Does the Inclusion of Non-Internet Households in a Web Panel Reduce Coverage Bias?

Published as:

Eckman S. Does the Inclusion of Non-Internet Households in a Web Panel Reduce Coverage Bias? Social Science Computer Review. 2016;34(1):41-58. doi:10.1177/0894439315572985

Stephanie Eckman

Institute for Employment Research (IAB)
Regensburger Strasse 104
D-90478 Nuremberg

stephanie.eckman@iab.de
+49 911 179 3165

Abstract

The LISS online panel has made extra efforts to recruit and retain households that were not regular users of the internet into the study. Households were provided with computers and/or internet when necessary. Including these cases made the panel more representative of the Dutch population, by bringing in respondents who were more likely to be older, to live in single-person homes and to have migration backgrounds. This paper replicates five published papers which used LISS data and explores how the conclusions in these papers would have been different had the LISS panel not included the non-internet households. There are strong demographic differences between the internet and non-internet households, and estimates of means would in many cases be biased if these households had not been included. However, across the five replicated studies, few of the published model estimates are substantively affected by the inclusion of these households in the LISS sample.

Keywords

Coverage; bias; online panels; LISS; replication

Acknowledgements

I appreciate assistance from Annette Scherpenzeel and others on the LISS team, and also from the authors of the replicated articles who were forthcoming with their thoughts and code: Rik Crutzen, Matthijs Kalmijn, Kasia Karpinska, Frederic Vermeulen, and Johan van Wilsem. Particular thanks to Ruben Bach and Markus Brendel for assistance with the replication work. The paper was also improved by the comments of the special issue editors and three reviewers. Frauke Kreuter and the participants of her research group provided feedback on the text. Jörg Drechsler, David Drukker, Tom Louis, and Joseph Schafer gave feedback on the analysis. Hannah Mautner, Svenja Wippich, Nesibe Elibol, and Natallia Hamanovich provided research assistance.

1 Introduction

Although web surveys offer cost and speed advantages over mail, telephone and face-to-face surveys, they also suffer from a unique disadvantage: undercoverage of the portion of the population that does not have the equipment or knowledge to use the internet (Couper, 2000). Whenever a portion of the population is systematically excluded from a survey, we worry that population estimates will be biased. For example, consider a web survey measuring support for lowering the voting age. If older people are less likely to support this change than younger persons, and older people are also more often missed in web surveys because they do not use the internet, the survey's estimate of support for the change in the voting age may overstate the true level of support in the population.

To combat this bias, the LISS study, a monthly web panel survey conducted since 2007 in the Netherlands, has made extraordinary efforts to bring non-internet households into the panel. Selected households that lacked the equipment needed to participate were provided with computers and/or internet connections to allow them to take part. These efforts were costly, but they did make the LISS panel more representative of the Dutch population by bringing in more respondents who were older, who lived in single-person homes and who had migration backgrounds (Leenheer and Scherpenzeel, 2013).

This paper extends the analysis of Leenheer and Scherpenzeel beyond demographics and household characteristics and explores how the inclusion of these non-internet households has affected estimates in five published articles based on LISS data. The counterfactual situation explored below is: how would the results of these published analyses have been different, had the survey not provided computers and internet access to the non-internet households?

2 Background

Although many surveys suffer from undercoverage, it is a particular concern in web surveys, because not all members of the population are able to participate. Some do not have access to a computer; others do not have access to the internet, or lack the skills to use the internet. These people cannot take part in the survey, whether they want to or not (Couper and Coutts, 2006). Whenever a survey undercovers some members of the population, coverage bias may result. If those who have the means to participate in a web survey have different characteristics than those who do not, then an estimate using only the covered cases may be biased.

Bias in means due to undercoverage is similar in form to bias due to nonresponse (Groves and Couper, 1998, Chapter 1). Undercoverage bias in the mean of a variable Y is the difference between the mean from the covered cases ($\overline{\mathrm{Y}}_{\mathrm{c}}$) and the mean from the entire population ($\overline{\mathrm{Y}}_{\mathrm{pop}}$):

$$\text{Bias}(\overline{Y}_c) = \overline{Y}_c - \overline{Y}_{pop} \qquad (1)$$

Undercoverage bias can also be expressed as a function of the undercoverage rate and the difference between the mean of Y on the covered cases ($\overline{Y}_c$) and the mean of Y on the undercovered cases ($\overline{Y}_{uc}$) (Lessler and Kalsbeek, 1992, p. 60):

$$\text{Bias}(\overline{Y}_c) = \frac{N_{uc}}{N_{pop}} [\overline{Y}_c - \overline{Y}_{uc}] \qquad (2)$$

Here N_{uc} is the number of people who are undercovered and N_{pop} is the total number of people in the population. (Throughout this paper, capital letters refer to population values and lower case letters to sample estimates. Thus *Bias* refers to true undercoverage bias in the population and *bias* is an estimate from sample data.)

Equation 2 can help us think through how much undercoverage bias may exist in data collected via web surveys. The first term, the undercoverage rate, can be derived from published statistics. In the Netherlands in 2007, the year of the initial LISS panel recruitment, 83% of all households had access to the internet (Eurostat, 2013), meaning the undercoverage rate was 17%. (As we shall see below, the unit of selection in the LISS survey is the household, and thus I have switched here from talking about covered and undercovered *persons* to speaking about *households*.)

The second term in Equation 2 changes with each variable we are interested in. Several studies have explored the differences between the internet and non-internet populations and found demographic, attitudinal and behavioral differences between the two. In the Netherlands, those without access to the internet tend to be older, to live alone and to have migration backgrounds (van der Laan, 2009; Leenheer and Scherpenzeel, 2013). Bosnjak et al (2013) found that Germans without internet access are less educated and slightly older than those with access; they also detect some personality differences between the two groups. Most studies of the characteristics of the internet and non-internet populations have been done in the US. Couper (2000) reports demographic differences from the Current Population Survey: internet households are different in terms of age, race, income, college education, and urbanicity (all in the expected directions). Beyond demographic differences, Robinson et al (2002) found that internet users have more tolerant attitudes. Dever et al (2008) explored health reports: internet users are more likely to be in good health, to have health insurance, and to exercise regularly. Zhang et al (2009) found more community-oriented behavior and higher reports of political engagement among internet users. In a survey of older Americans, Couper et al (2007) and Schonlau et al (2009) detected significant differences between internet users and non-users on financial and health variables. Thus the second term in Equation 2 is non-zero for many variables, and surveys that undercover non-internet users do suffer from coverage bias.

We might hope that as time passes, internet penetration will increase, the undercoverage rate (the first term in Equation 2) will decrease, and thus bias will get smaller and smaller.

However, it is also possible that, as the non-internet population shrinks, it will become more and more different from the rest of the population. In terms of Equation 2, this means that as the undercoverage rate (the first term) decreases, the difference between the covered and undercovered populations (the second term) increases, and undercoverage bias may increase (Couper, 2000; Scherpenzeel and Bethlehem, 2010). Thus, even in a high internet penetration country, such as The Netherlands, bias due to undercoverage may still be a problem.

Equations 1 and 2 capture bias due to undercoverage in means. Many researchers believe, or hope, that coverage bias in other statistics, such as subgroup means and regression coefficients, will be smaller (Groves, 1989, p. 86). Several studies have explored this question empirically, considering the differences between the internet and non-internet populations after controlling for demographic characteristics. The results of these investigations are mixed. Some studies find that conditioning on control variables can eliminate the bias on many variables (Robinson et al, 2002; Dever et al, 2008). Others find that bias still remains (Couper et al, 2007; Schonlau et al, 2007; Schonlau et al, 2009; Zhang et al, 2009). These diverse results are likely due to the different populations and time periods studied, as well as the different variables.

There is no simple formula like Equation 2 that helps us understand undercoverage bias in regression coefficients from a multivariate model. Bias depends on the relationship between the variables in the model and the mechanism behind undercoverage (Heckman, 1979; Goldberger, 1981; Groves, 1989, p. 94). In the context of web surveys which undercover non-internet households, bias in regression coefficients depends on whether the independent variables in the model are related to internet access. If variables that correlate with coverage are included in the model, undercoverage bias in the regression coefficients should be reduced or eliminated.

This paper explores bias due to undercoverage of non-internet households in real analyses run with LISS data and published in peer-reviewed journals. I replicate summary statistics and regression models from five papers and then remove the non-internet households and re-run the analyses, to study how the results would have been different had the non-internet households not been included in the LISS study. Models that include independent variables that are also good predictors of the support indicator should show little change in the estimated coefficients; models that do not contain such control variables should show more change. Due to previous work by Leenheer & Scherpenzeel (2013), we know that households that needed support to join the LISS panel are older and more likely to live alone and have a migration backgrounds. Controlling for factors such as these in a regression should reduce the susceptibility of a model to coverage bias.

3 Data

The LISS panel began in 2007 with the recruitment of 5,000 households from a simple random sample of 10,150 addresses selected from the address register of the Netherlands.

What makes the LISS panel different from most other internet panels is that it attempted to recruit all households, including non-internet households, which are usually excluded from web surveys. If a household agreed to participate, but had no computer or no internet connection, the study provided what was needed. When necessary, the panel set up a computer and installed a broadband internet connection. The panel sample has been refreshed every second year, and the same support was offered (van der Laan, 2009; Scherpenzeel and Das, 2010).

The panel maintained an indicator of which households required these forms of assistance in order to participate, and the availability of this flag makes the LISS data set an ideal medium for studying undercoverage bias. Below, I refer to the flagged households, those which needed assistance such as an internet connection or a computer, as *supported* households. As of 2010, 545 households required support, 9.5% of all households in the LISS panel (Leenheer and Scherpenzeel, 2013). In the analyses below, this percentage fluctuates a bit because the composition of the panel changes over time, as panel members do not respond every month.

