Volume 32, Issue 4

I would really love to participate in your survey! Bias problems in the measurement of well-being

Adrian Chadi
The Institute for Labour Law and Industrial Relations in the European Union (IAAEU)

Abstract

This paper argues that satisfaction data from surveys are biased by varying participant attitudes toward the interview itself. In this manner, interviewees in a German panel study report lower life satisfaction when there is evidence of transient influences like aversion. The empirical findings suggest that researchers of well-being should consider interview-specific factors in order to avoid drawing incorrect conclusions.

1. Introduction

Happiness research relies strongly on the validity of information reported in surveys. However, transient influences such as the interviewee's attitude toward the interview itself may cause considerable differences in reported well-being. While interview-specific factors account for a significant portion of the variance in satisfaction measured (Schwarz and Strack 1999, and Krueger and Schkade 2008), most researchers implicitly assume that there is no systematic bias in the collected well-being data. Indeed, in the case of random variation, sufficiently large samples do ensure that these effects are averaged out.

Transient mood effects constitute a problem for empirical research when certain groups of people are affected differently than others. Imagine, for example, Person 1, a prime-aged individual with a stressful full-time job who is confronted with the same panel survey year after year, but who is assumedly as happy as Person 2, a pensioner with much more time and possibly even some intrinsic motivation to participate in the survey. Obviously, these different circumstances could easily result in well-being data that are biased because of situational factors.

This study uses German panel data to examine the potential effects of varying participant attitudes toward the survey interview itself in order to determine (a) whether there is a significant impact on measured levels of well-being and (b) whether the consideration of interview-specific factors may even alter established conclusions about life satisfaction and its determinants.

2. Data and Empirical Strategy

The dataset used for this empirical analysis covers the period from 1994 to 2009 and is from the German Socio-Economic Panel Study (SOEP), a large representative survey of households (see Wagner *et al.* 2007). After excluding all missing values concerning the relevant variables used in the analysis, the unbalanced sample contains 279,132 observations.

In the SOEP questionnaire, respondents evaluate their general life satisfaction on a scale ranging from 0 to 10, with higher scores indicating greater well-being. One way to deal with the ordinal nature of the data is to transform these scores into a binary variable, which takes the value of one above a fixed threshold and zero below it. A binary distinction between only two states (happiness and unhappiness) allows the application of logit and probit methods. On further reflection, however, ordered logit and ordered probit methods seem superior, since, by use of the entire scale, the loss of a huge amount of information can be avoided. Eventually, if the empirical analysis is conducted on the basis of longitudinal panel data, it makes sense to exploit the fact that the same persons are observed more than once and to consider individual-specific effects. Following this line of thought, researchers of situation-specific response artefacts in satisfaction data, such as Taylor (2006) or Conti and Pudney (2011), implement random effects ordered probit (REOP) models.

In recent years, however, models that take account of fixed individual effects have become very popular in the happiness literature (e.g., Knabe and Raetzel 2010). Because of the strong correlation between satisfaction levels and stable personality traits (e.g., Lykken and Tellegen 1996), there is good reason to consider the influence of such time-invariant predispositions. Hence, in line with the methodological implications of Ferrer-i-Carbonell and Frijters (2004), the present study also conducts ordinary least squares (OLS) regressions, allowing for the consideration of time and individual fixed effects. Note that this implies that the linear effect of age and all time-invariant factors are omitted from the specifications.

The empirical model in this investigation is

$$LS_{it} = a\,X_{it} + b\,\mathrm{IF}_{it} + v_i + w_t + e_{it} \tag{1}$$

where X_{it} is a set of observed determinants of life satisfaction (LS_{it}) and IF_{it} are the interview-specific factors (with a and b as the corresponding coefficients), while w_t denotes time-varying circumstances in a specific year and e_{it} is the error term. Note that in contrast to the fixed-effects OLS method, the REOP method treats the individual-specific component v_i as random and estimates latent probabilities of the observed satisfaction scores.[1]

In order to examine the role of participant attitudes toward the SOEP interview, this study commences with an original idea, based on the intuitive argument that less motivated interviewees are more likely to skip questions, especially at the end. In fact, reported well-being is significantly lower for those skipping the question just prior to the final one on life satisfaction, which can be regarded as evidence of transient influences such as indifference when this also applies to the multivariate analysis.[2] Note that in most questionnaires, this penultimate question is in fact a whole battery of sub-questions either about changes in the family situation or about given payments (to relatives and others), which unenthusiastic participants are certainly like to skip.

