90

Economie et Statistiques

Working papers du STATEC

octobre 2016

Auteurs: Francesco Sarracino, Cesare F.A. Riillo and Malgorzata Mikucka

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

Abstract

We test whether web survey mode alters people's evaluations of their well-being compared to telephone survey mode. We use unique, nationally representative data from Luxembourg in which five questions about subjective well-being were asked using web and telephone surveys. Oaxaca decomposition and multinomial logit with Coarsened Exact Matching indicate that the survey mode alters peoples' 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-surveys are convenient and reliable 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.

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

October 13, 2016

Highlight: In knowledge intensive societies, can the Internet help to collect timely data about people's behaviours, attitudes, preferences, and values? **This study documents that web surveys are a convenient and reliable tool to collect information about people's well-being.**

Background: 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, may overcome such limitations: the wide penetration of Internet, combined with lower administration costs, time-effectiveness, and automatisation, create the 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, researchers know little about the comparability of subjective measures collected on nationally representative samples through telephone and web survey modes. This issue is relevant because current data-bases increasingly rely on mixed data collection modes, and it is important to know whether data issued from different modes are comparable.

Data and methods: We use unique data from the Luxembourgish Global Entrepreneurship Monitor from years 2013–2015 to test whether the web survey mode alters people's evaluations of their subjective well-being compared to telephone survey. 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 use 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.

Findings: 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 telephone surveys. Moreover, for all five variables, respondents of web survey are less likely to choose the neutral category ("neither agree nor disagree") than respondents of telephone survey. However, such differences have negligible impact on statistical analysis.

Conclusion: Present results support the view that web surveys are a convenient and reliable mode to collect subjective data, such as those on people's well-being. However, practitioners should be careful when using these data for descriptive purposes, especially when combining data collected with various survey modes. To minimize the risk of biasing the results we advice to gradually shift from telephone to web surveys.

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

Francesco Sarracino†, Cesare F.A. Riillo‡ and Małgorzata Mikucka§

September 8, 2016

Abstract

We test whether web survey mode alters people's evaluations of their well-being compared to telephone survey mode. We use unique, nationally representative data from Luxembourg in which five questions about subjective well-being were asked using web and telephone surveys. Oaxaca decomposition and multinomial logit with Coarsened Exact Matching indicate that the survey mode alters peoples' 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-surveys are convenient and reliable 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

*Małgorzata Mikucka was supported by the ARC grant funded by the Belgian-French Speaking Community within the project "*Family Transformations - Incentives and Norms*" (subvention: 15/19-063).

†Institut national de la statistique et des études économiques du Grand-Duché du Luxembourg (STATEC), 13, rue Erasme, L-2013, Luxembourg, Agence pour la normalisation et l'économie de la connaissance (ANEC), 19-21, boulevard Royal, 2449 Luxembourg City, Luxembourg, and LCSR National Research University Higher School of Economics, Russian Federation. Tel.: +352 247 88469; e-mail: Francesco.Sarracino@statec.etat.lu

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

§Université Catholique de Louvain (Belgium) and LCSR National Research University Higher School of Economics (Russian Federation).

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] Survey data contribute to the evaluations and policy decisions of statistical offices, international organizations, and governments.

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; face-to-face interviews, usually conducted as computer assisted personal interviews (CAPI) or as pen-and-paper interviews (PAPI). For a long time, telephone or 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 offer a solution to many of the shortcomings of other survey modes. 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). Web surveys can be cost-effective, reliable, and user friendly. They may cost less and be conducted more quickly than telephone or face to face surveys. Participants may choose to respond at their convenience, with possible benefits for the quality of the answers. The automatic data entry may reduce errors and speed up data preparation. 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. Yet, despite growing literature, one aspect received comparatively less attention: 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

[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.

popularity of mixed data collection mode (Roster et al., 2004; Schonlau et al., 2003; Nigg et al., 2009; Jäckle et al., 2015; Dillman et al., 2009). 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 alters 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 others, 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. Addi-

[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 and Oswald, 2008; van Reekum et al., 2007). Moreover, SWB measures are correlated with other proxies of SWB (Schwarz and Strack, 1999; Wanous and Hudy, 2001; Schimmack et al., 2010) and – more interestingly – they mirror the judgments about the respondent's happiness provided by acquaintances or clinical experts (Schneider and Schimmack, 2009; Kahneman and Krueger, 2006; Layard, 2005).

[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.

tionally, 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 generate 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 web surveys are a convenient and reliable tool to collect data on subjective well-being. Yet, to minimize the risk of biased conclusions in case of comparisons over time, we recommend to adopt a gradual shift from telephone to web survey.

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 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: population and response effects. Population 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, population effects result from the mismatch between the target population of a study (e.g., the adult population of a country) and the the frame population (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 and 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 population and response effects in the context of web and telephone surveys are discussed in detail below. However, while population effects are well known in the literature, response effects in the use of web surveys have received little attention. We aim to contribute to fill this gap.