There are additional sources of undercoverage that may lead to coverage bias in LISS estimates that are not explored here. First, the support flag is at the level of households, not persons. It may be the case that some households had a computer and internet access, but not all members of the households were able to use them. These household members are also undercovered, but are not flagged in the data set, and the results below do not capture bias due to this type of undercoverage. Another uncaptured source of undercoverage is due to missing addresses on the central address frame from which the sample was selected.

The panel sends monthly survey requests to all members, giving an incentive to those who participate (7.50€ per 30 minutes). The surveys consist of the LISS core study, which tracks members' living situations over time, plus additional modules proposed by outside researchers. All LISS data are available for download, after registration, from www.lissdata.nl. The support indicator is not included in these files, however, LISS researchers provided me with this flag.

4 Methods

Using the indicator of which households received support, I estimate the bias that would have been introduced had the LISS panel not made additional efforts to bring in the non-internet households. To explore this counterfactual situation, I first replicate a published analysis on the full LISS data set, and then run the same analysis again, excluding the supported households. If the two estimates differ in a meaningful way that would likely have changed the authors' conclusions, then the efforts to recruit and retain the non-internet households reduced coverage bias. This approach to estimating coverage bias using survey paradata is suggested in Eckman (2013). It relies on the assumption that the exclusion of the non-internet households would have had no other effects on the collected data, for example, that none of the funds used to provide support would have been diverted into increasing the recruitment rate and thus changing the make-up of the panel.

Analyses from five articles were replicated: van Wilsem (2011); van Wilsem (2013); Crutzen and Göritz (2011); Kalmijn (2013); Karpinska et al. (2013). These papers cover a variety of fields, from criminology to sports research, to family studies. Four criteria were used to choose papers for replication. First, papers must be published in peer-reviewed journals. Second, papers must include a table of summary statistics in addition to one or more regression models, because the summary statistics assist in replication. Third, papers must not involve a wording or other methodological experiment, because in such papers inference is not to the larger population but only between the experimental conditions. Four, the papers' models must run in Stata software. See Table 1 for more information on the replicated papers.

<Table 1 about here>

The summary statistics and the code provided by each author eased the replication task. Nevertheless, I encountered troubles with some studies matching the number of cases and the summary statistics. The LISS data may itself have changed slightly between the initial release some of the authors used and the later release now available for download from the website. Details about the successes and challenges in replicating each article are provided in Online Appendix A.

For each of the five replicated studies, I present two sets of estimates: means and model estimates. I first estimate the population mean of all variables and cases used in the model (for papers with more than one model, I use the final (full) model, or the model with the most cases): this estimate is $\bar{y}_{\mathrm{pop}}$. I then estimate the means using only those cases in households that did not require support to join the LISS panel: $\bar{y}_{\mathrm{c}}$. For each variable, the difference between the second mean and the first is an estimate of the undercoverage bias:

$$\mathrm{bias}(\bar{Y}_{\mathrm{c}}) = \bar{y}_{\mathrm{c}} - \bar{y}_{\mathrm{pop}} \tag{3}$$

That is, Equation 3 estimates the population undercoverage bias in Equation 1. To test whether undercoverage bias estimated in Equation 3 is significantly different from zero, I run a linear regression model where the dependent variable is the variable of interest (Y) and the sole independent variable is the support indicator: when the coefficient on the support variable is significant, the means between the supported and unsupported cases are different and undercoverage bias in the mean is significant.

Bias is difficult to compare across variables, however, because different variables are in different units. For this reason I also estimate absolute relative bias:

absolute relative bias$(\bar{Y}_c) = \left|\frac{\bar{y}_c - \bar{y}_{pop}}{\bar{y}_{pop}}\right|$ (4)

This measure expresses undercoverage bias as a percent of the mean and thus allows comparisons across variables. (To a first approximation, Equation 4 estimates the population absolute relative bias $\left|\frac{\bar{Y}_c - \bar{Y}_{pop}}{\bar{Y}_{pop}}\right|$, see Kendall and Stuart (1969), p. 351). Testing the significance of the relative bias estimates is not straightforward, however, because it is a ratio of two estimated quantities. Thus, I report both bias and absolute relative bias but test the significance of the bias only. In addition, a Hotelling test for each paper reports whether the supported and unsupported groups have different means across all the variables used in the model.

The next step in the analysis involves replicating the papers' regression model(s) and then re-estimating the models without the cases from the supported households. For each paper, I present a figure showing the model estimates (from the replication) when the entire data set is used, alongside the results when the supported households are excluded. These figures show visually how the point estimates and confidence intervals change. To get a more quantitative measure of the effect of the removal of the supported households on the model estimates, I use the overlap measure developed by Karr et al (2006), which is a measure of model agreement at the coefficient level. For each estimated coefficient, I find the intersection of the confidence intervals around the full-sample estimate and the reduced-sample estimate. I then calculate the percent of the full-sample confidence interval that is inside the intersection, and the percent of the reduced-sample confidence interval that is inside the intersection, and average the two percentages. The overlap measure ranges from 0% (the intersection is empty) to 100% (the two confidence intervals entirely coincide). The higher the percentage, the greater the agreement between the two models and thus the less the model estimates are affected by the removal of the supported cases. For each paper, I then present the distribution of these measures.

Finally, logistic regression models explore the ability of the models' independent variables to predict the support indicator. When a model includes controls that correlate with the support variable, it is less vulnerable to coverage bias in coefficient estimates. The area under the receiver operating characteristic (ROC) curve captures the models' ability to discriminate between the supported and unsupported households. This measure is also a percent, with higher values indicating better discrimination.

Online Appendix A discusses the issues encountered in the replication of each paper. Tables provide the replicated summary statistics and note any discrepancies with the original paper. Additional tables give the coefficient and standard error estimates (or *t*-statistics) from the original papers, the replications on the full data set, and the replications on the reduced data set.

This study does not explore the effects of nonresponse on any of the estimates, although nonresponse can also lead to bias in means and regression coefficients. Leenheer and

Scherpenzeel (2013) show that the rate of recruitment into the LISS panel was lower among non-internet households (35%) than among internet households (84%). However, under the assumption that the recruitment of cases into the LISS panel is independent (successful or unsuccessful recruitment of one household is not related to the recruitment of others), nonresponse does not bias the calculations of undercoverage bias in this paper.

If the papers' authors had used weights that adjusted for nonresponse and/or undercoverage, the counterfactual situation explored in this paper might impact the weights – had the non-internet households *not* been included in the data set, the adjusted weights would be different. However, no such complication occurs here because the panel data set does not contain any weights, and none of the papers replicated below used any researcher-derived weights.

5 Results

For each of the replicated papers, I briefly discuss the hypotheses, models and findings. I then present estimates of undercoverage bias in means. Figures compare the estimated coefficients from the replicated models to those from the models run on the counterfactual data set, the one from which the supported households have been dropped. I conclude with a discussion of how the paper's conclusions may have changed, had the analyses been run on the counterfactual data set.

Paper 1 – van Wilsem (2011)

Van Wilsem (2011) explores the characteristics that make people more likely to be threatened, both in the real world and online. Van Wilsem hypothesizes that a person's online behavior can make one susceptible to real-world threats, and vice versa. Using the LISS data, he runs multilevel multinomial regression models. The dependent variable is a four-category indicator of the threats a person received: no threats (the reference category), traditional threats only, digital threats only, and both kinds of threats. The important independent variables for the hypothesis tests are various measures of participation in online and offline activities, such as online shopping, use of a webcam, going to restaurants, etc. The model also includes demographic controls as independent variables, such as age, education, household income, household size, and others. Van Wilsem finds support for his hypothesis, noting that variables relating to digital activities, such as hours spent online and whether one has a social network profile, correlate with whether one has received traditional threats, and out-of-home activities, such as shopping and commuting, relate to whether one has received digital threats. He also notes that those who are offenders of digital crime are more likely to receive threats (traditional, digital, or both).

The first row of Table 2 summarizes the undercoverage bias in the means of all variables used in the van Wilsem (2011) model. Estimates of means of 69% of the 29 variables he uses would have significant undercoverage bias if the cases in the unsupported households were removed from the data set. The absolute relative bias ranges from just over 0% to 18%. One of the four values of the multinomial dependent variable shows bias: those living in unsupported households are more likely to receive digital threats than those in the sample as a whole, probably because those in unsupported households are more regular internet users. Many independent variables are also biased when we drop the supported households. (Estimates of means and bias for every variable are reported in Table B1 in Online Appendix B.) The people who live in supported households are older and have smaller households. A Hotelling test of the difference in means for all variables and cases used in the model also shows that the supported and unsupported groups have different characteristics on these variables (Table 2).

<Table 2 about here>

Figure 1 presents the estimated relative risk ratios (the exponentiated coefficients) from the replicated multinomial logit model on the full data set, and those estimated on only the cases which did not require support. The figure is composed of three panels, each of which contains results from one of the three parts of the multinomial model. The vertical line corresponds to a relative risk ratio of one: when a confidence interval overlaps this line, the estimated coefficient is not significantly different than zero.

For each variable along the vertical axis, there are two coefficient estimates and confidence intervals shown. The first (top) corresponds to the relative risk ratio from the replication (which are similar but not identical to the van Wilsem (2011) results, see Online Appendix A), and the second (bottom) to the ratio calculated from the unsupported households only. The replicated model is based on 6,370 cases, 5.9% of which required support.

