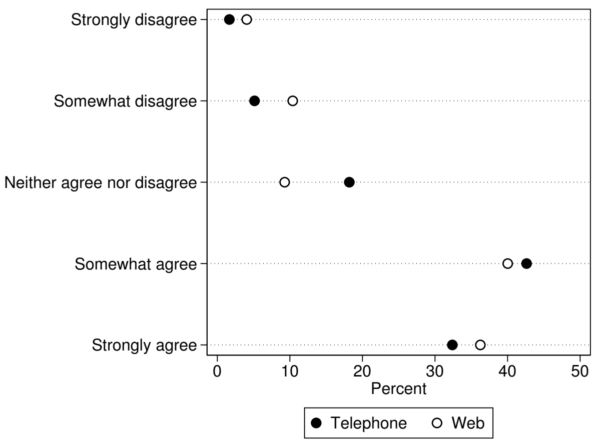

Figure B1. "So far I have obtained the important things I want in life"
Note:
$\text{Refusal}_{tel} = 0.04\%$; $\text{Don't know}_{tel} = 0.07\%$;
$\text{Refusal}_{web} = 0.43\%$; $\text{Don't know}_{web} = 0.61\%$;
$\mu_{tel} = 3.98$, $\mu_{web} = 3.97$;
t-test($\mu_{tel} - \mu_{web} = 0$) not significant.

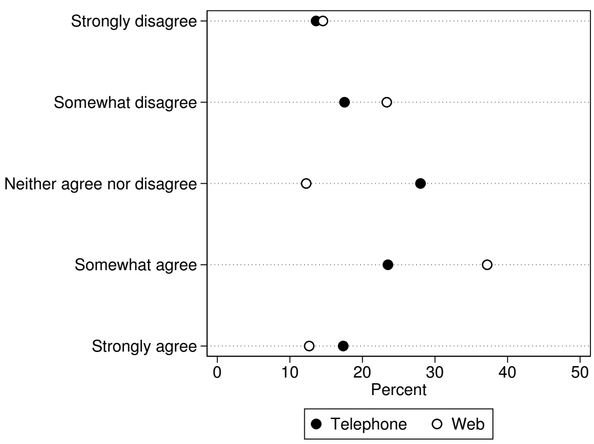

Figure B2. "If I could live my life again, I would not change anything"
Note:
$\text{Refusal}_{tel} = 0.18\%$; $\text{Don't know}_{tel} = 0.57\%$;
$\text{Refusal}_{web} = 0.58\%$; $\text{Don't know}_{web} = 3.70\%$;
$\mu_{tel} = 3.13$, $\mu_{web} = 3.11$;
t-test($\mu_{tel} - \mu_{web} = 0$) not significant

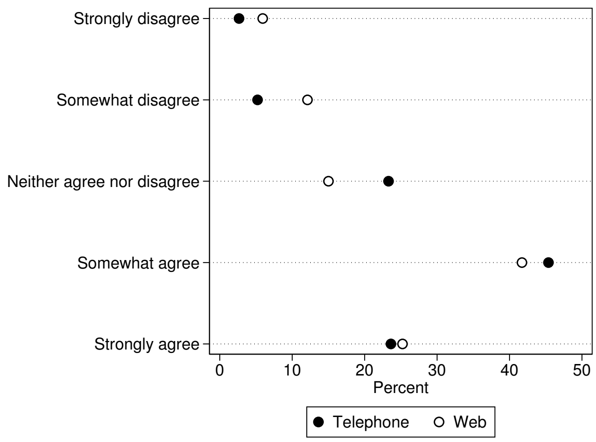

Figure B3. "The conditions of my life are excellent"
Note:
$\text{Refusal}_{tel} = 0.00\%$; $\text{Don't know}_{tel} = 0.49\%$;
$\text{Refusal}_{web} = 0.04\%$; $\text{Don't know}_{web} = 0.40\%$;
$\mu_{tel} = 3.82$, $\mu_{web} = 3.71$;
t-test($\mu_{tel} - \mu_{web} = 0$) significant ($p < 0.01$)