In the form of the interview mode and the date, the SOEP offers further useful but hidden information about people's potential aversion to being interviewed. With respect to the former, the oral interview is the first-choice option for data collectors and is thus conducted whenever possible, while in the event of refusal, other options are offered (Haisken-DeNew and Frick 2005). This configuration implies that motivated individuals are likely to give personal interviews, while self-completed questionnaires are at least occasionally associated with adverse attitudes, so that reported well-being levels among these participants may be biased downward. Similarly, this selection effect is also likely to be reflected in the date of the interview, with reluctant potential participants refusing the originally proposed appointment. In addition to their generally negative stance toward the interview itself, less motivated survey participants might specifically choose periods during which they prefer to conduct unpleasant activities (e.g., on a rainy Sunday), which would again affect the reported well-being levels.

Apart from these selection effects, situation-specific factors, such as the day of the interview (Taylor 2006), may also affect measured satisfaction levels more directly. Conti and Pudney (2011) find that face-to-face interviews generally produce more positive satisfaction responses, while, in the more private setting of a self-written questionnaire, participants are less reluctant to reveal negative emotions. According to this argument, respondents may have even greater reservations regarding reporting unhappiness in the Computer-Assisted Personal Interviewing (CAPI) scenario, in contrast to the traditional paper-and-pencil process (Schraepler 2007).

3. Results and Discussion

Tables 1 and 2 present the outcomes of the empirical analysis to investigate the role of survey-specific factors in the reporting of life satisfaction. While the second table also presents the OLS outcomes for a large set of variables that potentially affect satisfaction responses, the first table gives REOP results only concerning the examined interview factors. In both cases, Specification 1 indicates a strongly significant "prior-question" effect, so that skipping the second-to-last question in the SOEP is clearly associated with lower levels of reported well-being. Robustness checks confirm this finding to be unrelated to the question content. Therefore, transient influences such as aversion or even stress seem to be valid explanations of this remarkable finding.

---

[1] See Frechette (2001) for more on this and on his STATA method, which has been used in this study.

[2] As the identification of non-responses to the second-to-last question is rather difficult in the 1993 questionnaire, the data sample used here starts with the 1994 wave.

Table 1: Life Satisfaction and Random Effects Ordered Probit Regressions

Specification:	(1)	(2)	(3)	(4)
Prior question skipped	-0.064***	-0.063***	-0.063***	-0.044**
	(0.021)	(0.021)	(0.021)	(0.021)
Day of Interview				
Tuesday		-0.008	-0.008	-0.008
		(0.007)	(0.007)	(0.007)
Wednesday		-0.004	-0.004	-0.006
		(0.007)	(0.007)	(0.007)
Thursday		-0.004	-0.005	-0.005
		(0.007)	(0.007)	(0.007)
Friday		-0.017**	-0.017**	-0.013*
		(0.008)	(0.008)	(0.008)
Saturday		-0.036***	-0.035***	-0.027***
		(0.008)	(0.008)	(0.008)
Sunday		-0.048***	-0.048***	-0.034***
		(0.010)	(0.010)	(0.010)
Month of Interview				
February			-0.021***	-0.014*
			(0.007)	(0.007)
March			-0.024***	-0.014*
			(0.008)	(0.008)
April			-0.008	0.001
			(0.009)	(0.09)
May			-0.001	0.006
			(0.011)	(0.011)
June			0.009	0.013
			(0.012)	(0.012)
July			-0.009	-0.001
			(0.014)	(0.014)
August			0.005	0.009
			(0.017)	(0.017)
September			-0.005	0.005
			(0.022)	(0.022)
October			0.045	0.061*
			(0.032)	(0.032)
November			0.006	0.037
			(0.115)	(0.115)
December			-0.066	-0.106
			(0.094)	(0.094)
Survey Instrument				
Oral interview				0.102***
				(0.011)
CAPI				0.163***
				(0.012)
Written questionnaire, with interviewer				-0.032**
				(0.015)
Written questionnaire, no interviewer				-0.120***
				(0.011)
Written, by mail				-0.105***
				(0.015)
Log-likelihood	-452329	-452307	-452294	-451679

Notes: *(**/***) denotes significance at 10% (5%/1%) level. Standard errors are in parentheses. Covariates include variables for health, employment and family status as well as owner of dwelling, home resident, no children in household, person needing care, log household income p.c., sex, age (also squared), year and state. Reference categories are Monday, January and other survey instruments.