<Figure 1 about here>

The coefficients and their confidence intervals change slightly when the cases needing support are dropped. As discussed in the methods section, the percent of the overlap of the two confidence intervals is a measure of the similarity of the estimates across the two models. The distribution of this overlap measure across all variables in the model is shown in Figure A1 in the appendix. For all but one of the sixty-nine coefficients, the percent overlap is greater than 75%, indicating the model estimates do not change much. There is one variable where the overlap is 0%, meaning the confidence intervals do not overlap at all: Missing Value, Low Self Control in the third part of the model; its estimated coefficient and confidence interval can be seen near the bottom of the third panel of Figure 1. However, this variable is not one that van Wilsem discusses, so the large change in this estimate would likely not have affected his interpretation.

The patterns of significance that he draws upon in concluding that “actions in one domain can have consequences in the other” (van Wilsem, 2011, pp. 122-123) are largely unaffected by the removal of the supported households from the data set: there are still online activities which are significant in the traditional threat portion of the model, and offline activities that are significant in the digital threat portion. However, one offline activity, shopping, is significant in van Wilsem's model, but not in the full sample replication or with the reduced data set, as can be seen in Table A1.2 in Online Appendix A. This change somewhat reduces the strength of his conclusions, but it is not due to the removal of the supported households.

Another change is in the offender of digital crime variable in the traditional threat portion of the model: the estimated coefficient is significant in the replication and not in the model with only the unsupported households. Van Wilsem does interpret this coefficient, though it was not part of his initial hypotheses. He says "offenders run higher risks for each of the three types of threat victimization….[I]nterestingly, traditional threat victimization is also affected by digital offending. These findings suggest that sending digital threats and spreading computer viruses may lead to retaliation, in either a digital or a traditional way." Because the relative risk ratio on the offender variable is not significant after dropping the households that needed support, he may have come to a different conclusion here.

Paper 2 – van Wilsem (2013)

In this article, van Wilsem hypothesizes that persons with lower self-control will be at higher risk for internet consumer fraud victimization and uses a logistic model, with victimization as the dependent variable, to explore the factors associated with this event. The independent variables are measures of online activities (for example: internet shopping, use of social networking sites) and demographic controls (for example: age, education, household income, household composition). Only 2.5% of the cases in his data set, or 152 persons, experienced online victimization, and of these, only two were members of supported households, which is likely due to the fact that these households are not frequent internet shoppers.

As shown in Table 2, the means of 60% of the 15 variables in the full model change significantly when cases in the supported households are dropped. The relative bias in these variables ranges from 0% to 3.5%. There is significant undercoverage bias in age, education, and several variables associated with internet use, which is as expected (see Table B2 in Online Appendix B for details). The Hotelling test for differences between the supported and unsupported households suggests that they are significantly different on the variables used in the model (Table 2).

Although van Wilsem runs four nested logistic models, Figure 2 shows only the last, full model. This model is based on 6,157 cases, 4.5% of which are members of supported households. We see very little difference between the estimated odds ratios and confidence intervals from the replication with the full set of cases and the replication without the supported cases. The estimates are always very close and have the same significance level (see also Table A1.2 in Online Appendix A). Figure A1 shows that the percent overlap between the confidence intervals from the two models is greater than 94% for all variables.

Van Wilsem's interpretation of the significant odds ratio on low self-control as support for his hypothesis that low self-control is related to victimization would likely have been unaffected by the removal of the supported households from the LISS study.

<Figure 2 about here>

Paper 3 – Crutzen and Göritz (2011)

In this paper, Crutzen and Göritz investigate whether survey respondents' reports of their physical activity levels are biased by socially desirable responding. They use the Marlow-Crowne scale to measure each respondent's tendency to engage in socially desirable responding, and investigate whether this scale correlates with reports of physical activity. They run five linear regression models where five different types of physical activity are the dependent variables: walking, moderate-intensity activity, vigorous-intensity activity, sedentary behavior and total physical activity. Each model contains the social desirability scale as an independent variable, as well as age, sex, income, and education. The authors show and interpret standardized coefficients only for the social desirability scale. In each model, the coefficient is near zero and insignificant. These results lead Crutzen and Göritz to conclude that there is no social desirability bias in self-reports of physical activity in the LISS panel.

Due to missing data, the five models use different case bases, and the share of supported cases in the models varies from 6.5 to 8.1%. Using the cases in the sedentary model, which has the largest case base and the largest share of cases from supported households, the means of half of the variables have undercoverage bias that is significant and bias is up to 2% of the mean. The Hotelling test run on all variables and cases used in the model rejects the null hypothesis that the two groups of respondents have the same means on the variables used in the model (Table 2).

Figure 3 compares the standardized coefficients from the five models. The coefficient on the social desirability scale is still near zero and insignificant after the supported households are removed (Figure 3). We do see discrepancies in the significance of a few other coefficients in Figure 3, but Crutzen and Göritz do not discuss or even present these estimates in the paper. They would almost certainly have come to the same conclusion had they used only the cases in unsupported households.

<Figure 3 about here>

Paper 4 – Kalmijn (2013)

Kalmijn explores the effects of divorce on the relationship between adult children and their fathers. He puts forth three competing hypotheses for the negative influence of divorce on the father-child relationship: the swapping-families hypothesis; the stepmother hypothesis and the need hypothesis. He runs four linear models where the dependent variables are measures of the quality of the relationship between fathers and their children: contact frequency, perceived quality of relationship with father, and support from and to father. The independent variables relate to the current relationship status of the father and mother, and controls for the adult child's age, education, marital status, etc. (Note that it is the adult child who is the LISS respondent, not the father.) To distinguish between the swapping-families and the need hypothesis, he runs the same models on the mothers' relationships with the children. Kalmijn finds that divorced fathers have less frequent and lower quality contact with their children, and the effect is stronger for those fathers who are repartnered. The same pattern does not hold in the mother models. Kalmijn interprets these results as support for the swapping-families hypothesis.

The share of supported cases in the eight models is 3.4 - 4.8%. Tables 2 and B4 (in Online Appendix B) show the bias and absolute relative bias that would appear in the means of Model 1's dependent and independent variables, if the cases in supported households were not in the data set. In this model, the dependent variable, contact frequency with father, would be biased by more than 30% if the supported households were not covered, but the bias here does not reach significance. A Hotelling test on all cases and variables in Model 1 rejects the null hypothesis that the two groups of respondents are the same on these variables (Table 2).

Figures 4 and 5 present the results of Kalmijn's eight main models. Again, for each variable two coefficients and their confidence intervals are shown. The pairs of estimates are quite similar across all eight models, a point also made in Figure A1: the overlap measure is greater than 83% for all coefficients. The sign and significance of all but three estimated coefficients from the models are unchanged when the supported households are dropped from the data set. Two coefficients lose significance, and these are both in the mother models: in Model 6 the coefficient on child higher educated and in Model 8 that on mother repartnered. The seemingly-unrelated regression results testing for equivalence of coefficients in the mother and father models and eight additional regressions (not shown) are similarly unaffected. (One of Kalmijn's four seemingly-unrelated regression results, shown in his Table 3, is significant at the 5% level and only at the 10% level after I remove the supported cases.) I infer that Kalmijn would have come to the same conclusions about the mechanisms behind the degradation in relationships between divorced fathers and the adult children had he used the reduced data set.

<Figure 4 about here>

<Figure 5 about here>

Paper 5 – Karpinska et al

The fifth replicated paper, Karpinska et al (2013), uses a vignette study to explore managers' interest in retaining older workers. The authors put forth several hypotheses about organizational contexts that influence retention, and about the attributes of the workers and the managers themselves that correlate with retention. They run three linear multilevel models to test these hypotheses – the first on all managers, the second for those supervising low-skilled workers and the third for those supervising high-skilled workers. For each model, the dependent variable is the eleven point rating of how likely the manager would be to retain the worker described in the vignette. The independent variables capture the organizational context and worker characteristics manipulated in the vignettes (for example, whether there is a labor shortage and worker age, health, and manageability), and the managers' characteristics (for example, age, gender, and stereotypes held about older workers). The authors find support for many of their hypotheses. With respect to organizational context, managers who work in firms facing labor shortages are more likely to recommend retention. With respect to worker characteristics, "the human capital of older workers, their health status, flexibility, attitudes towards retirement, training motivation, and manageability contribute to their retention chances" (Karpinska et al 2013, p. 9). With respect to manager characteristics, the models show that norms about the appropriate age for retirement affect the retention recommendation, as do managers' perception of the soft skills of older workers.

Only LISS respondents who reported that they were in managerial roles were asked to participate in the study. For this reason, the sample size for this study is small, 238 respondents (1,190 vignettes). Five percent of these vignettes came from respondents who lived in supported households, and thus would not have appeared in the data set, had the LISS study not offered support for non-internet households. In Table 2 we see that only seven percent of the 26 variables used in Model 1 would suffer from significant undercoverage bias if we were to drop the cases in supported households, fewer than in any of the other papers. The absolute relative bias estimates for this paper are also low, and no variable would be biased by more than 2.1%. Nonetheless, the Hotelling test here again rejects the null hypothesis that the two groups of respondents are the same on these variables.

<Figure 6 about here>

Figure 6 summarizes the results of the replications of the three models. In Models 1 and 2, the signs of the coefficients and the significance patterns across the original model, the full-sample replication and the replication without the supported households are the same (Table A5.2, Online Appendix A). In the third model, one coefficient loses its significance, the indicator for employee's age in the vignette, which is not one the authors interpret. The final panel of Figure A1 shows that all confidence intervals in the two replications overlap by 87% or more, showing little change in the model estimates due to undercoverage of supported households. It is likely that the authors would have come to the same conclusions had they used only the unsupported cases in their analysis.

