Personality Has Minor Effects on Panel Attrition

David Richter, John L. Körtner and Denise Saßenroth
SOEPpapers on Multidisciplinary Panel Data Research
at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at http://www.diw.de/soeppapers

Editors:
Jürgen Schupp (Sociology)
Gert G. Wagner (Social Sciences, Vice Dean DIW Graduate Center)

Conchita D’Ambrosio (Public Economics)
Denis Gerstorf (Psychology, DIW Research Director)
Elke Holst (Gender Studies, DIW Research Director)
Frauke Kreuter (Survey Methodology, DIW Research Professor)
Martin Kroh (Political Science and Survey Methodology)
Frieder R. Lang (Psychology, DIW Research Professor)
Henning Lohmann (Sociology, DIW Research Professor)
Jörg-Peter Schräpler (Survey Methodology, DIW Research Professor)
Thomas Siedler (Empirical Economics)
C. Katharina Spieß (Empirical Economics and Educational Science)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | soeppapers@diw.de
Abstract

In light of the recent interest in using longitudinal panel data to study personality development, it is important to know if personality traits are related to panel attrition. We analyse the effects of personality on panel drop-out separately for an ‘older’ subsample (started in 1984), a relatively ‘young’ subsample (started in 2000), and a ‘new’ subsample (started in 2009) of the German Socio-Economic Panel (SOEP) study. We found that openness slightly decreases the probability of panel drop-out in all three samples. For the ‘older’ subsample only, we found a small negative effect of agreeableness on panel drop-out. We control for age, sex, education, migration background, and the number of inhabitants in the region of the respondents.

(word count: 116)
Personality has minor effects on panel attrition

Personality dispositions like those captured by the Big Five model (John, Naumann, & Soto, 2008) are considered to be relatively stable patterns that distinguish the individual from others. Recent research on personality has used longitudinal data to investigate the development of the core personality dispositions at specific ages and to study changes in personality traits over the adult life span (e.g., Lucas & Donnellan, 2011; Specht, Schmukle, & Egloff, 2011). It is therefore important to know if, and to what extent, personality traits are related to panel attrition.

If the reasons for panel attrition are related to the investigated variables, the sample could be biased (Groves et al., 2009). A relationship between personality traits and attrition (e.g., when respondents with lower values for a certain trait have a higher probability of dropping out than respondents with higher values for the same trait) could limit the quality and accuracy of the data. Recent research using the Big Five taxonomy has provided valuable initial insights into the effects of personality traits on panel attrition.

Roberts et al. (2006) conducted a meta-analysis of 92 longitudinal studies to investigate the development of personality across the life course, but found no systematic relationship between attrition ($M = 44\%$, range 0\% to 93\%) and mean-level change in personality traits. Salthouse (2013), in contrast, found higher levels of agreeableness ($d = .11$) and openness ($d = .11$) among respondents ($N = 2,082$) returning for the second round of an ongoing panel survey on cognitive functioning than among non-returning respondents ($N = 1,698$). In addition, older respondents with higher levels of extraversion had a higher probability of returning, indicating interaction effects of age and panel attrition.

Lugtig (2014) used a latent class framework to capture different patterns of panel commitment in a Dutch online panel, the Longitudinal Internet Studies for the Social Sciences (LISS; $N = 8,148$). He was able to empirically separate nine different groups of respondents
and to follow a different and distinct process of attrition for each one. Loyal stayers (37% of the sample), defined as respondents who participated in almost all waves of the panel, were found to be more conscientious ($d = .06$), less extraverted ($d = .15$), and less agreeable ($d = .07$) than fast attritors (19% of the sample) who dropped out immediately after the start of the panel. Also with the LISS panel, Saßenroth (2013) revealed negative effects of neuroticism ($d = .10$) and openness ($d = .09$) on the duration of panel membership.

In summary, research on the effects of personality on panel attrition is sparse, the results are inconsistent, and the effects identified are rather small. The present study attempts to fill this gap by examining the effects of personality on panel drop-out in a large sample of 11,436 German adults surveyed annually in the Socio-Economic Panel Study (SOEP).

The SOEP is a frequently used source of longitudinal data for the study of personality development (e.g., Lucas & Donnellan, 2011; Specht, Schmukle, & Egloff, 2011; but see Cobb-Clark & Schurer, 2012, for a study using the Australian HILDA panel). Our study provides insight into the reliability of studies on personality development using SOEP data.