2.1 Population 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 (Dillman et al., 2009; Dever et al., 2008; Bethlehem, 2010; Jäckle et al., 2015). Research typically showed that telephone samples over-represented the older respondents (Roster et al., 2004), whereas web surveys under-represented older respondents (Bech and Kristensen, 2009), over-represented younger people (Roster et al., 2004; Flemming and Sonner, 1999), those with higher incomes (Couper, 2000), more geographically mobile (Roster et al., 2004), and men (Flemming and Sonner, 1999). Results for education were mixed. Some studies showed that telephone surveys over-represented adults with a college degree (Berrens et al., 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 (Yeager et al., 2011; Chang and Krosnick, 2009). In some cases, self-selection (Bethlehem, 2010) may substantially affect data quality (Bethlehem, 2015; Hudson et al., 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 and Sonner, 1999). For example, non-probability web sample may be biased toward people engaged and knowledgeable about the survey's topic (Chang and Krosnick, 2009), or towards people using e-mail and Internet more often than others (Yun and Trumbo, 2000).

2.2 Response effects

2.2.1 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 et al., 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 (Yun and Trumbo, 2000; Kiesler and Sproull, 1986). This explains why respondents participating to web surveys provided more accurate answers to questions testing their knowledge (Fricker et al., 2005), and their answers were less affected by random measurement errors (Chang and Krosnick, 2009) than those of telephone survey respondents.

2.2.2 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') (Tourangeau et al., 2004; Fricker et al., 2005; Zhang and Conrad, 2014). However, while some research showed that satisficing is more frequent in web than in telephone surveys (Fricker et al., 2005), other studies demonstrated the opposite (Chang and Krosnick, 2009). Overall, satisficing is more frequent among lower educated respondents (Zhang and Conrad, 2014; Malhotra, 2008), and among respondents who prefer a different survey mode than the mode actually used (Smyth et al., 2014).

2.2.3 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 et al., 2000). Self-administered web surveys provide privacy to the respondent and are more likely to report some socially undesirable opinions or behavior. This may limit the comparability of interviewer- and self-administered modes (Klausch et al., 2013). Studies confirmed that web surveys typically yield less socially desirable answers than telephone surveys (Nigg et al., 2009; Kreuter et al., 2008; Christian et al., 2008; Chang and Krosnick, 2009; Bason, 2000). 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 (Parks et al., 2006; Bason, 2000).

2.2.4 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. Research demonstrated 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; Dillman et al., 2009; Christian et al., 2008; Roster et al., 2004). On the contrary, respondents in telephone surveys tend to give extreme positive answers more often than respondents in web surveys (Christian et al., 2008; Dillman et al., 2009).

However, some studies found no difference in acquiesence between respondents participating in web and telephone surveys (Fricker et al., 2005).

2.2.5 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 (Dillman and Christian, 2005; Christian et al., 2008).

2.3 Relationships between variables

Some studies assessed the quality of answers obtained through web and telephone surveys by investigating relationships among variables (Roster et al., 2004; Yun and Trumbo, 2000; Chang and Krosnick, 2009; Dolan and Kavetsos, 2012). Research compared various survey modes and provided various results. For example, Yun and Trumbo (2000) observed nearly identical R-squares in regression models on data collected by traditional mail vs. electronic channel (web and e-mail surveys combined); they also noted few statistical differences between beta coefficients. However, a comparison of predictors of well-being between face to face and telephone surveys showed differences in significance of coefficients and estimated sizes of effects (Dolan and Kavetsos, 2012).

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 Global Entrepreneurship Monitor (GEM, 2016), a survey of the Statistical Office of Luxembourg (STATEC) 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 data-base 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] homemaker; [f] student; and [g] other not working), and income

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. $N = 5644$
Note: *** $p < 0.01$;

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.

Table 2: Descriptive statistics of variables used in analysis

variable	mean	sd	min	max	obs
Women	0.52	–	0	1	5644
Age (in years)	42.69	12.66	18	64	5644
Age squared / 100	19.82	10.58	3.24	40.96	5644
Elementary education	0.04	–	0	1	5502
Secondary education	0.48	–	0	1	5502
Master craftsman	0.06	–	0	1	5502
Bachelor	0.25	–	0	1	5502
Master	0.15	–	0	1	5502
Full-time employed	0.55	–	0	1	5545
Part-time employed	0.13	–	0	1	5545
Self-employed	0.05	–	0	1	5545
Retired or disabled	0.11	–	0	1	5545
Homemaker	0.05	–	0	1	5545
Student	0.08	–	0	1	5545
Other, not working	0.03	–	0	1	5545
Income (log)	10.82	0.55	9.01	11.42	4363
Year	–	–	2013	2015	5644

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

4 Methods

To test whether the use of web surveys alters 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 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., Helliwell and Barrington-Leigh, 2010; Sarracino, 2013; Bartolini and Sarracino, 2015).

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:

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

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; Oaxaca and Ransom, 1994). Computations are based on the package *nldecompose* which allows to decompose the gap for a categorical variable using STATA (Sinning et al., 2008).

In the vector of explanatory variables (the $\mathbf{X}$ 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.

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 with Coarsened Exact Matching (CEM), a technique previously applied in studies on well-being (e.g., Sidney et al., 2015).