Table 2: Life Satisfaction and Fixed Effects OLS Regressions

Specification:	(0)	(1)	(2)	(3)	(4)
Health status					
Good	-0.348***	-0.348***	-0.348***	-0.348***	-0.335***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Satisfactory	-0.755***	-0.755***	-0.754***	-0.754***	-0.734***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Not so good	-1.311***	-1.311***	-1.311***	-1.311***	-1.291***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Bad	-2.354***	-2.355***	-2.354***	-2.354***	-2.333***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Employment status					
Full-time employed	-0.021	-0.021	-0.022	-0.021	-0.020
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Regular part-time	-0.081***	-0.081***	-0.082***	-0.081***	-0.080***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Irregular part-time	-0.104***	-0.104***	-0.105***	-0.104***	-0.098***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Sheltered workshop	-0.339	-0.340	-0.341	-0.341	-0.350
	(0.241)	(0.241)	(0.241)	(0.241)	(0.242)
Not employed	-0.073***	-0.073***	-0.074***	-0.073***	-0.077***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Registered unemployed	-0.524***	-0.524***	-0.524***	-0.524***	-0.524***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Family status					
Married	0.190***	0.190***	0.190***	0.190***	0.191***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Married, but separated	-0.191***	-0.191***	-0.191***	-0.191***	-0.192***
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
Spouse in native country	0.353**	0.353**	0.355**	0.353**	0.383**
	(0.156)	(0.157)	(0.157)	(0.157)	(0.157)
Divorced	0.105***	0.105***	0.105***	0.105***	0.100***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Widowed	-0.244***	-0.244***	-0.244***	-0.243***	-0.251***
	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
Further Controls					
Owner of dwelling	0.071***	0.072***	0.072***	0.072***	0.072***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Home resident	-0.380***	-0.380***	-0.379***	-0.377***	-0.377***
	(0.131)	(0.131)	(0.131)	(0.131)	(0.132)
No children in the household	-0.072***	-0.071***	-0.071***	-0.071***	-0.067***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Person needing care in the household	-0.427***	-0.427***	-0.427***	-0.427***	-0.424***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Log household income per capita	0.268***	0.268***	0.269***	0.269***	0.271***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Age squared / 10000	0.684**	0.681**	0.670*	0.655*	0.389
	(0.342)	(0.342)	(0.342)	(0.343)	(0.343)
Prior question skipped		-0.066***	-0.066**	-0.065**	-0.051**
		(0.026)	(0.026)	(0.026)	(0.026)
Day of Interview					
Tuesday			-0.012	-0.012	-0.012
			(0.008)	(0.008)	(0.008)
Wednesday			-0.004	-0.004	-0.005
			(0.008)	(0.008)	(0.008)

(continued from previous page)

Thursday			-0.008	-0.008	-0.008
			(0.009)	(0.009)	(0.009)
Friday			-0.021**	-0.021**	-0.018**
			(0.009)	(0.009)	(0.009)
Saturday			-0.029***	-0.029***	-0.022**
			(0.009)	(0.009)	(0.009)
Sunday			-0.038***	-0.038***	-0.027**
			(0.012)	(0.012)	(0.012)
Month of Interview					
February				-0.024***	-0.018**
				(0.009)	(0.009)
March				-0.038***	-0.030***
				(0.010)	(0.010)
April				-0.030**	-0.021*
				(0.012)	(0.012)
May				-0.025*	-0.017
				(0.014)	(0.014)
June				-0.017	-0.012
				(0.015)	(0.015)
July				-0.035**	-0.024
				(0.018)	(0.018)
August				-0.017	-0.009
				(0.022)	(0.021)
September				-0.012	-0.000
				(0.027)	(0.027)
October				-0.002	0.014
				(0.039)	(0.039)
November				0.025	0.059
				(0.138)	(0.138)
December				-0.048	-0.075
				(0.134)	(0.134)
Survey Instrument					
Oral interview					0.110***
					(0.016)
CAPI					0.188***
					(0.018)
Written questionnaire, with interviewer					-0.037*
					(0.020)
Written questionnaire, no interviewer					-0.137***
					(0.016)
Written, by mail					-0.108***
					(0.026)
Adjusted R^2	0.1029	0.1030	0.1030	0.1030	0.1061