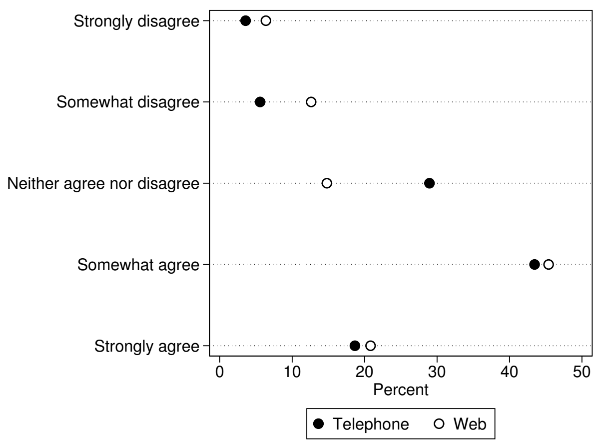

Figure B4. "In most ways my life is close to my ideal"
Note:
$\text{Refusal}_{tel} = 0.00\%$; $\text{Don't know}_{tel} = 0.67\%$;
$\text{Refusal}_{web} = 0.32\%$; $\text{Don't know}_{web} = 0.83\%$;
$\mu_{tel} = 3.68$, $\mu_{web} = 3.64$;
t-test($\mu_{tel} - \mu_{web} = 0$) significant ($p < 0.05$)

Appendix C
Marginal effects of the correlates of subjective well-being.

(Tables follow on next page)

Table C1

Average marginal effects of the correlates of 'the conditions of my life are excellent' from a multinomial logit with all variables interacted with survey mode.

	Telephone sample	Web sample	Difference Teleph.-Web
Women	-0.01	-0.03^{**}	-0.02
	(0.014)	(0.016)	(0.022)
Age: (ref: 18-29)			
age 30-39	-0.02	-0.01	0.01
	(0.025)	(0.024)	(0.035)
age 40-47	-0.04^{*}	-0.03	0.01
	(0.024)	(0.023)	(0.033)
age 48-55	-0.00	-0.06^{**}	-0.05
	(0.026)	(0.022)	(0.034)
age 56-64	0.04	0.05	0.01
	(0.032)	(0.034)	(0.047)
Education (ref: elementary)			
secondary	-0.01	-0.05	-0.04
	(0.023)	(0.033)	(0.040)
master craftsman	-0.03	-0.08^{**}	-0.05
	(0.032)	(0.035)	(0.048)
bachelor	0.03	-0.07^{**}	-0.10^{**}
	(0.026)	(0.031)	(0.041)
master	0.03	-0.02	-0.05
	(0.028)	(0.035)	(0.045)
Employment status (ref: full-time employed)			
part-time employed	-0.02	0.04	0.07^{**}
	(0.018)	(0.027)	(0.032)
self-employed	-0.02	-0.04	-0.02
	(0.029)	(0.033)	(0.044)
retired	-0.00	0.00	0.01
	(0.028)	(0.033)	(0.044)
home maker	0.05	0.02	-0.04
	(0.042)	(0.050)	(0.066)
student	0.10^{**}	0.16^{***}	0.06
	(0.047)	(0.049)	(0.068)
not working, other	-0.07^{*}	-0.14^{***}	-0.07
	(0.037)	(0.028)	(0.047)
Household income (ref: first quintile)			
second income quintile	0.12^{***}	0.16^{***}	0.03
	(0.023)	(0.026)	(0.035)
third income quintile	0.18^{***}	0.25^{***}	0.07^{**}
	(0.024)	(0.026)	(0.036)
fourth income quintile	0.21^{***}	0.30^{***}	0.09^{**}
	(0.027)	(0.029)	(0.039)
fifth income quintile	0.32^{***}	0.49^{***}	0.17^{***}
	(0.026)	(0.027)	(0.037)
Year (ref: 2013)			
year 2014	-0.05^{***}	-0.01	0.04^{*}
	(0.015)	(0.018)	(0.024)
year 2015	-0.01	-0.03	-0.01
	(0.016)	(0.017)	(0.024)
Obs.	2168	2400	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. The marginal effects are computed with reference to the category "strongly agree". Estimates are computed interacting all variables with the survey mode. Weighted estimations and robust standard error in parenthesis.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table C2
Average marginal effects of the correlates of 'in most ways my life is close to my ideal' from a multinomial logit with all variables interacted with survey mode.