The author(s) of each of the papers included variables correlated to internet access in their models, even when they were not part of their hypotheses. In the LISS sample, age, living alone and migration background are related to support (Leenheer & Scherpenzeel, 2013), and all papers included one or more related variables. Paper 1 included age and household size; Paper 2, age, household size and household composition; Paper 3, age and marital status; Paper 4, age and marital status; Paper 5, age. Papers 1, 2 and 3 also control for household income, which is likely strongly correlated with age, living alone and migration background, and thus with internet access. Table 3 shows how well the independent variables used in these papers discriminate between the supported and unsupported households in logistic regressions. The area under the ROC curves for these five logistic models are all high: Hosmer and Lemeshow (2000, p. 162) state that values between 0.7 and 0.8 indicate acceptable discrimination, and values between 0.8 and 0.9, excellent discrimination. The inclusion of variables that predict support protects the papers' substantive models from undercoverage bias when the supported households are dropped. Notably, the authors of these studies did not include these variables in order to reduce undercoverage bias: they did not need to, as the LISS data that they used included both internet and non-internet households. Nevertheless, all of the models control for age and other demographics, and thereby reduce or eliminate the effects of undercoverage.

<Table 3 about here>

6 Conclusions and Implications

The results above demonstrate that undercoverage of the non-internet households would in most cases not have introduced bias into model estimates in these papers and would likely not have changed the authors' interpretations of the model results. Although there are significant differences between the supported and unsupported cases on the variables used in the models, as shown in Table 2 and Online Appendix B, the model estimates in these papers are affected to a lesser extent. Perhaps in Paper 1, the author would have come to different conclusions about the role of offline activities in digital threats, as one of the two variables he relies on is no longer significant when the supported households are excluded. In the other papers, however, the changes seem to be too small to affect the authors' conclusions. It appears that the analyses in these papers were protected from undercoverage bias by the small undercoverage rates and the authors' inclusion of control variables correlated with undercoverage.

However, we should be cautious in extrapolating from these five studies. The analyses replicated here span diverse research fields and several different types of regression models and the dependent variables they explore range from strongly related to internet access (paper 2, victim of internet fraud) to unrelated to internet access (paper 4, quality of father-son relationship). Nevertheless, the fact that the conclusions in these papers probably would not change in the counterfactual scenario does not mean that other papers based on LISS data would also be unaffected. These findings also do not rule out the possibility that undercoverage bias in model estimates could occur in other web surveys and other countries. The Netherlands has a particularly high internet penetration rate (Eurostat, 2013). In populations with lower penetration rates, undercoverage bias in model estimates is more likely. Studies that focus on the older population or migrants, or compare these groups to others, are particularly likely to be biased, if the supported households are not included.

For users of web surveys that do not have the resources to include non-internet households (by offering support or using a different mode), the results in this paper offer a few pieces of advice. Estimates of means do seem to be more vulnerable to undercoverage bias than estimates from multivariate models, as suggested by Groves (1989, p. 86). Those who design these web surveys should measure correlates of internet coverage. Analysts of such data sets should include these variables in their substantive models in addition to the predictors of interest, because this approach may provide some protection from undercoverage bias.

More theoretical statistical work and/or simulations could provide guidance on when and which model coefficients are susceptible to undercoverage bias. These analyses could also be carried out with the LISS data and the support flag, but are outside of the scope of this paper.

The only way to be certain that a web survey does not suffer from undercoverage bias is to cover the non-internet portion of the population, as the LISS study has done. The support offered by the LISS study is a form of insurance against possible coverage bias. This insurance is expensive, however, and more studies like this one are needed to understand whether it is worth the cost, especially as the internet penetration rate rises.

Table 1: Papers Replicated

Article	**Journal** (all peer reviewed)	**Citations in Google Scholar**[a]
van Wilsem, J. (2011). Worlds Tied Together? Online and Non-Domestic Routine Activities and their Impact on Digital and Traditional Threat Victimization	European Journal of Criminology	15
van Wilsem, J. (2013). 'Bought it, but Never Got it' Assessing Risk Factors for Online Consumer Fraud Victimization	European Sociological Review	15
Crutzen, R. and Göritz, A. S. (2011). Does Social Desirability Compromise Self-Reports of Physical Activity in Web-Based Research?	International Journal of Behavioral Nutrition and Physical Activity	8
Kalmijn, M. (2013). Relationships Between Fathers and Adult Children: The Cumulative Effects of Divorce and Repartnering	Journal of Family Issues	4
Karpinska, K., Henkens, K. and Schippers J. (2013). Retention of Older Workers: Impact of Managers' Age Norms and Stereotypes	European Sociological Review	6

[a] As of January 5, 2015

Table 2: Comparison of Means in Analysis Variables

Paper	Model	% Cases in Supported HHs	% Model Variables with Significant Undercoverage bias[a] (n)	absolute relative bias[b] Min. (%)	Median (%)	Max. (%)	Hotelling Test
1	N/A	5.9	68.9 (29)	0.022	1.61	18.07	F(26, 6343) = 30.0*
2	Full	4.5	60.0 (15)	0.031	1.73	3.47	F(15,6141) = 24.9*
3	Sedentary	8.1	50.0 (6)	0.0036	0.20	2.11	F(6,5087) = 53.8*
4	Model 1	3.8	21.4 (14)	0.0027	0.92	31.43	F(14,2265) = 2.6*
5	Model 1	5.0	6.9 (29)	0.0045	0.63	2.09	F(15,1174) = 2.5*

* $p < 0.05$
[a] Bias estimated as shown in Equation 3, significance testing as described in Methods section
[b] Estimated as shown in Equation 4
NOTE: See tables in Online Appendix B for estimates of full sample means and means on unsupported cases, sample sizes, and estimates of bias and absolute relative bias.

Table 3: Discrimination of Models of Support Indicator, Using Independent Variables from the Published Models

	Model	Area Under ROC[a]
1	N/A	0.84
2	Full	0.80
3	Sedentary	0.73
4	Model 1	0.70
5	Model 1	0.71

[a] Area under the receiver operating characteristic curve

Figure 1: Comparison of Relative Risk Ratios from Multinomial Model: Full data set vs. Unsupported Households Only, van Wilsem (2011)

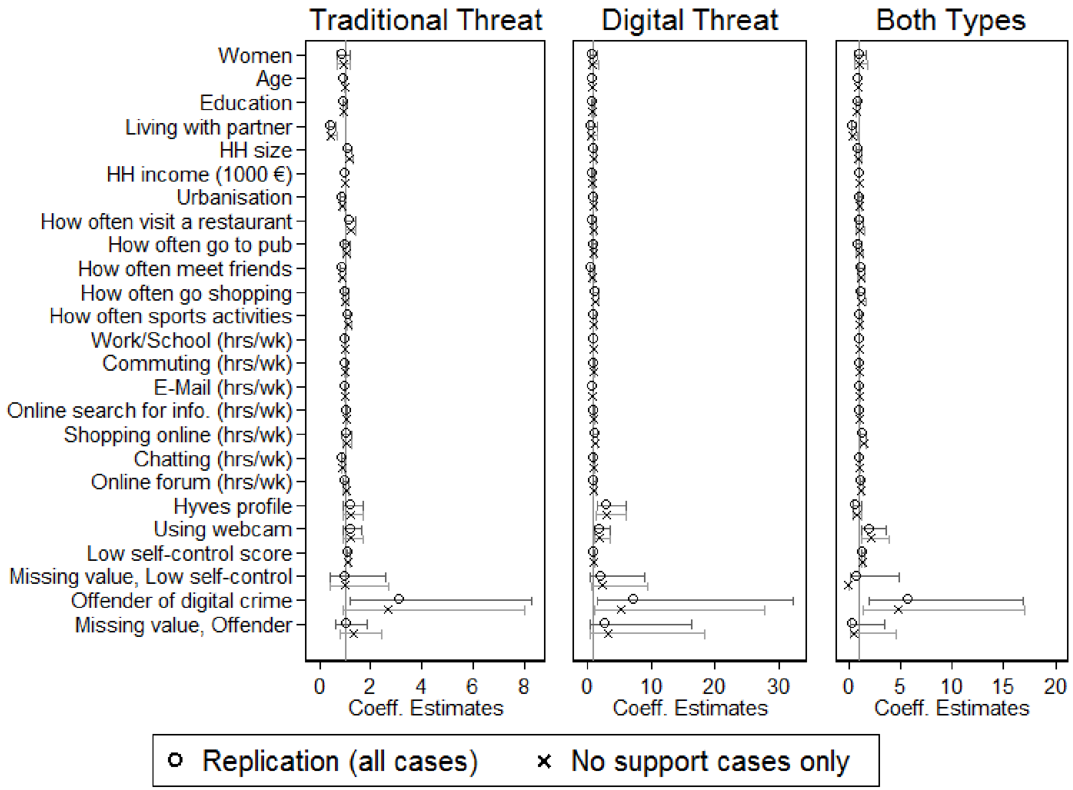


Figure 2: Comparison of Odds-Ratios: Full data set vs. Unsupported Households Only, van Wilsem (2013)