Notes: *(**/***) denotes significance at 10% (5%/1%) level. Robust standard errors are in parentheses. Additional covariates are year and state dummy variables. The omitted reference groups are people in very good health, those in vocational training and single persons. Reference categories regarding the interview-specific factors are Monday, January and other survey instruments.

Specification 2 suggests that there are considerable day-of-the-week effects, so that participants report significantly different levels of well-being depending on the day of the interview. At first glance, it is counter-intuitive that people report being less happy on Fridays and Saturdays, and particularly on Sundays. Referring to the present argument, however, this

makes sense when differences in reported well-being are interpreted as a consequence of temporary mood effects. In particular, the unenthusiastic survey participants may report downward-biased satisfaction levels when they decide, for whatever reason, to do the unpleasant job on the weekend, during their precious leisure time.

As mentioned in Section 2, a first refusal to do the interview can lead not only to a different interview mode, but also to a different interview date. The selection bias, resulting from underreported well-being levels, is therefore expected to be reflected in several different factors. In comparison to the previous specifications, the results for the final Specification 4 substantiate this argument, as the significant coefficients indicating transient influences become smaller.

In fact, the relevance of the interview mode and the consequences associated with whether an interviewer is present or not are impressive. In addition, happiness reports become even more positive when a computer is used (in the case of CAPI), which underlines the privacy argument, according to which the circumstances of the interview determine people's reluctance or willingness to provide sensitive information. However, in order to explain the enormous interview-mode effects, attitudes toward the interview itself certainly have to be considered as well.

After discussing the relevance of interview-specific factors, the second research question can also be addressed by the results in Table 2. So far, most happiness researchers have assumed that transient influences induce only random variation in the dependent variable. By and large, the empirical results are indeed robust with respect to the inclusion of interview-specific controls. However, the following example may demonstrate the potential consequences of not considering the specific circumstances of the data collection.
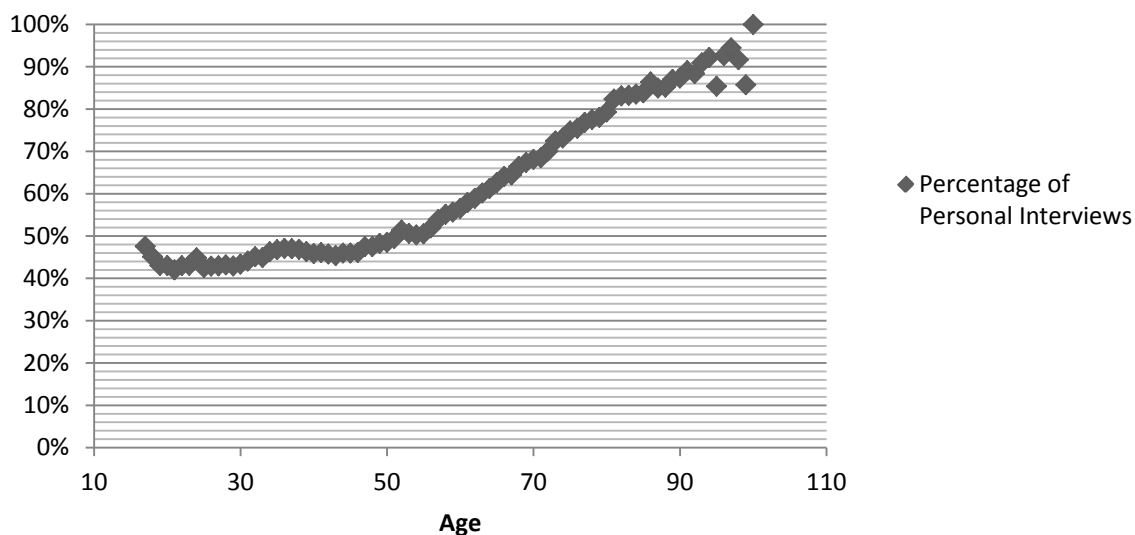