CEM approximates the randomized control trial, i.e. it considers the survey mode as a treatment and it estimates its effect on subjective well-being after matching similar observations from the two samples. In particular, 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. In this way CEM ensures that the "treated" group (the web survey respondents) and the "control" group (the telephone survey respondents) are similar not only on average, but on the whole distribution of observable variables. 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). To make the two groups of respondents comparable, CEM identifies a sub-set of observations for which the covariates are balanced, i.e. the treated and control groups are similar on the distribution of the covariates. This allows to compare the outcomes of the treated and control groups approximating an experimental setting. 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 observations matched using CEM are subsequently analyzed using multinomial logit, which is particularly 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.

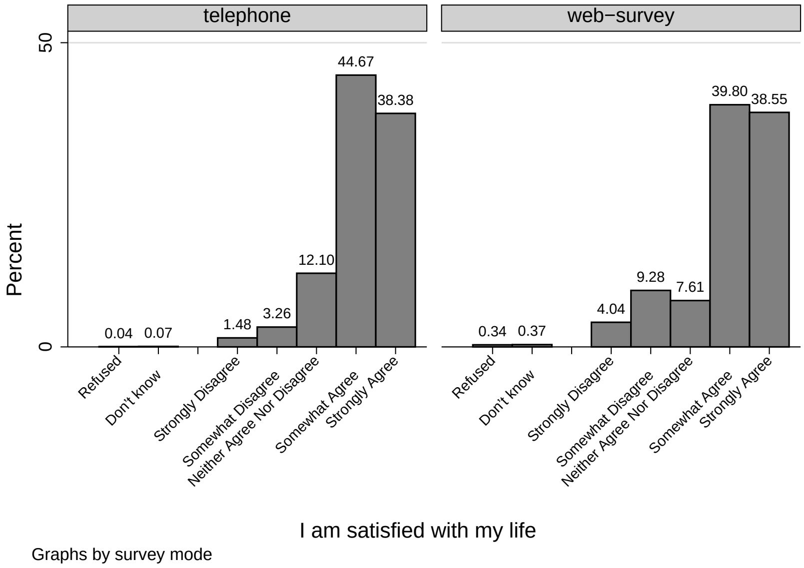
5 Results

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

[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, is violated in our data (Long and Freese, 2006).

are more frequent among respondents of the web survey. The Kolmogorov-Smirnov test for equality of distributions confirms that life satisfaction does not have the same distribution across survey modes. The combined test reports a difference between distribution functions of 0.092 which is significant at 99%.

Figure 1: Distribution of life satisfaction by survey mode.



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 4.07% of respondents *"strongly disagreed"*, and 9.35% *"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%).

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. For instance, it is plausible that respondents of the web survey are on average younger, more educated, and a greater share of them are employed. We report the distribution of life satisfaction, gender, education, employment, age, income, and occupation by survey mode in Appendix A. Table 3 shows the marginal effects of the socio-demographic controls on the probability of participating in the web survey. Variables are incrementally included in the model to test their relative importance.

Table 3: Predictors of participation in the web survey. Marginal effects after 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)
Elementary education			*reference*			
Secondary education			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)
Full-time employed			*reference*			
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)
Income (log)					0.04**	(2.74)
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. * $p < .10$, ** $p < .05$, *** $p < .01$.

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 declines 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

Coefficients in Table 5 report the effect of web survey mode 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 in Table 3, and a consequent lower average subjective well-being for people who used a web survey.

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

	coefficient	z-stat	p-values
Difference:	0.151	3.363	0.001
Decomposition:			
Explained	−0.001	−0.615	0.538
Telephone	−0.246	−5.799	0.000
Web-survey	0.399	19.415	0.000
Observations	4230		

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

Table 5: 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
survey mode (web)	0.027***	0.047***	−0.043**	−0.046	0.014
	(0.007)	(0.011)	(0.014)	(0.024)	(0.023)
Observations	4230	4230	4230	4230	4230

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.
CEM robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

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, $p < 0.05$).

We run the Blinder-Oaxaca decomposition for these two variables (see Table 6). 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.

Table 6: Blinder-Oaxaca decomposition of subjective well-being gap by survey mode. Alternative measures of well-being.

	coefficient	z-stat	p-values
The conditions of my life are excellent:			
Difference:	0.149	4.704	0.000
Decomposition:			
Explained	−0.009	−1.357	0.175
Telephone	−0.332	−3.836	0.000
Web-survey	0.490	5.308	0.000
Observations	4229		
In most ways my life is close to my ideal:			
Difference:	0.062	2.374	0.018
Decomposition:			
Explained	−0.003	−1.199	0.231
Telephone	−0.483	−10.962	0.000
Web-survey	0.548	13.801	0.000
Observations	4220		

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

The counterfactual analysis using multinomial logit and CEM (Table 7) 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.

Table 7: Web survey mode as a predictor of subjective well-being scores. Marginal effects from multinomial logit after CEM.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
The conditions of my life are excellent:					
survey mode (web)	0.030***	0.066***	−0.069***	−0.059*	0.032
	(0.007)	(0.012)	(0.018)	(0.024)	(0.020)
Observations	4227	4227	4227	4227	4227
In most ways my life is close to my ideal:					
survey mode (web)	0.024**	0.045***	−0.140***	0.033	0.038*
	(0.008)	(0.012)	(0.019)	(0.024)	(0.018)
Observations	4221	4221	4221	4221	4221
So far I have obtained the important things I want in life:					
survey mode (web)	0.022**	0.029*	−0.120***	−0.020	0.089***
	(0.007)	(0.014)	(0.018)	(0.024)	(0.022)
Observations	4229	4229	4229	4229	4229
If I could live my life again, I would not change anything:					
survey mode (web)	0.016	0.098***	−0.195***	0.093***	−0.012
	(0.015)	(0.018)	(0.020)	(0.022)	(0.015)
Observations	4193	4193	4193	4193	4193

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.
CEM robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

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 respondents. 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.