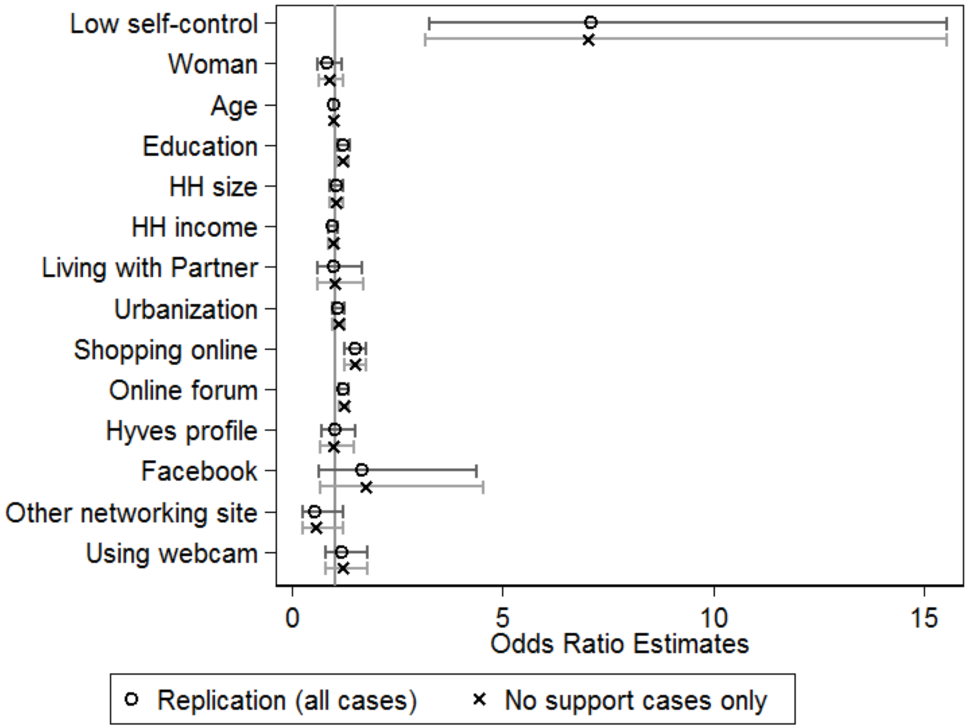


Figure 3: Comparison of Coefficients from Five Models: Full data set vs. Unsupported Households Only, Crutzen and Görtiz (2011)

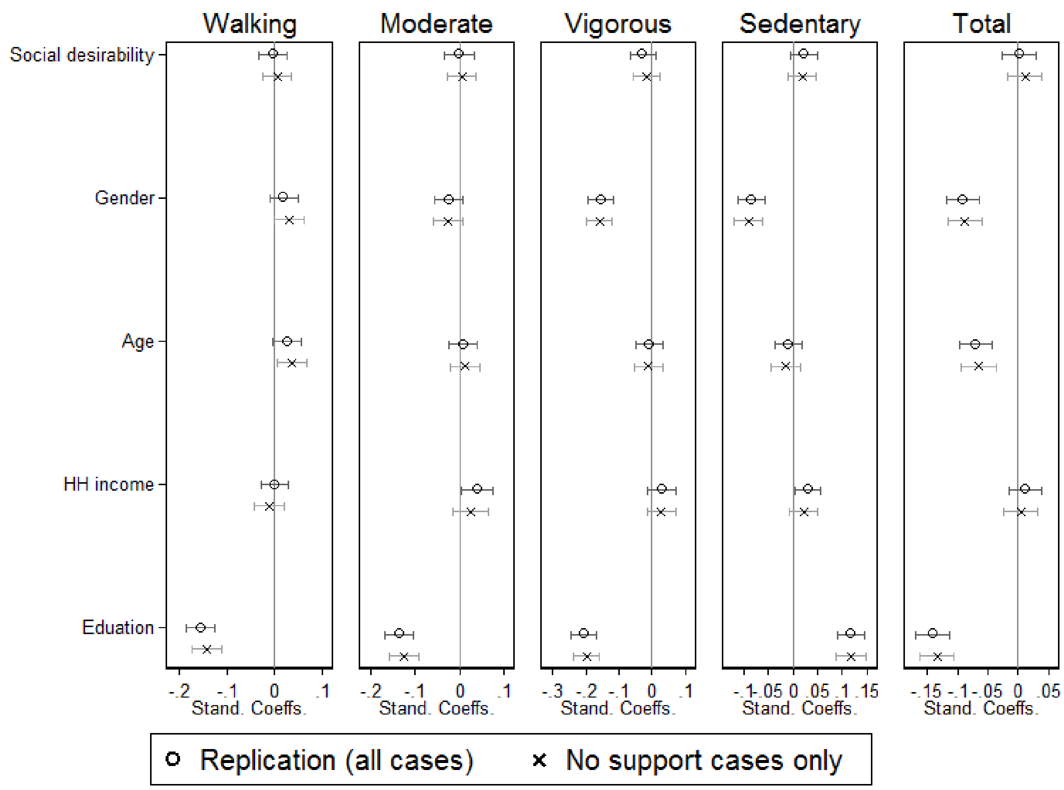


Figure 4: Comparison of Coefficients from Father Models: Full data set vs. Unsupported Households Only, Kalmijn (2013)

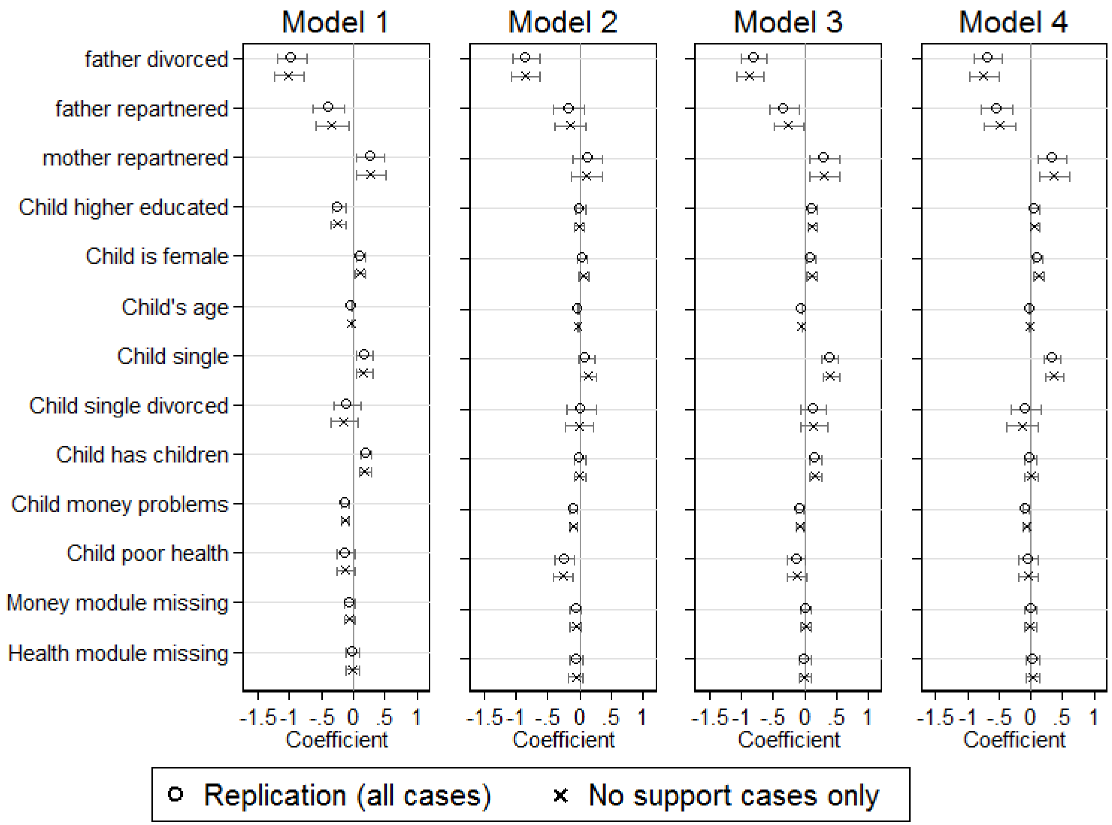


Figure 5: Comparison of Coefficients from Mother Models: Full data set vs. Unsupported Households Only, Kalmijn (2013)

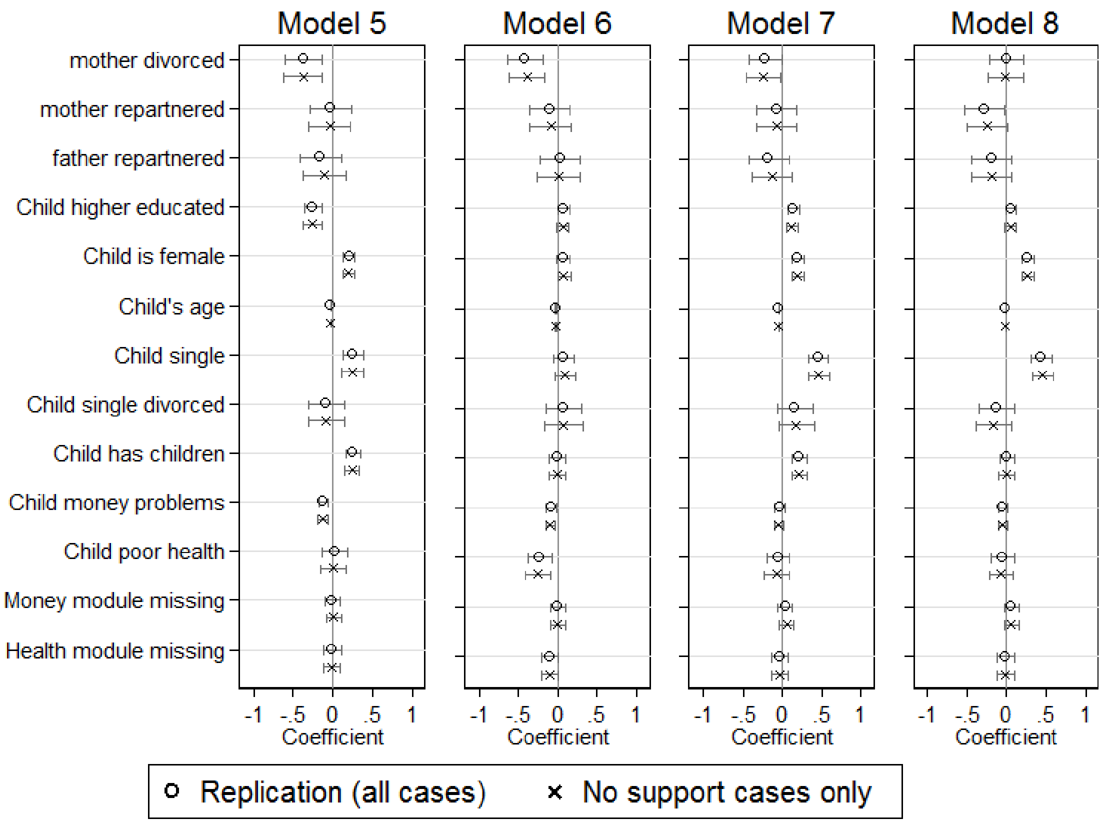


Figure 6: Comparison of Coefficients from Multilevel Models: Full data set vs. Unsupported Households Only, Karpinska et al (2013)