	Telephone sample	Web sample	Difference Teleph.-Web
Women	0.01	0.01	−0.01
	(0.012)	(0.015)	(0.019)
Age: (ref: 18-29)			
age 30 - 39	−0.05**	0.02	0.07**
	(0.020)	(0.024)	(0.032)
age 40 - 47	−0.07***	−0.03	0.05*
	(0.018)	(0.021)	(0.028)
age 48 - 55	−0.07***	−0.04*	0.03
	(0.019)	(0.021)	(0.028)
age 56 - 64	−0.02	0.04	0.05
	(0.025)	(0.031)	(0.040)
Education (ref: elementary)			
secondary	−0.04	−0.03	0.01**
	(0.023)	(0.029)	(0.037)
master craftsman	−0.01	−0.03	−0.02
	(0.032)	(0.034)	(0.047)
bachelor	−0.01	−0.04	−0.04
	(0.024)	(0.029)	(0.037)
master	−0.02	−0.00	0.02
	(0.024)	(0.032)	(0.041)
Employment status (ref: full-time employed)			
part-time employed	−0.01	0.06**	0.06
	(0.017)	(0.025)	(0.030)
self-employed	−0.00	−0.03	−0.03***
	(0.025)	(0.031)	(0.040)
retired	−0.02	0.04	0.06
	(0.021)	(0.035)	(0.041)
home maker	0.02	0.01	−0.0
	(0.033)	(0.049)	(0.058)
student	−0.03	0.06	0.09*
	(0.026)	(0.041)	(0.049)
not working, other	−0.09***	−0.13***	−0.05**
	(0.023)	(0.023)	(0.033)
Household income (ref: first quintile)			
second income quintile	0.07***	0.14***	0.07***
	(0.020)	(0.027)	(0.034)
third income quintile	0.09***	0.22***	0.12***
	(0.021)	(0.029)	(0.035)
fourth income quintile	0.11***	0.26***	0.15***
	(0.023)	(0.032)	(0.039)
fifth income quintile	0.19***	0.36***	0.17***
	(0.025)	(0.031)	(0.040)
Year (ref: 2013)			
year 2014	−0.01	−0.04**	−0.02
	(0.013)	(0.016)	(0.021)
year 2015	−0.00	−0.04***	−0.04*
	(0.014)	(0.016)	(0.021)
Obs.	2168	2400	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. The marginal effects are computed with reference to the category "strongly agree". Estimates are computed interacting all variables with the survey mode. Weighted estimations and robust standard error in parenthesis.
* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table C3

Average marginal effects of the correlates of 'If I could live my life again, I would not change anything' from a multinomial logit with all variables interacted with survey mode.