6 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 compare the coefficients of a standard happiness regression run on three different samples: the telephone survey respondents, the web survey respondents, and the pooled data-set. Table 8 reports the marginal effects for the fifth category of life satisfaction (*"strongly agree"*) after ordered probit regression.

The results are consistent with previous results from the literature. Ageing decreases the probability do be strongly satisfied with life, and the coefficients document the usual U-shaped pattern. The category "Other, not working", which captures the unemployed, is a significant and negative predictor of being satisfied with one's life. Income level is a significant positive predictor.

The coefficients are fairly consistent across the two samples. Only the coefficients of belonging to the fourth and fifth income quintile differ across the two samples: in the web sample the correlations between income and life satisfaction are stronger than in the telephone sample. (Results not shown, available upon request.) In sum, we find that survey mode affects regression coefficients, but does not substantially change the results. The analysis of the alternative measures of subjective well-being confirm this result (see Appendix C).

Table 8: Predictors of life satisfaction in web and telephone samples. Marginal effects after ordered probit regression.

	telephone survey		web survey		pooled sample	
	dy/dx	*t*	*dy/dx*	*t*	*dy/dx*	*t*
women	0.029	(1.50)	0.016	(0.87)	0.022	(1.63)
age: 18 - 29			*reference*			
age: 30 - 39	−0.102**	(−2.78)	−0.050	(−1.72)	−0.052*	(−2.23)
age: 40 - 47	−0.135***	(−3.79)	−0.064*	(−2.22)	−0.074**	(−3.29)
age: 48 - 55	−0.128***	(−3.52)	−0.083**	(−2.82)	−0.080***	(−3.46)
age: 56 - 64	−0.053	(−1.25)	−0.016	(−0.42)	−0.013	(−0.47)
elementary education			*reference*			
secondary education	−0.019	(−0.53)	0.005	(0.12)	−0.004	(−0.14)
master craftsman	0.049	(0.94)	−0.030	(−0.54)	0.000	(0.01)
bachelor	−0.040	(−1.06)	−0.031	(−0.74)	−0.040	(−1.44)
master	−0.049	(−1.22)	−0.027	(−0.61)	−0.038	(−1.27)
full-time employed			*reference*			
part-time employed	0.003	(0.10)	0.053	(1.81)	0.028	(1.35)
self-employed	0.010	(0.23)	−0.009	(−0.24)	−0.012	(−0.43)
retired or disabled	−0.014	(−0.37)	0.041	(1.02)	0.018	(0.64)
home maker	0.027	(0.52)	0.069	(1.22)	0.056	(1.52)
student	−0.002	(−0.05)	0.041	(0.90)	0.043	(1.27)
not working, other	−0.172**	(−2.89)	−0.236***	(−4.44)	−0.198***	(−4.97)
first income quintile			*reference*			
second income quintile	0.098***	(3.46)	0.152***	(5.64)	0.115***	(5.92)
third income quintile	0.165***	(5.80)	0.208***	(7.37)	0.177***	(8.79)
fourth income quintile	0.148***	(4.77)	0.251***	(8.39)	0.193***	(8.90)
fifth income quintile	0.226***	(7.61)	0.354***	(12.36)	0.284***	(13.82)
year 2013			*reference*			
year 2014	−0.028	(−1.30)	0.003	(0.14)	−0.014	(−0.94)
year 2015	0.020	(0.89)	−0.018	(−0.87)	−0.008	(−0.55)
Observations	2072		2158		4230	

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. * $p < .10$, ** $p < .05$, *** $p < .01$.
The results are robust to using sample weights.
We also run a model interacting all variables with the survey mode. We found significant differences only for the income variables (fourth and fifth quartile)

7 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 provided reasons to question the comparability of the answers to subjective questions administered using different survey modes (Dolan and Kavetsos, 2012). This paper used unique data from the Luxembourgish Global Entrepreneurship Monitor from years 2013–2015 to test whether the web survey mode alters people's evaluations of their subjective well-being compared to telephone survey. 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. Moreover, 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 probit happiness regressions run separately on the samples of 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 and occupation. 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.