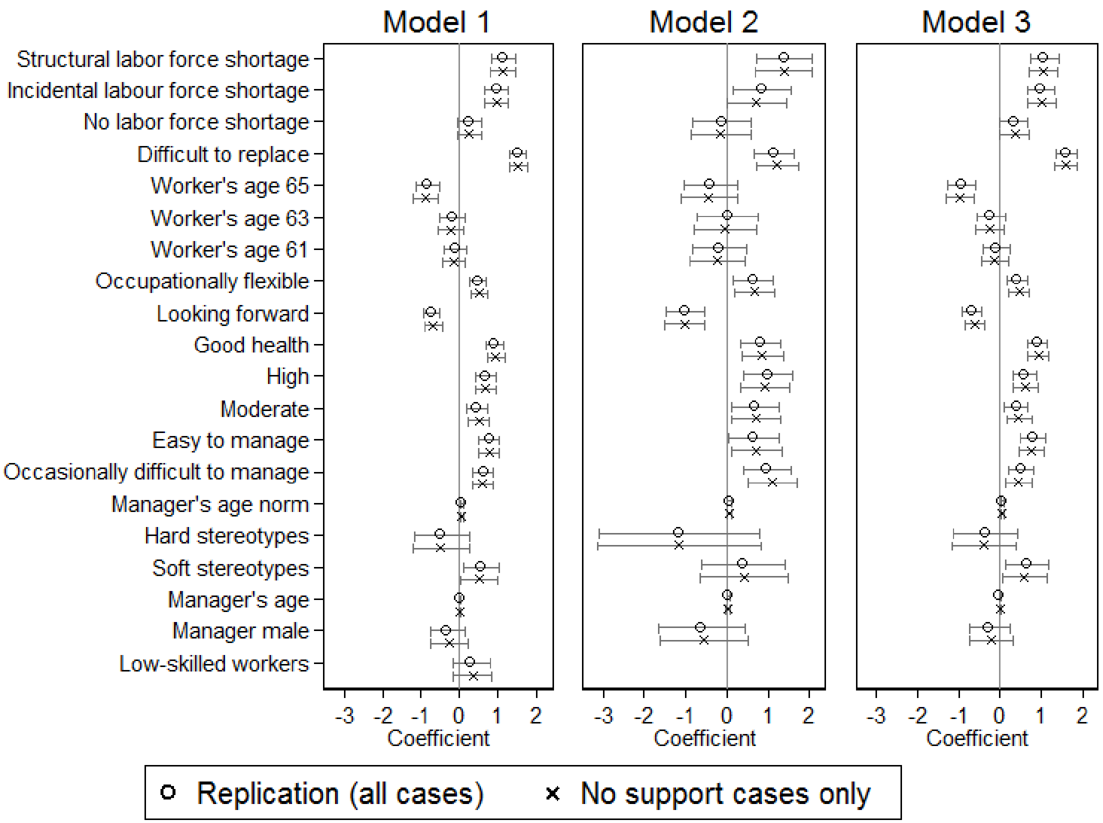
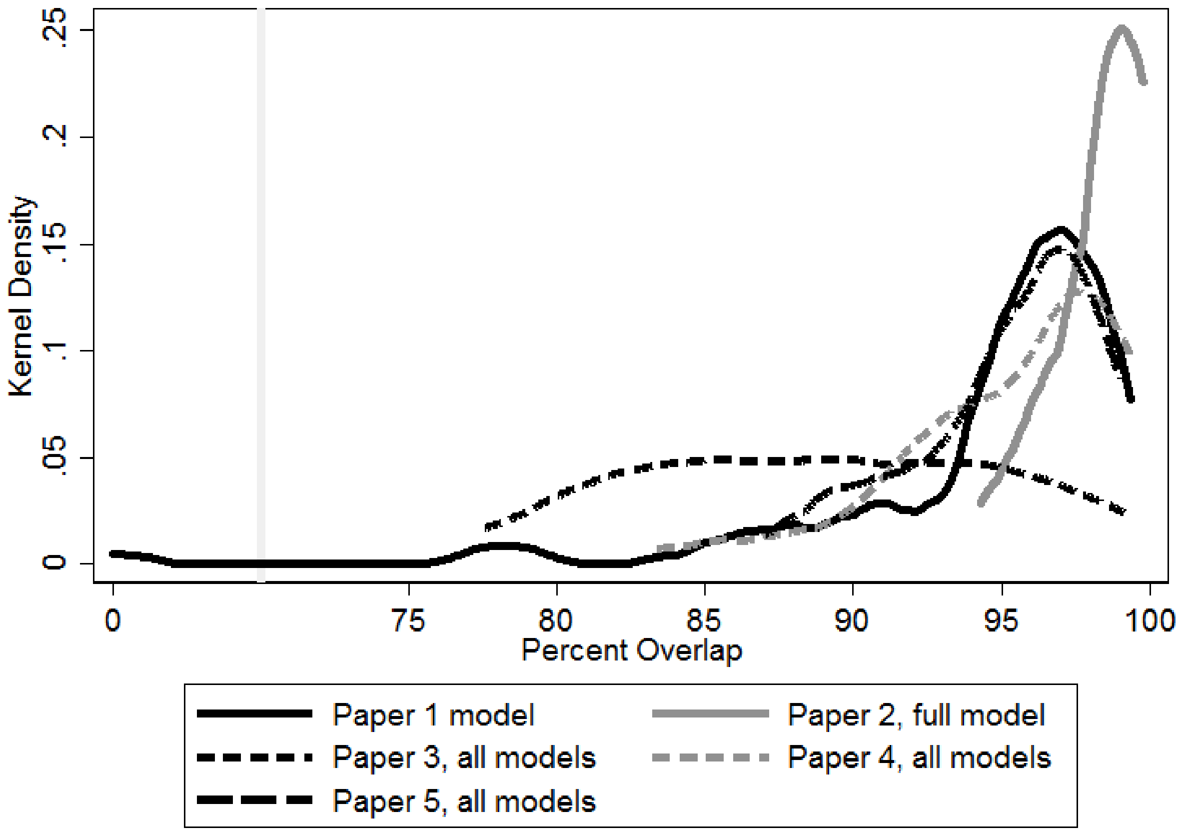


Figure A1: Density of the Overlap of Confidence Intervals on Model Estimates, By Paper

Commented [S1]: To be included as appendix in print version



References

Bosnjak, M., Haas, I., Galesic, M., Kaczmirek, L., Bandilla, W. & Couper, M. P. (2013). Sample Composition Discrepancies in Different Stages of a Probability-Based Online Panel. *Field Methods, 25(4)*, 339-360.

Brendel, M. (2012). Coverage-Bias in Internetpanels. *Master's Thesis at Institute for Statistics, Ludwig Maximilian University*.

Couper, M. P. (2000). Web Surveys: A Review of Issues and Approaches. *Public Opinion Quarterly, 64(4),* 464–494.

Couper, M. P. & Coutts, E. (2006). "Online-Befragungen: Probleme und Chancen Versheidener Arten von Online-Erhebungen." In A. Diekman (Eds.), *Methoden der Sozialforschung* (pp.217-243). Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 44. Wiesbaden: VS Verlag für Sozialwissenschaften.

Couper, M. P., Kapteyn, A., Schonlau, M. & Winter, J. (2007). Noncoverage and Nonresponse in an Internet Survey. *Social Science Research, 36 (1)*, 131-148.

Crutzen, R. & Göritz, A. S. (2011). Does Social Desirability Compromise Self-Reports of Physical Activity in Web-Based Research? *International Journal of Behavioral Nutrition and Physical Activity, 8(1)*, 31.

Dever, J. A., Rafferty, A. & Valliant, R. (2008). Internet Surveys: Can Statistical Adjustment Eliminate Coverage Bias? *Survey Research Methods, 2 (2)*, 47-62.

Eckman, S. (2013). Paradata in Coverage Research. In: F. Kreuter (Eds.), *Improving Surveys Through Paradata: Analytic Uses of Process Information*. New York: John Wiley and Sons.

Eurostat. (2013). "Households with Internet Access at Home" Code: 00088. http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=0&language=en&pcode=tin00088 (accessed August 31, 2014).

Goldberger, A. S. (1981). Linear Regression After Selection. *Journal of Econometrics, 15(3)*, 357‑366.

Groves, R. G. (1989). *Survey Errors and Survey Costs*. New York: Wiley-Interscience.

Groves, R. G. & Couper, M. P. (1998). *Nonresponse in Household Interview Surveys*. New York: Wiley-Interscience.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica, 47(1)*, 153-162.

Hosmer, D. W. & Lemeshow, S. (2000). Applied Logistic Regression (2^{nd} ed). New York: Wiley-Interscience.