	Telephone sample	Web sample	Difference Teleph.-Web
Women	0.00	0.02	0.01
	(0.011)	(0.011)	(0.015)
Age: (ref: 18-29)			
age 30 - 39	−0.04**	−0.01	0.03
	(0.019)	(0.016)	(0.025)
age 40 - 47	−0.08***	−0.03	0.06**
	(0.017)	(0.015)	(0.023)
age 48 - 55	−0.08***	−0.04***	0.04*
	(0.017)	(0.015)	(0.022)
age 56 - 64	−0.05***	−0.00	0.05*
	(0.020)	(0.020)	(0.028)
Education (ref: elementary)			
secondary	−0.05**	0.00	0.05
	(0.022)	(0.023)	(0.032)
master craftsman	−0.00	−0.00	−0.00
	(0.031)	(0.031)	(0.044)
bachelor	−0.02	−0.00	0.02
	(0.021)	(0.023)	(0.031)
master	−0.02	0.03	0.05
	(0.022)	(0.027)	(0.035)
Employment status (ref: full-time employed)			
part-time employed	−0.01	0.04**	0.05**
	(0.015)	(0.018)	(0.023)
self-employed	0.06	0.01	−0.05
	(0.037)	(0.023)	(0.044)
retired	−0.03*	0.04	0.08**
	(0.017)	(0.027)	(0.032)
home maker	0.03	−0.01	−0.04
	(0.034)	(0.031)	(0.046)
student	−0.03	0.00	0.03
	(0.022)	(0.025)	(0.034)
not working, other	−0.00	−0.06**	−0.06
	(0.035)	(0.023)	(0.042)
Household income (ref: first quintile)			
second income quintile	0.01	0.06***	0.05*
	(0.017)	(0.021)	(0.027)
third income quintile	0.04**	0.09***	0.042
	(0.018)	(0.022)	(0.029)
fourth income quintile	0.05***	0.12***	0.06**
	(0.020)	(0.024)	(0.032)
fifth income quintile	0.08***	0.15***	0.08**
	(0.021)	(0.026)	(0.034)
Year (ref: 2013)			
year 2014	−0.00	−0.00	−0.00
	(0.011)	(0.012)	(0.017)
year 2015	0.00	0.00	−0.00
	(0.013)	(0.012)	(0.017)
Obs.	2168	2400	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. The marginal effects are computed with reference to the category "strongly agree". Estimates are computed interacting all variables with the survey mode. Weighted estimations and robust standard error in parenthesis.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table C4

Average marginal effects of the correlates of 'so far I have obtained the important things I want in my life' from a multinomial logit with all variables interacted with survey mode.

	Telephone sample	Web sample	Difference Teleph.-Web
Women	0.03*	0.05**	0.02
	(0.018)	(0.021)	(0.028)
Age: (ref: 18-29)			
age 30 - 39	0.02	−0.02	−0.04
	(0.035)	(0.033)	(0.048)
age 40 - 47	0.04	0.00	−0.03
	(0.034)	(0.032)	(0.047)
age 48 - 55	0.04	−0.01	−0.05
	(0.035)	(0.032)	(0.048)
age 56 - 64	0.09**	0.04	−0.06
	(0.041)	(0.042)	(0.059)
Education (ref: elementary)			
secondary	0.02	0.07	0.05
	(0.032)	(0.042)	(0.053)
master craftsman	0.10*	0.04	−0.06
	(0.051)	(0.056)	(0.076)
bachelor	0.02	0.06	0.04
	(0.034)	(0.046)	(0.057)
master	0.02	0.03	0.01
	(0.038)	(0.049)	(0.062)
Employment status (ref: full-time employed)			
part-time employed	0.02	0.03	0.02
	(0.026)	(0.034)	(0.043)
self-employed	0.04	−0.01	−0.05
	(0.040)	(0.042)	(0.058)
retired or disabled	−0.02	0.03	0.05
	(0.032)	(0.045)	(0.055)
home maker	0.11**	0.03	−0.08
	(0.047)	(0.066)	(0.081)
student	0.06	0.03	−0.03
	(0.054)	(0.055)	(0.078)
not working, other	−0.18***	−0.12**	0.06
	(0.030)	(0.054)	(0.061)
Household income (ref: first quintile)			
second income quintile	0.05*	0.16***	0.11***
	(0.027)	(0.030)	(0.040)
third income quintile	0.07***	0.22***	0.15***
	(0.027)	(0.031)	(0.041)
fourth income quintile	0.17***	0.30***	0.13***
	(0.031)	(0.032)	(0.044)
fifth income quintile	0.17***	0.38***	0.21***
	(0.030)	(0.029)	(0.042)
Year (ref: 2013)			
year 2014	−0.04*	−0.01	0.02
	(0.019)	(0.024)	(0.031)
year 2015	0.02	−0.02	−0.03
	(0.021)	(0.023)	(0.031)
Obs.	2168	2400	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. The marginal effects are computed with reference to the category "strongly agree". Estimates are computed interacting all variables with the survey mode. Weighted estimations and robust standard error in parenthesis.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Appendix D
Web survey mode as a predictor of subjective well-being. Complete results.