We use Cox proportional hazards regressions (Cox, 1972) to estimate the effects of personality traits on panel drop-out. Based on past studies, we refrain from stating specific hypotheses as previous results are very mixed. We include several control variables that have been shown to influence either personality traits or panel survival in previous studies.

Studies on personality development over the adult life span have shown that personality traits change with age (e.g., Lucas & Donnellan, 2011; Specht et al., 2011). Thus, we control for age, as well as for age$^2$ and age$^3$ to test for non-linear effects. Variables that have shown effects on panel survival in previous waves of the SOEP as well as in other longitudinal studies are sex, migration background, education, and geographical region of the respondents (e.g., Groves & Couper, 1998; Groves & Lyberg 1988; Kroh, 2013).
Material and Methods

Sample

To test the effect of personality on panel attrition we use data from the Socio-Economic Panel Study (SOEP). The SOEP is an ongoing longitudinal study of households in Germany. Details on sampling strategies, response rates, attrition, and representativeness of the sample can be found in Wagner, Frick, and Schupp (2007) and on the SOEP website (http://www.diw.de/gsoep). Households are selected using a multistage probability design, and all adult members of selected households (i.e., those older than age 16) are asked to participate. Since 1984, respondents have been surveyed yearly.

The multi-cohort nature of the SOEP – new respondents are recruited into the SOEP on a regular basis – enabled us to utilize three subsamples of the SOEP, each selected to be representative of the population of Germany, using a multistage probability design for our analyses: a ‘new’ sample surveyed first in 2009 with personality measured on the same occasion (SOEP Sample I, \( N = 2,400 \)), a ‘younger’ sample with short duration of panel membership surveyed first in 2000 with personality measured in the year 2005 after five years of panel membership (SOEP Sample F, \( N = 5,826 \)), and an ‘older’ sample with a long duration of panel membership, consisting of those surveyed first in 1984 with personality measured in the year 2005 after 21 years of panel membership (SOEP Samples A & B, \( N = 3,210 \)).

Analysing three different subsamples separately allows us to test whether personality traits influence panel attrition differently in newly started surveys than in longer running panel surveys. We refrain from the alternative approach of analysing all samples together while using duration of panel membership as a control variable, as a) time in study does not overlap for the three subsamples and therefore time and cohort effects (i.e., initial differences in the
composition of the subsamples) cannot be disentangled and b) time in study might have non-linear relationship with panel attrition.

We excluded all respondents that died or moved abroad between 2005 and 2012. Furthermore, we excluded all respondents that lived in their respective households in 1984 (‘older’ sample) or 2000 (‘younger’ sample) as children, even if they provided measures of their personality in 2005, because these respondents (‘older’ sample: \(N = 1,027\); ‘younger’ sample: \(N = 537\)) were not able to decide whether or not to join the panel until their first individual interview at the age of 17. Finally, respondents with missing data in the control variables described below were excluded as well (\(N = 9\)).

Material

In 2005 and 2009, the Big Five personality traits (BFI; John, Donahue, & Kentle, 1991) were measured in the SOEP using a short form with 15 items (Lang, John, Lüdtke, Schupp, & Wagner, 2011). Three items were used to measure each dimension. Participants responded to these items on a 7-point scale that ranged from ‘1’ (does not apply) to ‘7’ (does apply). The inter-item correlations are reported by Lucas & Donnellan (2011).

Control variables were measured with the yearly SOEP questionnaires. Age (‘new’ sample: \(M = 49.81\) years, \(SD = 17.51\) years, \(range = 17-94\) years; ‘younger’ sample: \(M = 52.10\) years, \(SD = 15.88\) years, \(range = 21-96\) years; ‘older’ sample: \(M = 57.84\) years, \(SD = 12.51\) years, \(range = 37-94\) years) was generated as the differences between the reported birth year and the year of our study. Sex indicates if the respondent was male (‘new’ sample: 47.62% male; ‘younger’ sample: 46.93%; ‘older’ sample: 48.07%). For education, a binary variable was used to indicate if the respondent held a high school diploma or not (‘new’ sample: 15.50% high school diploma; ‘younger’ sample: 17.73%; ‘older’ sample: 11.68%). Migration background was measured with a binary variable indicating if the respondent was born in a country other than Germany or had parents born in another country (‘new’ sample:
18.42% migration background; ‘younger’ sample: 9.96%; ‘older’ sample: 21.68%). The regional level indicates if the respondents were from a big city or from a less populated region (‘new’ sample: 31.04% region with more than 500,000 inhabitants; ‘younger’ sample: 29.54%; ‘older’ sample: 34.92%). All non-binary variables were z-standardized before the analyses.