Kalmijn, M. (2013). Relationships Between Fathers and Adult Children: The Cumulative Effects of Divorce and Repartnering. *Journal of Family Issues*, 20 (10), 1-23.

Karpinska, K., Henkens, K. & Schippers J. (2013). Retention of Older Workers: Impact of Managers' Age Norms and Stereotypes. *European Sociological Review,* 29 (6), 1323-1335.

Karr, A. F., Kohnen, C., Ognian, A., Reiter, J. P. & Sanil, A. P. (2006). A Framework for Evaluating the Utility of Data Altered to Protect Confidentiality. *The American Statistician, 60(3)*, 224 -232.

Kendall, M. G. & Stuart, A. (1969). *The Advanced Theory of Statistics*. London: Griffin.

Leenheer, J. & Scherpenzeel, A. (2013). Does it Pay Off to Include Non-Internet Households in an Internet Panel? *International Journal of Internet Science, 8 (1)*, 17-29.

Lessler, J. T. & Kalsbeek, W. D. (1992). *Nonsampling Error in Surveys*. New York: John Wiley and Sons.

Robinson, J. P., Neustadtl, A. & Kestnbaum, M. (2002). The Online "Diversity Divide": Public Opinion Differences among Internet Users and Nonusers. *IT & Society*, *1 (1)*, 284-302.

Scherpenzeel, A. & J. Bethlehem (2010), How Representative are Online Panels? Problems of Coverage and Selection and Possible Solutions. In: M. Das, P. Ester & L. Kaczmirek (Eds.), *Social Research and the Internet: Advances in Applied Methods and New Research Strategies*. London: Routledge.

Scherpenzeel, A.C. & Das, M. (2010). True Longitudinal and Probability-Based Internet Panels: Evidence from the Netherlands. In: M. Das, P. Ester & L. Kaczmirek (Eds.), *Social Research and the Internet: Advances in Applied Methods and New Research Strategies*. London: Routledge.

Schonlau M., Van Soest, A. & Kapteyn, A. (2007). Are `Webographic' or Attitudinal Questions Useful for Adjusting Estimates from Web Surveys using Propensity Scoring? *Survey Research Methods, 1 (3)*, 155-163.

Schonlau M., Van Soest, A., Kapteyn, A. & Couper, M.(2009). Selection Bias in Web surveys and the Use of Propensity Scores. *Sociological Methods and Research, 37 (3)*, 291-318.

van der Laan, J. (2009). Representativity of the LISS Panel. CBS, Discussion paper (09041).

van Wilsem, J. (2011). Worlds tied together? Online and Non-Domestic Routine Activities and their Impact on Digital and Traditional Threat Victimization. *European Journal of Criminology, 8 (2)*, 115-127.

van Wilsem, J. (2013). Bought it, but Never Got it’ Assessing Risk Factors for Online Consumer Fraud Victimization. *European Sociological Review, 29 (2)*, 168-178.

Zhang, C., Callegaro, M., Thomas, M. & DiSogra, C. (2009). Do we Hear Different Voices? Investigating the Differences between Internet and Non-Internet Users on Attitudes and Behaviors. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 6063-6076.

Online Appendix A: Replication Results

The following sections and tables present the success of the replication of each papers' summary statistics and model estimates.

A.1 – Details of Replication of van Wilsem (2011)

Replication of the summary statistics table was quite successful for this article. I find slight differences in the Age and Household Income variables, as shown in Table A1.1, which may be due to updates to the LISS data itself. There are more discrepancies in the variables at the bottom of this table: see the rows for Low self-control score, Missing value Low self-control, and Missing value Offender of digital crime. However, these small issues in the table do not affect the models or the author's substantive results.

Model replication was rather successful, with the exception of the coefficient on the Offender of digital crime variable (see the bottom of Table A1.2). Here my coefficients are attenuated towards zero, when compared to what the author reports. This variable has a very low rate of occurrence, which makes replication difficult: The published article reports a mean of 1%, which I have replicated, but I may be off by a few cases, which could impact the estimated coefficient in the models.

For more details on this replication, see Brendel (2012).

Table A1.1: Replication of Summary Statistics, van Wilsem (2011)

Variable	Mean	Std Dev	Min	Max	N
Victim, offline	5.1%	0.22	0	1	6,835
Victim, online	1.0%	0.09	0	1	6,835
Victim, both	1.3%	0.12	0	1	6,835
Women	54 %	0.5	0	1	6,896
Age	45.79	15.74	15	94	6,896
	45.55	*15.75*	*16*		
Education (6 categories)	3.44	1.52	1	6	6,896
Living with partner	80%	0.4	0	1	6,896
HH size	2.79	1.31	1	9	6,896
HH income (1000 €)	3.53	1.31	0	347	6,896
	3.21	*11.41*			
Urbanisation (5 categories)	3.01	1.28	1	5	6,894
How often visit a restaurant (5 point scale)	2.26	0.8	1	5	6,858
How often go to a pub (5 point scale)	2.07	0.99	1	5	6,834
How often meet with friends(5 point scale)	3.03	1.06	1	5	6,814
How often go shopping(5 point scale)	3.12	1.1	1	5	6,820
How often sports activities(5 point scale)	3.13	1.73	1	5	6,693
Work/School (hrs/wk)	23.43	18.36	0	61	6,759
Commuting (hrs/wk)	3.24	4.1	0	21	6,759
E-Mail (hrs/wk)	3.41	5.17	0	30	6,896
Online search for information (hrs/wk)	2.1	2.97	0	20	6,896
Shopping online (hrs/wk)	0.38	0.65	0	4	6,896
Chatting (hrs/wk)	0.94	2.7	0	18	6,896
Online forum (hrs/wk)	0.33	1.03	0	7	6,896
Hyves profile	25%	0.43	0	1	6,890
Using webcam	15%	0.36	0	1	6,855
Low self-control score	1.2	0.16	0	10	6,735
		1.61			
Missing value "Low self-control"	2%	0.27	0	1	6,896
		0.15			
Respondent is offender of digital crime	1%	0.09	0	1	6,358
Missing value "Offender of digital crime"	8%	0.15	0	1	6,896
		0.27			

Replicated values identical, unless shown *in italics*

Table A1.2: Comparison of Estimated Relative Risk Ratios from Multinomial Logit Models: Original; Replication; Replication without Supported Cases

Variables	Traditional only vs. No victimization			Digital only vs. No victimization			Both types vs. No victimization		
	original[a]	repl.	repl.ns[b]	original[a]	repl.	repl.ns[b]	original[a]	repl.	repl.ns[b]
Women	0.90	0.92	0.93	0.84	0.87	0.90	0.94	1.00	1.03
		(0.12)	(0.13)		(0.29)	(0.30)		(0.28)	(0.29)
Age	0.98*	0.98*	0.98*	0.94*	0.94*	0.95*	0.96*	0.96*	0.96*
		(0.01)	(0.01)		(0.02)	(0.2)		(0.01)	(0.01)
Education level	0.96	0.97	0.97	0.87	0.88	0.88	0.88	0.88	0.83
		(0.04)	(0.42)		(0.09)	(0.09)		(0.08)	(0.08)
Living with partner	0.43*	0.47*	0.47*	0.60	0.64	0.63	0.31*	0.37*	0.38*
		(0.08)	(0.09)		(0.29)	(0.29)		(0.13)	(0.13)
HH size	1.17*	1.13*	1.15*	1.06	1.08	1.07	0.95	0.92	0.91
		(0.07)	(0.07)		(0.13)	(0.13)		(0.12)	(0.13)
HH income (1000 €)	1.01	1.01*	1.01*	0.99	0.91	0.91	1.00	1.01	1.01
		(0.00)	(0.00)		(0.10)	(0.10)		(0.01)	(0.01)
Urbanization (5 categories)	0.89*	0.91*	0.90*	1.06	1.06	1.07	1.04	1.04	1.06
		(0.04)	(0.04)		(0.12)	(0.12)		(0.10)	(0.11)
How often to visit a restaurant	1.22*	1.20*	1.22*	0.94	0.96	0.98	1.00	1.00	1.04
		(0.10)	(0.10)		(0.15)	(0.15)		(0.17)	(0.18)
How often to go to a pub	1.04	1.04	1.04	1.06	1.06	1.08	0.96	0.96	0.98
		(0.07)	(0.07)		(0.15)	(0.16)		(0.14)	(0.15)
How often to meet friends	0.90	0.91	0.91	0.78	0.77	0.77	1.16	1.16	1.15
		(0.06)	(0.66)		(0.16)	(0.16)		(0.17)	(0.18)
How often to go shopping	1.01	1.01	1.02	1.40*	1.37	1.33	1.18	1.15	1.21
		(0.06)	(0.07)		(0.23)	(0.22)		(0.17)	(0.19)
How often sports activities	1.14*	1.12*	1.14*	1.16	1.16	1.14	1.05	1.04	1.05
		(0.04)	(0.4)		(0.10)	(0.11)		(0.08)	(0.08)
Work/School (hrs/wk)	1.00	1.00	1.00	1.00	1.01	1.00	1.00	1.00	1.00
		(0.005)	(0.00)		(0.01)	(0.13)		(0.01)	(0.01)
Commuting (hrs/wk)	1.03*	1.03	1.03	1.07*	1.06*	1.07*	1.02	1.02	1.03