Table D1
Web survey mode as a predictor of life satisfaction scores. Marginal effects from multinomial logit with CEM.

I am satisfied with my life:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Web survey mode	0.02**	0.05***	−0.04*	−0.06*	0.03
	(0.007)	(0.012)	(0.015)	(0.024)	(0.023)
Woman	0.01	−0.01	−0.02	−0.01	0.04
	(0.007)	(0.014)	(0.018)	(0.028)	(0.027)
Age (in years)	−0.00	0.00	0.01	0.02	−0.03**
	(0.002)	(0.006)	(0.006)	(0.011)	(0.010)
Age squared / 100	0.00	−0.00	−0.01	−0.02	0.03*
	(0.003)	(0.007)	(0.008)	(0.013)	(0.012)
Education (ref: elementary)					
secondary	−0.01	0.06	−0.05	0.06	−0.06
	(0.014)	(0.050)	(0.034)	(0.078)	(0.071)
master craftsman	−0.24***	0.08	−0.07	0.11	0.12
	(0.043)	(0.058)	(0.056)	(0.100)	(0.092)
bachelor	−0.00	0.04	−0.03	0.04	−0.05
	(0.015)	(0.051)	(0.036)	(0.081)	(0.074)
master	−0.00	0.08	−0.05	0.05	−0.08
	(0.016)	(0.051)	(0.041)	(0.085)	(0.077)
Employment status (ref: full-time employed)					
part-time employed	−0.01	−0.00	0.03	−0.03	0.01
	(0.013)	(0.021)	(0.024)	(0.045)	(0.044)
self-employed	0.06**	−0.63***	−1.19***	1.02***	0.73***
	(0.022)	(0.068)	(0.093)	(0.155)	(0.155)
retired	−0.01	−0.05	0.04	−0.03	0.05
	(0.013)	(0.032)	(0.039)	(0.061)	(0.057)
home maker	0.01	0.05	0.03	−0.34**	0.26**
	(0.015)	(0.032)	(0.049)	(0.107)	(0.092)
student	−0.02	−0.00	−0.02	0.09	−0.06
	(0.020)	(0.048)	(0.045)	(0.079)	(0.075)
not working, other	−0.18***	0.14*	−1.24***	1.05***	0.23
	(0.035)	(0.059)	(0.101)	(0.206)	(0.239)
Household income (log)	−0.01	−0.05***	−0.08***	−0.07*	0.21***
	(0.008)	(0.012)	(0.016)	(0.031)	(0.031)
Year (ref: 2013)					
year 2014	0.00	−0.00	0.01	−0.02	0.01
	(0.007)	(0.013)	(0.019)	(0.031)	(0.029)
year 2015	0.00	−0.03	0.03	−0.01	0.01
	(0.007)	(0.014)	(0.018)	(0.029)	(0.028)
Observations	2,134	2,134	2,134	2,134	2,134

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home. CEM for survey data computed according to Riillo (2017) Robust standard errors in parentheses.
* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table D2

Web survey mode as a predictor of 'the conditions of my life are excellent' scores. Marginal effects from multinomial logit with CEM.