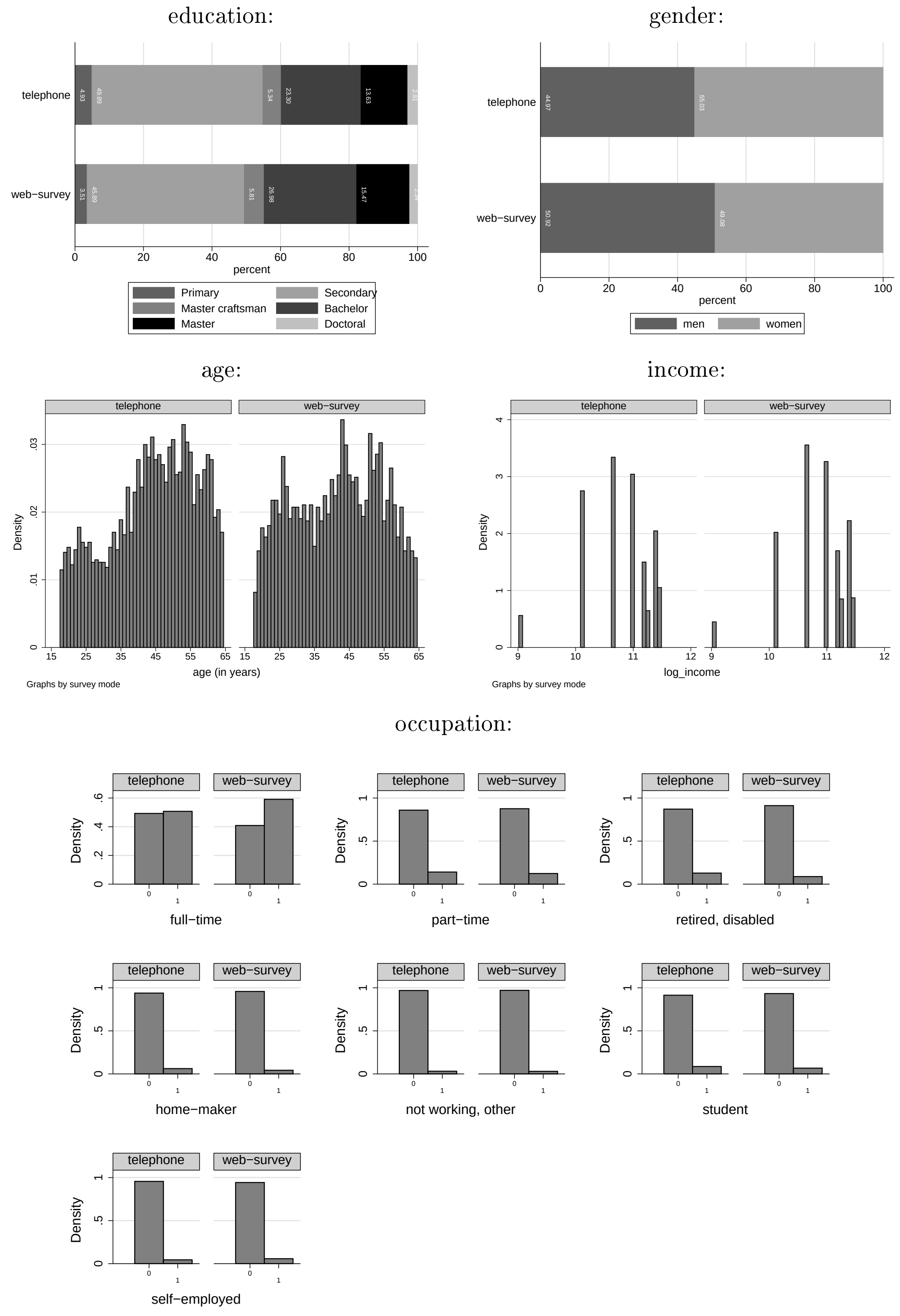
Previous studies suggest three mechanisms to explain these results. First, web respondents tend to provide less socially desirable answers than telephone respondents, which, in many social contexts may result in declaring lower life satisfaction. Second, web respondents tend to choose more often negative 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 respondents. Available studies, however, do not provide any explanation for why web respondents choose the neutral category less often than telephone respondents. 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. On the other hand, it is interesting to 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 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. To minimize the risk of biasing the results it is probably desirable to gradually shift from telephone to web surveys.

A Distribution of socio-demographic variables in the telephone and web samples

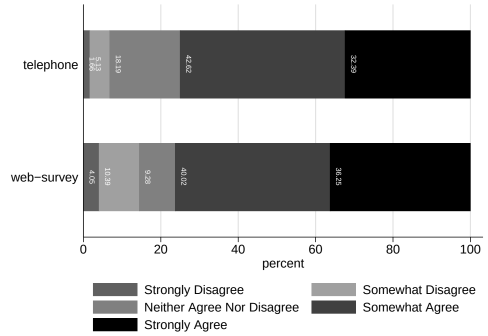
Figure 2: Distribution of socio-demographic variables in the telephone and web samples.



B Distribution of alternative measures of well-being in the telephone and web samples

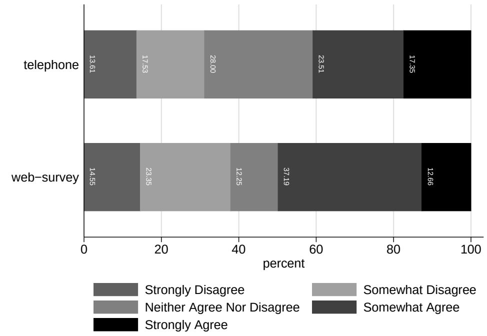
Figure 3: Distribution of alternative measures of well-being in the telephone and web samples

"So far I have obtained the important things I want in life"