Analyses

We used Cox proportional hazards regressions (Cox, 1972) to analyse the effects of the Big Five personality traits and the control variables by determining hazard ratios for overall survival. Hazard was modelled as

\[ H(t) = H_0(t) \times \exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k), \]

where \( x_1 \ldots x_k \) were the predictor variables and \( H_0(t) \) was the baseline hazard at time \( t \), representing the hazard – the instantaneous risk of dropping out at any time – for a person with the value 0 for all the predictor variables. The coefficients \( \beta_1 \ldots \beta_k \) were estimated by Cox regressions. The quantity \( \exp(\beta_k) \) can be interpreted as the relative risk of panel drop-out for an individual with an increase of one standard deviation in the value of the personality trait compared with another individual, given that both individuals are the same on all other covariates. For the dichotomous control variables, the quantity \( \exp(\beta_k) \) can be interpreted as the relative risk of panel drop-out for an individual with the risk factor present compared with an individual with the risk factor absent, given that both individuals are the same on all other covariates. A hazard ratio > 1 indicates a higher risk of panel drop-out, whereas a hazard ratio < 1 indicates a lower risk.

Drop-out was defined as a refusal to participate without any successful attempt to contact the former respondent again later, so temporary drop-outs were not taken into account.¹ The SOEP can obtain information about temporary drop-outs in subsequent waves and with additional questionnaires, hence the more serious problem are final drop-outs. Thus,
the analyses could be estimated through 2011 only because we do not yet know if 2012 drop-outs are temporary or final.

Possible predictors of panel drop-out were entered in three steps. In model 1, the influence of the five z-standardized personality traits was estimated. In model 2, variables associated with personality development – respondents’ age, age^2, age^3 – and sex were entered into the model 1 equation. Age was centred before age^2 and age^3 were calculated. In model 3, variables that have shown effects on panel drop-out – education, migration background, and regional level – were included. All analyses were computed with Stata 13.

Results

Table 1 shows the results of the Cox proportional hazards regressions models for the three subsamples of the SOEP.

Model 1 shows a significant negative effect of openness on panel attrition (or – in other words – a positive effect on panel membership) for all three subsamples of the SOEP. For example, in the ‘new’ sample a hazard ratio of .905 (p < .05) emerged for openness. This means that an individual with an openness score one standard deviation above the mean of the sample has a 9.5% lower probability per wave to drop out of the panel. In addition, significant negative effects of conscientiousness and extraversion on panel attrition emerged for the ‘older’ sample (i.e., the sample with ‘long’ duration of panel membership) but not for the ‘new’ or the ‘younger’ samples.

When age and sex were added to the analyses, the significant negative effect of openness on panel drop-out remained almost unchanged. However, in the ‘older’ sample, a significant negative effect of agreeableness on panel drop-out emerged whereas the negative effects of conscientiousness and extraversion found in Model 1 were no longer significant. Age had complex effects on panel drop-out: depending on the subsample, linear, quadratic,
and cubic age terms were significant. Sex of the respondents had no significant effect on panel drop-out.

The additional control variables in Model 3 confirm the effects found in Model 2: openness had a significant negative effect on panel drop-out in all three subsamples, whereas agreeableness had a significant negative effect in the ‘older’ sample only. Education had a significant negative effect on panel drop-out in the ‘younger’ sample, meaning that respondents with a high school diploma are more likely to stay in the sample. Migration status had a positive effect on panel drop-out in the ‘younger’ and in the ‘older’ sample, meaning that respondents with a migration background were more likely to drop out of the panel. Finally, regional level had a positive effect on panel drop-out in the ‘new’ and in the ‘older’ sample, indicating that respondents from urban areas were more likely to drop out.

Discussion

An influence of personality traits on panel attrition could be problematic for the increasingly popular use of long-running panel studies like the SOEP for research on personality development. If the reasons behind panel attrition are related to the investigated variables (i.e., personality traits), the sample could be biased (Groves et al., 2009) and analyses would be flawed.

Our survival analyses show small effects of personality traits on panel attrition in the German Socio-Economic Panel (SOEP) study. We used three subsamples of the SOEP for our analyses, each selected to be representative of the population of Germany. Our results showed no effects on panel drop-out for the Big Five dimension of neuroticism in a ‘new’ sample surveyed first in 2009 with personality measured on the same occasion, in a ‘younger’ sample with a short duration of panel membership surveyed first in 2000 with personality measured after five years of panel membership, or in an ‘older’ sample with a long duration of panel
membership surveyed first in 1984 with personality measured after 21 years of panel membership.