Figure 1: Average percentages of personal interviews by age

While some people are simply uninterested in giving survey interviews, others may have a more positive attitude and possibly even delight in being questioned about their lives. In the SOEP, the latter group is more likely to be interviewed personally, which already implies a bias effect related to the lower level of privacy. On the other hand, especially younger people often initially refuse to participate, so that the percentage of interviews conducted in writing is

much higher in that group. Therefore, the interview mode, just like the date of the interview, captures information about the various attitudes toward being interviewed. The unambiguous relationship between age and the percentage of personal interviews (oral interviews, including those with computer assistance), shown in Figure 1, indicates a serious danger of obtaining biased results.

Indeed, while there is no significant relationship between age-squared and life satisfaction in the appropriately specified model (Specification 4 in Table 2), the outcomes for a model without any variables for transient factors (Specification 0 in Table 2) incorrectly suggests a significantly positive effect resulting merely from being older, which is in fact a common finding in the empirical literature on well-being. Although the relationship between age and happiness is certainly a complex issue that has been, for a reason, discussed intensely of late (e.g., Glenn 2009), this finding could make a very important contribution to this debate.

4. Conclusion

This empirical study confirms the concern that satisfaction data are biased by the individually specific situation in which the survey interview takes place. Transient factors, such as varying attitudes to being interviewed, play a significant role in the emergence of subjective data, which are all too often used unquestioningly by empirical researchers. The significant impact of interview-specific factors on the determination of well-being is reason enough to consider these aspects in empirical analyses whenever possible.

The example of the age-happiness relationship hints at the potential consequences of empirical researchers ignoring the fact that the commonly used determinants of happiness explain only a small portion of the variance in subjective well-being data. Moreover, researchers should question whether prime-agers are really unhappy with their lives or simply unhappy about spending their time on interviews, whereas older people may report being happier, but perhaps only because they actually like to participate in survey interviews.

References

Conti, G., and S. Pudney (2011) "Survey design and the analysis of satisfaction" *Review of Economics and Statistics* **93**, 1087-1093.

Ferrer-i-Carbonell, A., and P. Frijters (2004) "How important is methodology for the estimates of the determinants of happiness?" *Economic Journal* **114**, 641-659.

Frechette, G. (2001) "Random-effects ordered probit" *Stata Technical Bulletin* **59**, 23–27.

Glenn, N. (2009) "Is the apparent U-shape of well-being over the life course a results of inappropriate use of control variables? A commentary on Blanchflower and Oswald" *Social Science & Medicine* **69**, 481-485.

Haisken-DeNew, J., and J. Frick (2005) "Desktop companion to the German Socio-Economic Panel (SOEP)" German Institute for Economic Research: Berlin.

Knabe, A., and S. Raetzel (2010) "Better an insecure job than no job at all? Unemployment, job insecurity and subjective wellbeing" *Economics Bulletin* **30**, 2486-2494.

Krueger, A., and D. Schkade (2008) "The reliability of subjective well-being measures" *Journal of Public Economics* **92**, 1833-1845.

Lykken, D., and A. Tellegen (1996) "Happiness is a stochastic phenomenon" *Psychological Science* **7**, 186-189.

Schraepler , J. (2007) "A study of mode-effects of a change from PAPI to CAPI" *Journal of Applied Social Science Studies* **127**, 113-125.

Schwarz, N., and F. Strack (1999) "Reports of subjective well-being: judgment processes and their methodological implications" in *Well Being: The Foundations of Hedonic Psychology* by Kahneman, D., Diener, E., and N. Schwarz (Eds.), Russell Sage: New York, 61-84.

Taylor, M. (2006) "Tell me why I don't like Mondays: investigating day of the week effects on job satisfaction and psychological well-being" *Journal of the Royal Statistical Society Series A* **169**, 127-142.

Wagner, G., Frick, J., and J. Schupp (2007) "The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements" *Journal of Applied Social Science Studies (Schmollers Jahrbuch)* **127**, 139-169.