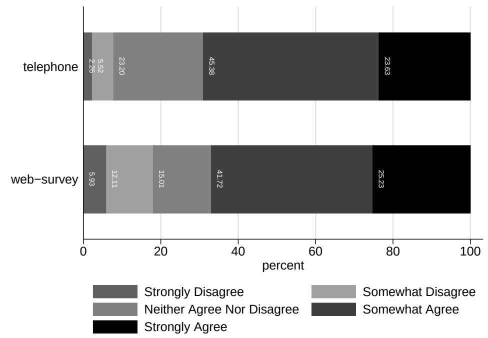
Note: Refusal = 0.27%; Don't know = 0.37%
$\mu_{tel} = 3.98$, $\mu_{web} = 3.97$;
t-test($\mu_{tel} - \mu_{web} = 0$) not significant

"If I could live my life again, I would not change anything"



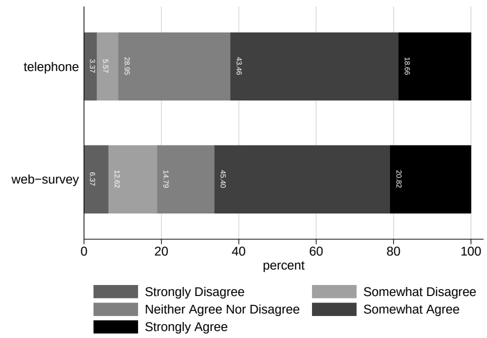
Note: Refusal = 0.41%; Don't know = 2.16%
$\mu_{tel} = 3.13$, $\mu_{web} = 3.11$;
t-test($\mu_{tel} - \mu_{web} = 0$) not significant

"The conditions of my life are excellent"



Note: Refusal = 0.25%; Don't know = 0.25%
$\mu_{tel} = 3.82$, $\mu_{web} = 3.71$;
t-test($\mu_{tel} - \mu_{web} = 0$) significant ($p < 0.01$)

"In most ways my life is close to my ideal"



Note: Refusal = 0.37%; Don't know = 0.53%
$\mu_{tel} = 3.68$, $\mu_{web} = 3.64$;
t-test($\mu_{tel} - \mu_{web} = 0$) significant ($p < 0.05$)

C Effect of the survey mode on the equations of subjective well-being

Table 9: Predictors of strongly agreeing that "The conditions of my life are excellent". Marginal effects after ordered probit regression. Weighted data.

	telephone survey		web survey		pooled sample	
	dy/dx	*t*	*dy/dx*	*t*	*dy/dx*	*t*
Women	−0.00863	(−0.57)	−0.0334*	(−2.19)	−0.0231*	(−2.13)
Age: 18 - 29			*reference*			
Age: 30 - 39	−0.0260	(−0.96)	−0.0110	(−0.47)	−0.00872	(−0.49)
Age: 40 - 47	−0.0454	(−1.70)	−0.0282	(−1.25)	−0.0264	(−1.52)
Age: 48 - 55	−0.00700	(−0.26)	−0.0530*	(−2.30)	−0.0223	(−1.26)
Age: 56 - 64	0.0349	(1.11)	0.0448	(1.52)	0.0448*	(2.10)
Elementary education			*reference*			
Secondary education	−0.0151	(−0.61)	−0.0469	(−1.44)	−0.0253	(−1.27)
Master craftsman	−0.0200	(−0.55)	−0.0816	(−1.95)	−0.0492	(−1.78)
Bachelor	0.0291	(1.08)	−0.0600	(−1.77)	−0.0183	(−0.87)
Master	0.0259	(0.90)	−0.0191	(−0.52)	0.00661	(0.29)
Full-time employed			*reference*			
Part-time employed	−0.0236	(−1.15)	0.0332	(1.41)	0.00765	(0.49)
Self-employed	−0.0206	(−0.63)	−0.0344	(−1.03)	−0.0333	(−1.39)
Retired or disabled	−0.00920	(−0.32)	−0.00114	(−0.04)	0.00178	(0.09)
Home maker	0.0477	(1.29)	0.0109	(0.24)	0.0471	(1.66)
Student	0.0856*	(2.05)	0.144***	(3.89)	0.123***	(4.44)
Not working, other	−0.0765	(−1.62)	−0.154***	(−3.52)	−0.120***	(−3.71)
First income quintile			*reference*			
Second income quintile	0.128***	(5.86)	0.135***	(6.01)	0.121***	(7.77)
Third income quintile	0.175***	(7.97)	0.215***	(9.41)	0.187***	(11.88)
Fourth income quintile	0.205***	(8.95)	0.250***	(10.15)	0.218***	(12.84)
Fifth income quintile	0.288***	(12.65)	0.392***	(16.34)	0.336***	(20.34)
Year 2013			*reference*			
Year 2014	−0.0534**	(−3.22)	−0.00882	(−0.51)	−0.0300*	(−2.50)
Year 2015	−0.0121	(−0.70)	−0.0207	(−1.24)	−0.0226	(−1.86)
Observations	2074		2155		4229	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015.
Note: Estimations for individuals with Internet access and fixed phone line at home.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Table 10: Predictors of strongly agreeing that "In most ways my life is close to my ideal". Marginal effects after ordered probit regression. Weighted data.