For conscientiousness, a small negative effect on panel drop-out (i.e., a positive effect on panel membership) emerged in the ‘older’ subsample. This is consistent with previous findings by Lugtig (2014) who investigated positive effects of conscientiousness on survey cooperation in the LISS panel.

Also for extraversion, a small negative effect on panel drop-out emerged in the ‘older’ subsample. This effect, however, is contrary to previous findings by Lugtig (2014), who found negative effects of extraversion on survey cooperation.

However, the effects of conscientiousness and extraversion vanished after inclusion of age and sex as control variables. As conscientiousness increases in young adulthood and decreases in older adulthood (Specht et al., 2011), the effect of conscientiousness on panel drop-out can be attributed mainly to age differences in panel stability. Consistently, Kroh (2013) found households of younger adults (16 to 29 years old) and households of older adults (older than 60 years) to be more likely to drop out of the SOEP than middle-aged respondents. Extraversion decreases slowly and steadily during adulthood (Specht et al., 2011). Therefore, attributing the effect of extraversion on panel drop-out to age differences in panel stability is possible as well.

For agreeableness, we found a small negative effect on panel drop-out (i.e., a positive effect on panel membership) in the ‘older’ sample after including age and sex in the model. The effect of agreeableness on panel drop-out was robust after including respondents’ education, migration background, and regional level in the analysis. Consistent with previous research (Salthouse, 2013), people high in agreeableness appear more willing to participate in a social survey, possibly because of their stronger sense of obligation to contribute to society. However, the effect of agreeableness on panel drop-out was rather small – an individual with an agreeableness score one standard deviation above the mean of the sample had an 8% lower
probability per wave of dropping out of the panel – and emerged only in the long-running panel that started in 1984.

The most robust effect was found for openness. People high in openness showed a lower probability of dropping out of the panel in all three subsamples, regardless of whether or not control variables were included. This finding corresponds to the study of Salthouse (2013). As people high in openness are considered to be interested in new experiences and intellectually curious, they may also have a greater willingness to take part in a social survey. Although responding to a social survey may not typically be considered entertaining or exciting, the diverse questions asked may nevertheless provide an interesting experience and social insights that are particularly valuable to those high in openness.

With regard to the control variables, the results were consistent with previous studies on panel attrition (e.g., Groves & Couper, 1998; Groves & Lyberg 1988; Kroh, 2013). Respondents with a higher level of education were less likely to drop out, whereas respondents with a migration background and respondents from larger cities were more likely to drop out.

Taken together, our results seem to indicate that personality has minor effects on panel membership that are limited to the personality trait of openness. For respondents in new or ‘young’ panels with only a few waves, other variables such as education or age seem to be better predictors of panel drop-out. Personality traits seem to gain importance as the panel ages and the number of survey waves increases. In our ‘older’ panel, running since 1984, the personality trait of agreeableness showed a significant but relatively small effect.

However, varying effects of personality traits in the different samples could also be due to initial differences between the samples. For instance, the initial response rates dropped from 61% and 68% for the ‘older’ sample (SOEP Samples A and B) in 1984 to 52% for the ‘younger’ sample (SOEP Sample F) in 2000 and to 33% for the ‘new’ sample (SOEP Sample
I) in 2009. However, we cannot test this at the moment because the ‘younger’ samples first need to be surveyed for as long a period as the older sample.

Nonetheless, our study can be seen as providing initial insights that point to the need for additional samples with varying ‘ages’ to explore the effects of personality on panel attrition more thoroughly. In addition, as measures of personality are included in other panel studies like the British Household Panel Study (BHPS) or the Household, Income and Labour Dynamics in Australia (HILDA) survey as well, further research could also study the effects of personality traits on panel attrition from an international perspective. However, the lack of refresher samples in the BHPS and HILDA makes it impossible to study subsamples with varying durations of panel membership with these data.