The conditions of my life are excellent:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Web survey mode	0.03***	0.06***	−0.07***	−0.06*	0.04
	(0.008)	(0.013)	(0.019)	(0.025)	(0.021)
Woman	0.03**	0.00	−0.03	0.00	0.00
	(0.008)	(0.014)	(0.022)	(0.028)	(0.023)
Age (in years)	0.01	−0.01	0.01	0.01	−0.02
	(0.003)	(0.005)	(0.008)	(0.011)	(0.009)
Age squared / 100	−0.01	0.01	−0.01	−0.01	0.02
	(0.004)	(0.006)	(0.010)	(0.013)	(0.011)
Education (ref: elementary)					
secondary	−0.03*	0.14	0.03	−0.04	−0.09
	(0.014)	(0.072)	(0.055)	(0.078)	(0.058)
master craftsman	−0.03	0.15	−0.00	0.03	−0.15
	(0.025)	(0.080)	(0.072)	(0.099)	(0.081)
bachelor	−0.03	0.13	0.03	−0.08	−0.05
	(0.015)	(0.073)	(0.058)	(0.081)	(0.060)
master	−0.03	0.12	0.04	−0.08	−0.05
	(0.019)	(0.075)	(0.062)	(0.085)	(0.063)
Employment status (ref: full-time employed)					
part-time employed	−0.01	−0.01	0.06	0.09	−0.12**
	(0.011)	(0.022)	(0.033)	(0.047)	(0.045)
self-employed	0.02	−0.01	−0.13	0.11	0.01
	(0.020)	(0.074)	(0.161)	(0.161)	(0.152)
retired	0.02	−0.10**	−0.03	0.12*	−0.01
	(0.017)	(0.034)	(0.047)	(0.060)	(0.050)
home maker	0.01	−0.05	−0.01	−0.03	0.08
	(0.014)	(0.048)	(0.065)	(0.096)	(0.084)
student	−0.02	−0.08	−0.13	0.11	0.12
	(0.025)	(0.044)	(0.074)	(0.084)	(0.066)
not working, other	−0.22***	0.23**	−2.07***	1.55***	0.51**
	(0.036)	(0.080)	(0.112)	(0.172)	(0.194)
Household income (log)	−0.03***	−0.10***	−0.15***	0.07*	0.21***
	(0.008)	(0.014)	(0.022)	(0.032)	(0.031)
Year (ref: 2013)					
year 2014	−0.00	0.01	0.01	0.02	−0.03
	(0.008)	(0.016)	(0.024)	(0.031)	(0.026)
year 2015	−0.00	0.01	0.01	0.01	−0.03
	(0.008)	(0.015)	(0.022)	(0.030)	(0.025)
Observations	2,805	2,805	2,805	2,805	2,805

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home. CEM for survey data computed according to Riillo (2017). CEM matched and regression observations differ because of missing values in the dependent variable. Robust standard errors in parentheses.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table D3
Web survey mode as a predictor of 'in most ways my life is close to my ideal' scores. Marginal effects from multinomial logit with CEM.