	telephone survey		web survey		pooled sample	
	dy/dx	*t*	*dy/dx*	*t*	*dy/dx*	*t*
Women	0.0135	(1.03)	0.00741	(0.54)	0.00838	(0.87)
Age: 18 - 29			*reference*			
Age: 30 - 39	−0.0517*	(−2.09)	0.0207	(0.94)	−0.000292	(−0.02)
Age: 40 - 47	−0.0872***	(−3.60)	−0.0267	(−1.27)	−0.0431**	(−2.73)
Age: 48 - 55	−0.0787**	(−3.14)	−0.0376	(−1.76)	−0.0450**	(−2.77)
Age: 56 - 64	−0.0160	(−0.57)	0.0363	(1.36)	0.0204	(1.06)
Elementary education			*reference*			
Secondary education	−0.0393	(−1.58)	−0.0254	(−0.91)	−0.0293	(−1.57)
Master craftsman	−0.0162	(−0.45)	−0.0289	(−0.80)	−0.0242	(−0.95)
Bachelor	−0.00748	(−0.29)	−0.0465	(−1.60)	−0.0300	(−1.53)
Master	−0.0279	(−1.01)	−0.00588	(−0.19)	−0.0162	(−0.78)
Full-time employed			*reference*			
Part-time employed	−0.00580	(−0.31)	0.0474*	(2.22)	0.0193	(1.35)
Self-employed	−0.00310	(−0.11)	−0.0271	(−0.88)	−0.0231	(−1.07)
Retired or disabled	−0.0228	(−0.93)	0.0285	(0.95)	0.00235	(0.12)
Home maker	0.00853	(0.26)	0.0151	(0.37)	0.0183	(0.73)
Student	−0.0339	(−1.03)	0.0440	(1.25)	0.0172	(0.73)
Not working, other	−0.120**	(−2.92)	−0.165***	(−3.88)	−0.146***	(−4.93)
First income quintile			*reference*			
Second income quintile	0.0727***	(3.79)	0.116***	(5.46)	0.0902***	(6.35)
Third income quintile	0.0914***	(4.82)	0.173***	(7.77)	0.129***	(8.90)
Fourth income quintile	0.111***	(5.57)	0.202***	(8.67)	0.154***	(10.00)
Fifth income quintile	0.167***	(8.36)	0.277***	(11.98)	0.220***	(14.44)
Year 2013			*reference*			
Year 2014	−0.0161	(−1.13)	−0.0309	(−1.90)	−0.0248*	(−2.30)
Year 2015	−0.00408	(−0.27)	−0.0363*	(−2.31)	−0.0223*	(−2.04)
Observations	2071		2149		4220	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015.
Note: Estimations for individuals with Internet access and fixed phone line at home.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Table 11: Predictors of strongly agreeing that "If I could live my life again, I would not change anything". Marginal effects after ordered probit regression. Weighted data.

	telephone survey		web survey		pooled sample	
	dy/dx	*t*	*dy/dx*	*t*	*dy/dx*	*t*
Women	−0.000929	(−0.07)	0.0165	(1.72)	0.00980	(1.26)
Age: 18 - 29			*reference*			
Age: 30 - 39	−0.0463	(−1.93)	−0.0180	(−1.27)	−0.0174	(−1.37)
Age: 40 - 47	−0.105***	(−4.37)	−0.0340*	(−2.35)	−0.0502***	(−3.92)
Age: 48 - 55	−0.110***	(−4.48)	−0.0510***	(−3.37)	−0.0629***	(−4.76)
Age: 56 - 64	−0.0712**	(−2.61)	−0.0138	(−0.75)	−0.0256	(−1.66)
Elementary education			*reference*			
Secondary education	−0.0702**	(−2.97)	0.00423	(0.21)	−0.0318*	(−2.01)
Master craftsman	−0.0207	(−0.59)	−0.000512	(−0.02)	−0.0157	(−0.71)
Bachelor	−0.0366	(−1.49)	0.0124	(0.60)	−0.0157	(−0.96)
Master	−0.0321	(−1.21)	0.0325	(1.47)	−0.000445	(−0.03)
Full-time employed			*reference*			
Part-time employed	−0.00951	(−0.52)	0.0188	(1.28)	0.00588	(0.50)
Self-employed	0.0420	(1.30)	0.00455	(0.23)	0.0152	(0.84)
Retired or disabled	−0.0293	(−1.26)	0.0393	(1.88)	0.00502	(0.32)
Home maker	0.0287	(0.91)	−0.00190	(−0.07)	0.0184	(0.87)
Student	−0.0408	(−1.30)	−0.0127	(−0.57)	−0.0115	(−0.63)
Not working, other	−0.00314	(−0.08)	−0.0584	(−1.90)	−0.0326	(−1.33)
First income quintile			*reference*			
Second income quintile	0.0130	(0.71)	0.0448**	(3.12)	0.0291*	(2.51)
Third income quintile	0.0495**	(2.67)	0.0714***	(4.74)	0.0603***	(5.07)
Fourth income quintile	0.0580**	(2.92)	0.0818***	(5.37)	0.0703***	(5.71)
Fifth income quintile	0.0831***	(4.36)	0.112***	(6.99)	0.0996***	(8.04)
Year 2013			*reference*			
Year 2014	−0.00112	(−0.08)	0.00620	(0.55)	0.00256	(0.29)
Year 2015	0.00561	(0.38)	0.000368	(0.03)	0.00122	(0.14)
Observations	2154		2323		4477	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015.
Note: Estimations for individuals with Internet access and fixed phone line at home.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Table 12: Predictors of strongly agreeing that "So far I have obtained the important things I want in life". Marginal effects after ordered probit regression. Weighted data.