**Conclusions**

Effects of personality traits on panel attrition are either non-significant or comparably small. Therefore, it can be assumed that recent research on personality and personality development using longitudinal data from the SOEP is not biased with regard to selective panel attrition. Nevertheless, research on personality development using panel data should check thoroughly for personality-related selective attrition, for example, by comparing means and standard deviations of the personality traits of non-continuers with continuers.
References


Footnotes

1 We also estimated models considering temporary drop-outs and the results did not change significantly.
Table 1. **Hazard models to test the effect of personality on panel attrition (hazard ratios with 95% confidence intervals).**

<table>
<thead>
<tr>
<th>Age of SOEP Sample</th>
<th>New Sample Model 1</th>
<th>New Sample Model 2</th>
<th>New Sample Model 3</th>
<th>Younger Sample Model 1</th>
<th>Younger Sample Model 2</th>
<th>Younger Sample Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR, LB, UB</td>
<td>HR, LB, UB</td>
<td>HR, LB, UB</td>
<td>HR, LB, UB</td>
<td>HR, LB, UB</td>
<td>HR, LB, UB</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>.962 .891 1.038</td>
<td>.964 .892 1.042</td>
<td>.965 .892 1.043</td>
<td>1.003 .960 1.047</td>
<td>1.018 .964 1.064</td>
<td>1.013 .969 1.059</td>
</tr>
<tr>
<td>Extraversion</td>
<td>.980 .899 1.069</td>
<td>.965 .884 1.053</td>
<td>.963 .883 1.052</td>
<td>.992 .947 1.039</td>
<td>.994 .948 1.041</td>
<td>.994 .948 1.041</td>
</tr>
<tr>
<td>Openness</td>
<td>.905* .835 .981</td>
<td>.908* .837 .985</td>
<td>.906* .835 .984</td>
<td>.919*** .878 .962</td>
<td>.925** .884 .969</td>
<td>.936** .893 .980</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>1.040 .962 1.125</td>
<td>1.034 .954 1.120</td>
<td>1.031 .952 1.118</td>
<td>.986 .943 1.033</td>
<td>.972 .928 1.018</td>
<td>.969 .925 1.014</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>1.002 .921 1.090</td>
<td>1.020 .936 1.111</td>
<td>1.024 .940 1.117</td>
<td>.976 .932 1.021</td>
<td>.992 .957 1.049</td>
<td>.996 .951 1.042</td>
</tr>
<tr>
<td>Age</td>
<td>.839* .718 .982</td>
<td>.848* .722 .995</td>
<td>.941 .868 1.020</td>
<td>.931* .859 1.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age²</td>
<td>1.089* 1.019 1.164</td>
<td>1.090* 1.019 1.165</td>
<td>1.210*** 1.167 1.256</td>
<td>1.206*** 1.162 1.252</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age³</td>
<td>1.041 .980 1.106</td>
<td>1.038 .976 1.104</td>
<td>1.029* .999 1.061</td>
<td>1.031* 1.001 1.062</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (1=male)</td>
<td>1.022 .875 1.193</td>
<td>1.021 .874 1.193</td>
<td>1.068 .977 1.168</td>
<td>1.075 .983 1.175</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education (1=high school diploma)</td>
<td></td>
<td>.973 .801 1.183</td>
<td></td>
<td>.849** .752 .858</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Migrant (1=migration background)</td>
<td>1.004 .812 1.241</td>
<td>1.352*** 1.189 1.537</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional Level (1=big city)</td>
<td>1.246** 1.065 1.458</td>
<td></td>
<td>1.352*** 1.189 1.537</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: HR = hazard ratio, LB = lower bound, UB = upper bound.
Table 1. continued

<table>
<thead>
<tr>
<th>Age of SOEP Sample</th>
<th>Older Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
</tr>
<tr>
<td></td>
<td>HR</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>1.030</td>
</tr>
<tr>
<td>Extraversion</td>
<td>.943*</td>
</tr>
<tr>
<td>Openness</td>
<td>.884***</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>.952</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.948*</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Age²</td>
<td></td>
</tr>
<tr>
<td>Age³</td>
<td></td>
</tr>
<tr>
<td>Gender (I=Male)</td>
<td></td>
</tr>
<tr>
<td>Education (I=High school diploma)</td>
<td></td>
</tr>
<tr>
<td>Migrant (I=Migration background)</td>
<td></td>
</tr>
<tr>
<td>Regional Level (I=Big city)</td>
<td></td>
</tr>
</tbody>
</table>

Note.
‘New’ refers to SOEP Sample 1, started in 2009 (N = 2,400), ‘Younger’ refers to SOEP Sample F, started in 2000 (N = 5,826) and surveyed for five years before the drop-out analyses started; ‘Older’ refers to SOEP Samples A and B (N = 3,210) started in 1984 and surveyed for 21 years before the drop-out analyses started.
HR Hazard Ratio. LB Lower Bound. UB Upper Bound.
* p < .05. ** p < .01. *** p < .001.