In most ways my life is close to my ideal:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Web survey mode	0.02**	0.04***	−0.14***	0.03	0.05*
	(0.009)	(0.013)	(0.020)	(0.024)	(0.019)
Woman	0.00	−0.02	−0.01	0.02	0.01
	(0.011)	(0.015)	(0.024)	(0.028)	(0.022)
Age (in years)	0.00	0.01	0.02	−0.00	−0.02**
	(0.004)	(0.006)	(0.009)	(0.011)	(0.008)
Age squared / 100	−0.00	−0.01	−0.02	0.00	0.02*
	(0.005)	(0.008)	(0.011)	(0.012)	(0.010)
Education (ref: elementary)					
secondary	0.05	0.03	0.04	−0.18**	0.06
	(0.034)	(0.045)	(0.057)	(0.070)	(0.064)
master craftsman	0.07	−0.03	0.07	−0.10	−0.00
	(0.043)	(0.059)	(0.074)	(0.095)	(0.086)
bachelor	0.06	0.03	−0.00	−0.17*	0.08
	(0.035)	(0.046)	(0.059)	(0.074)	(0.066)
master	0.06	0.05	−0.04	−0.14	0.08
	(0.037)	(0.047)	(0.064)	(0.078)	(0.067)
Employment status (ref: full-time employed)					
part-time employed	−0.02	0.01	−0.01	0.04	−0.02
	(0.019)	(0.022)	(0.038)	(0.045)	(0.036)
self-employed	0.06*	0.01	−0.09	0.11	−0.08
	(0.023)	(0.086)	(0.161)	(0.168)	(0.187)
retired	−0.00	−0.02	−0.03	−0.03	0.08
	(0.022)	(0.035)	(0.050)	(0.059)	(0.048)
home maker	0.03	−0.01	−0.07	−0.06	0.11
	(0.021)	(0.048)	(0.073)	(0.097)	(0.077)
student	−0.01	−0.08	0.12*	0.01	−0.04
	(0.036)	(0.057)	(0.059)	(0.080)	(0.059)
not working, other	0.11**	0.21**	0.30	1.32***	−1.93***
	(0.037)	(0.069)	(0.166)	(0.165)	(0.110)
Household income (log)	−0.03**	−0.07***	−0.13***	0.08*	0.15***
	(0.011)	(0.016)	(0.023)	(0.032)	(0.028)
Year (ref: 2013)					
year 2014	0.00	0.00	0.02	−0.02	0.00
	(0.011)	(0.015)	(0.025)	(0.031)	(0.024)
year 2015	0.01	0.01	0.03	−0.03	−0.01
	(0.010)	(0.016)	(0.024)	(0.029)	(0.023)
Observations	2,792	2,792	2,792	2,792	2,792

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home. CEM for survey data computed according to Riillo (2017). CEM matched and regression observations differ because of missing values in the dependent variable. Robust standard errors in parentheses.
* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table D4

Web survey mode as a predictor of 'so far I have obtained the important things I want in life' scores. Marginal effects from multinomial logit with CEM.

So far I have obtained the important things I want in life:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Web survey mode	0.02**	0.03*	−0.12***	−0.02	0.09***
	(0.007)	(0.015)	(0.018)	(0.024)	(0.022)
Woman	0.008	−0.01	−0.03	0.01	0.02
	(0.007)	(0.017)	(0.020)	(0.028)	(0.026)
Age (in years)	0.00	0.00	−0.01	0.01	−0.01
	(0.003)	(0.006)	(0.007)	(0.011)	(0.010)
Age squared / 100	−0.00	−0.00	0.01	−0.01	0.01
	(0.003)	(0.007)	(0.009)	(0.013)	(0.012)
Education (ref: elementary)					
secondary	−0.02*	0.04	−0.03	−0.01	0.03
	(0.010)	(0.046)	(0.044)	(0.077)	(0.074)
master craftsman	−0.24***	0.03	−0.00	0.17	0.05
	(0.042)	(0.066)	(0.059)	(0.099)	(0.095)
bachelor	−0.02	0.02	−0.04	−0.02	0.06
	(0.011)	(0.047)	(0.047)	(0.080)	(0.076)
master	−0.02	0.07	−0.04	−0.01	−0.00
	(0.012)	(0.052)	(0.049)	(0.083)	(0.079)
Employment status (ref: full-time employed)					
part-time employed	−0.01	−0.01	0.01	−0.00	0.02
	(0.010)	(0.024)	(0.033)	(0.045)	(0.042)
self-employed	−0.17***	0.19***	−1.59***	0.88***	0.69***
	(0.031)	(0.056)	(0.096)	(0.145)	(0.143)
retired	−0.01	0.00	−0.09	0.04	0.04
	(0.014)	(0.030)	(0.044)	(0.060)	(0.057)
home maker	−0.01	−0.00	0.02	−0.15	0.14
	(0.015)	(0.043)	(0.057)	(0.096)	(0.085)
student	0.00	−0.04	−0.06	0.03	0.07
	(0.016)	(0.056)	(0.052)	(0.080)	(0.074)
not working, other	−0.13***	0.37***	0.57***	2.15***	−2.97***
	(0.029)	(0.061)	(0.110)	(0.149)	(0.119)
Household income (log)	−0.01	−0.04**	−0.08***	−0.07*	0.20***
	(0.008)	(0.015)	(0.021)	(0.030)	(0.030)
Year (ref: 2013)					
year 2014	−0.01	0.03	0.02	0.01	−0.04
	(0.007)	(0.019)	(0.022)	(0.031)	(0.028)
year 2015	−0.01	0.01	0.03	−0.03	0.00
	(0.007)	(0.016)	(0.021)	(0.029)	(0.027)
Observations	2,803	2,803	2,803	2,803	2,803