	telephone survey		web survey		pooled sample	
	dy/dx	t	dy/dx	t	dy/dx	t
Women	0.0425*	(2.40)	0.0433*	(2.44)	0.0414**	(3.27)
Age: 18 - 29			*reference*			
Age: 30 - 39	0.0296	(0.88)	−0.0244	(−0.94)	−0.00650	(−0.32)
Age: 40 - 47	0.0388	(1.17)	−0.0184	(−0.70)	0.000241	(0.01)
Age: 48 - 55	0.0425	(1.26)	−0.0208	(−0.79)	0.000867	(0.04)
Age: 56 - 64	0.101**	(2.68)	0.00910	(0.27)	0.0428	(1.73)
Elementary education			*reference*			
Secondary education	−0.0190	(−0.60)	0.0733*	(1.98)	0.0259	(1.07)
Master craftsman	0.0744	(1.62)	0.0474	(1.00)	0.0494	(1.50)
Bachelor	−0.00545	(−0.16)	0.0729	(1.88)	0.0276	(1.07)
Master	−0.0130	(−0.35)	0.0532	(1.29)	0.0160	(0.58)
Full-time employed			*reference*			
Part-time employed	0.0199	(0.76)	0.0296	(1.05)	0.0222	(1.14)
Self-employed	0.0367	(0.96)	0.0116	(0.32)	0.0136	(0.51)
Retired or disabled	−0.0259	(−0.79)	0.0447	(1.17)	0.00958	(0.38)
Home maker	0.0895*	(2.13)	0.0324	(0.60)	0.0723*	(2.19)
Student	0.0611	(1.22)	−0.0143	(−0.32)	0.0152	(0.46)
Not working, other	−0.190***	(−3.79)	−0.108*	(−2.21)	−0.143***	(−4.10)
First income quintile			*reference*			
Second quintile	0.0613*	(2.42)	0.135***	(5.48)	0.0962***	(5.42)
Third quintile	0.0775**	(3.12)	0.191***	(7.37)	0.134***	(7.39)
Fourth quintile	0.171***	(6.23)	0.253***	(9.00)	0.209***	(10.49)
Fifth quintile	0.173***	(6.44)	0.336***	(12.26)	0.255***	(13.19)
Year 2013			*reference*			
Year 2014	−0.0408*	(−2.11)	−0.00387	(−0.18)	−0.0212	(−1.48)
Year 2015	0.0173	(0.84)	−0.0120	(−0.60)	−0.000931	(−0.06)
Observations	2168		2388		4556	

Source: Global Entrepreneurship Monitor, Luxembourg, 2013-2015.
Note: Estimations for individuals with Internet access and fixed phone line at home.
* $p < .10$, ** $p < .05$, *** $p < .01$.

References

Bartolini, S. and 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. *American Association of Public Opinion Research*, pages 1074–1080.

Bech, M. and 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., and 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.

Blanchflower, D. and Oswald, A. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine*, 66:pp. 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., and Gates, R. (2007). A comparison of reliability between telephone and web-based surveys. *Journal of Business Research*, 60(7):758–764.

Chang, L. and 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., and 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., and Valliant, R. (2008). Internet surveys: Can statistical adjustments eliminate coverage bias. *Survey Research Methods*, 2(2):47–62.

Dillman, D. A. and 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., and 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. and Kavetsos, G. (2012). Happy talk: Mode of administration effects on subjective well-being. *Journal of Happiness Studies*, pages 1–19.

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

Fricker, S., Galesic, M., Tourangeau, R., and 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. and 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., and 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., and 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., and Porro, G. (2012). Causal inference without balance checking: Coarsened Exact Matching. *Political Analysis*, 20(1):1–24.

Jäckle, A., Lynn, P., and 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., and 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. and Krueger, A. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20:3–24.

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

Klausch, T., Hox, J. J., and 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., and 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.* The Penguin Press, New York.

Long, J. S. and Freese, J. (2006). *Regression models for categorical dependent variables using Stata.* 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., Horwath, C. C., and 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. and 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., and 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.

Roster, C. A., Rogers, R. D., Albaum, G., and 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., and 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. and 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., and 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. and Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. D. and Schwarz, N., editors, *Well-being: the foundations of hedonist psychology*. Russell Sage Foundation, New York.

Sidney, J., Coberley, C., Pope, J., and 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., et al. (2008). The Blinder-Oaxaca decomposition for nonlinear regression models. *The Stata Journal*, 8(4):480–492.

Smyth, J., Olson, K., and 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., and 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., Schaefer, H., Alexander, A., and 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. and Hudy, M. (2001). Single-item reliability: a replication and extension. *Organizational Research Methods*, 4:pp. 361 – 375.

We Are Social (2016). Digital in 2016. Technical report, We Are Social. available at: `http://wearesocial.com/uk/special-reports/digital-in-2016`.

Yeager, D. S., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., and 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. and Trumbo, C. W. (2000). Comparative response to a survey executed by post, e-mail, & web form. *Journal of Computer-Mediated Communication*, 6(1):0.

Zhang, C. and 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.