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home. CEM for survey data computed according to Riillo (2017). CEM matched and regression observations differ because of missing values in the dependent variable. Robust standard errors in parentheses.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table D5

Web survey mode as a predictor of 'if I could live my life again, I would not change anything' scores. Marginal effects from multinomial logit with CEM.

If I could live my life again, I would not change anything:	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Web survey mode	0.01	0.10***	−0.19***	0.09***	−0.010
	(0.015)	(0.018)	(0.020)	(0.023)	(0.016)
Woman	0.010	−0.01	−0.04	−0.02	0.06**
	(0.020)	(0.022)	(0.022)	(0.027)	(0.018)
Age (in years)	0.02*	0.00	0.01	−0.01	−0.01
	(0.008)	(0.008)	(0.009)	(0.010)	(0.007)
Age squared / 100	−0.02	0.00	−0.01	0.01	0.01
	(0.009)	(0.010)	(0.010)	(0.012)	(0.008)
Education (ref: elementary)					
secondary	−0.02	−0.09	0.11	−0.14	0.14
	(0.041)	(0.054)	(0.071)	(0.074)	(0.087)
master craftsman	−0.03	−0.15	0.12	−0.12	0.18
	(0.056)	(0.075)	(0.086)	(0.093)	(0.099)
bachelor	−0.08	−0.05	0.10	−0.10	0.14
	(0.045)	(0.057)	(0.072)	(0.077)	(0.088)
master	−0.07	−0.06	0.06	−0.10	0.17
	(0.048)	(0.060)	(0.075)	(0.081)	(0.088)
Employment status (ref: full-time employed)					
part-time employed	−0.03	0.01	−0.02	0.07	−0.03
	(0.030)	(0.038)	(0.037)	(0.042)	(0.032)
self-employed	−0.06	0.22*	−0.09	−0.15	0.07
	(0.128)	(0.108)	(0.158)	(0.192)	(0.088)
retired	−0.01	−0.01	0.00	0.03	−0.01
	(0.039)	(0.048)	(0.049)	(0.057)	(0.040)
home maker	0.02	−0.06	−0.06	0.04	0.06
	(0.047)	(0.078)	(0.094)	(0.096)	(0.059)
student	0.02	0.02	0.01	0.07	−0.12*
	(0.070)	(0.065)	(0.064)	(0.073)	(0.056)
not working, other	−1.33***	0.36**	0.27	0.56***	0.14
	(0.093)	(0.132)	(0.149)	(0.154)	(0.119)
Household income (log)	−0.08***	−0.09***	0.01	0.11***	0.05*
	(0.020)	(0.024)	(0.024)	(0.030)	(0.023)
Year (ref: 2013)					
year 2014	−0.02	−0.01	−0.00	0.03	0.00
	(0.020)	(0.024)	(0.025)	(0.030)	(0.020)
year 2015	0.04	−0.03	0.01	−0.02	0.00
	(0.019)	(0.023)	(0.024)	(0.028)	(0.019)
Observations	2,753	2,753	2,753	2,753	2,753

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015. *Note:* Estimations for individuals with Internet access and fixed phone line at home. CEM for survey data computed according to Riillo (2017). CEM matched and regression observations differ because of missing values in the dependent variable. Robust standard errors in parentheses.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$