		(0.02)	(0.02)		(0.03)	(0.03)		(0.03)	(0.03)
Email (hrs/wk)	0.99	0.99	0.99	0.95	0.95	0.95	1.01	1.01	1.02
		(0.01)	(0.12)		(0.04)	(0.04)		(0.02)	(0.02)
Online information retrieval (hrs/wk)	1.05*	1.04*	1.05*	1.05	1.06	1.06	1.01	1.02	0.98
		(0.02)	(0.02)		(0.04)	(0.04)		(0.04)	(0.05)
Online product purchasing (hrs/wk)	1.06	1.54	1.06	1.33*	1.28	1.26	1.37*	1.33*	1.40*
		(0.10)	(0.10)		(0.02)	(0.18)		(0.14)	(0.14)
Online chatting (hrs/wk)	0.90*	0.91*	0.92*	1.04	1.04	1.04	1.02	1.02	1.00
		(0.03)	(0.03)		(0.04)	(0.4)		(0.03)	(0.03)
Online forums (hrs/wk)	1.03	1.03	1.04	1.09	1.12	1.12	1.22*	1.20*	1.18*
		(0.06)	(0.6)		(0.10)	(0.09)		(0.08)	(0.09)
Hyves profile	1.31*	1.25	1.22*	2.90*	3.10*	3.01*	0.66	0.69	0.70
		(0.20)	(0.20)		(1.11)	(1.08)		(0.19)	(0.19)
Using webcam	1.24	1.23	1.26	1.90*	1.96*	1.98*	2.14*	2.06*	2.14*
		(0.20)	(0.20)		(0.60)	(0.62)		(0.59)	(0.63)
Low self-control score	1.14*	1.13*	1.11*	1.04	1.05	1.04	1.38*	1.36*	1.35*
		(0.04)	(0.04)		(0.08)	(0.08)		(0.08)	(0.08)
Missing value “Low self-control”	0.88	0.99	1.03	1.76	2.16	2.39	0.54	0.80	0.00*
		(0.48)	(0.51)		(1.57)	(1.68)		(0.73)	(0.00)
Respondent is offender of digital crime	4.26*	3.16*	2.71	11.43*	8.49*	5.39*	10.69*	5.67*	4.79*
		(1.55)	(1.50)		(6.00)	(4.50)		(3.20)	(3.10)
Missing value “Offender of digital crime”	1.10	1.08	1.37	3.01	3.41	3.23	0.46	0.43	0.55
		(0.31)	(0.39)		(2.75)	(2.85)		(0.45)	(0.59)
N	6,373	6,370	5,997	6,373	6,370	5,997	6,373	6,370	5,997
Pseudo-R^2	--	0.116	0.113	--	0.116	0.113	--	0.116	0.113

Exponentiated coefficients

* $p < 0.05$

[a] Pseudo-R^2 and standard errors not given in article

[b] Replication without supported cases

A.2 – Details of Replication of van Wilsem (2013)

Replication of this paper's reported summary statistics was entirely successful, aside from the case number. Although the paper reports 6,201 cases in the summary table, the replication found slightly fewer cases, 6,157, as shown in Table A2.1. I am not able to account for this difference, but given that all other summary statistics match, I believe I was working with the right data set.

The models also replicated well, with only small discrepancies in some coefficients, as shown in Table A2.2. Again here, though, I have 6,157 cases and the paper reports 6,201.

For more details on this replication, see Brendel (2012).

Table A2.1: Replication of Summary Statistics, van Wilsem (2013)

Variable	Mean	Std Dev[a]	Minimum	Maximum	N
Fraud Victim	0.02	--	0	1	6,201
		0.16			*6,157*
Women	0.53	--	0	1	6,201
		0.50			*6,157*
Age	45.73	15.20	15	94	6,201
	44.39	*15.18*	*16*	*89*	*6,157*
Education (6 categories)	3.54	1.50	1	6	6,201
	3.55				*6,157*
HH size	3.18	1.33	1	9	6,201
	2.82				*6,157*
HH income (1000 €)	3.18	12.82	0	347	6,201
		11.07			*6,157*
Living with partner	0.80	--	0	1	6,201
		0.40			*6,157*
Urbanization (5 categories)	3.04	1.28	1	5	6,201
	3.00				*6,157*
Shopping online (hrs/wk)	0.40	0.67	0	4	6,201
	0.41				*6,157*
Online forum (hrs/wk)	0.35	1.08	0	7	6,201
		1.06			*6,157*
Hyves profile	0.27	--	0	1	6,201
		0.44			*6,157*
Facebook	0.02	--	0	1	6,201
		0.12			*6,157*
Other Networking Site	0.04	--	0	1	6,201
		0.20			*6,157*
Using webcam	0.16	--	0	1	6,201
		0.37			*6,157*
Low self-control score	1.12	0.16	1	2	6,201
					6,157

Replicated values identical, unless shown *in italics*

[a] In published paper, standard deviation not reported for every variable

Table A2.2: Comparison of Estimated Odds Ratios from Logit Models:
Original; Replication; Replication without Supported Cases

Variables	Model 1 Low self-control		Model 2 LSC+ demographics		Model 3 LSC+routine online act		Model 4 Full model		
	original[a]	repl.	original[a]	repl.	original[a]	repl.	original[a]	repl.	repl. ns[b]
Low self-control	10.08*	10.11* (3.68)	7.76*	7.76* (3.05)	6.97*	7.00* (2.69)	7.11*	7.10* (2.84)	7.02* (2.85)
Female	-	-	0.76	0.77 (0.13)	-	-	0.83	0.84 (0.14)	0.87 (0.15)
Age	-	-	0.97*	0.97* (0.01)	-	-	0.98*	0.98* (0.01)	0.98* (0.01)
Education level	-	-	1.19*	1.18* (0.07)	-	-	1.21*	1.20* (0.07)	1.2* (0.07)
Household size	-	-	1.03	1.03 (0.08)	-	-	1.04	1.04 (0.08)	1.04 (0.8)
Household income	-	-	0.98	0.96 (0.06)	-	-	0-98	0.96 (0.06)	0.96 (0.06)
Partner present in household	-	-	1.02	1.04 (0.27)	-	-	0.98	1.00 (0.26)	1.00 (0.26)
Degree of urbanism	-	-	1.06	1.06 (0.07)	-	-	1.09	1.09 (0.07)	1.10 (0.07)
Internet shopping	-	-		-	1.53*	1.53* (0.129)	1.48*	1.48* (0.13)	1.48* (0.13)
Visiting online forums	-	-		-	1.22*	1.22* (0.06)	1.21*	1.21* (0.06)	1.21* (0.06)
Social network site: Hyves	-	-		-	1.24	1.23 (0.23)	1.02	1.01 (0.20)	0.98 (0.20)
Social network site: Facebook	-	-		-	1.89	1.88 (0.92)	1.65	1.67 (0.82)	1.72 (0.85)

Social network site: Other site	-	-		-	0.55	0.55 (0.23)	0.54	0.54 (0.22)	0.54 (0.22)
Webcam	-	-		-	1.18	1.17 (0.24)	1.19	1.18 (0.25)	1.18 (0.25)
N	6,201	6,157	6,201	6,157	6,201	6,157	6,201	6,157	5,879
-2 log likelihood	1,394.8	1,392.6	1,360.7	1,359	1,340.8	1,339.2	1,320.4	1,319.4	1295.4
Nagelkerke R^2	0.026	0.0231	0.053	0.0466	0.068	0.0606	0.083	0.0744	0.0725

Exponentiated coefficients, standard errors in parentheses

* $p < 0.05$

[a] Standard errors not reported in the original article

[b] Replication without supported cases

A.3 – Details of Replication of Crutzen and Görtiz (2011)

Replication of this article, both the summary statistics and the model coefficients, was entirely successful. See Tables A3.1 and A3.2.

Table A3.1: Replication of Summary Statistics, Crutzen and Görtiz (2011)

Variable		Mean	Std Dev	Min.	Max.	N
Age		47.07	15.97	16	95	5,495
Sex	Female	53.98%	0.50	0	1	5,495
Personal net monthly income		1300[a]				5,495
Level of education	Primary school	9.70%				533
	Intermediate secondary education (US: junior high school)	26.44%				1,453
	Higher secondary education/preparatory university education (US: senior high school)	10.99%				604
	Intermediate vocational education (US: junior college)	23.02%				1,265
	Higher vocational education (US: college)	22.27%				1,224
	University	7.57%				416
Social desirability (Marlowe-Crowne Scale)						5,495
IPAQ[b]	Walking	341.84	367.25	0	1260	5,495
	Moderate-intensity activity	298.74	354.79	0	1260	5,495
	Vigorous-intensity activity	157.01	266.79	0	1260	5,495
	Sedentary behavior	2470.85	1364.84	0	6720	5,495
	Total physical activity[c]	3579.16	3525.33	0	19278	5,495

Replicated values identical for all variables

[a] Median

[b] In minutes per week, see Crutzen and Görtiz (2011) for details

[c] In MET-minutes per week: see Crutzen and Görtiz (2011) for details

Table A3.2: Comparison of Estimated Model Coefficients:
Original; Replication; Replication without Supported Cases

Variables	Model 1 Walking			Model 2 Moderate-intensity activity			Model 3 Vigorous-intensity activity			Model 4 Sedentary behavior			Model 5 Total physical activity[b]		
	Orig.[a]	Repl.	Repl.ns[c]	Orig.[a]	Repl.	Repl.ns[c]	Orig.[a]	Repl.	Repl.ns[c]	Orig.[a]	Repl.	Repl.ns[c]	Orig.[a]	Repl.	Repl.ns[c]
Marlowe-Crowne-Scale	0.00	0.003 (0.015)	0.01 (0.02)	0.00	0.00017 (0.017)	0.00 (0.02)	-0.03	-0.028 (0.020)	-0.02 (0.02)	0.02	0.023 (0.014)	0.02 (0.01)	0.00	0.002 (0.014)	0.00 (0.01)
Female	--	0.02 (0.015)	0.03* (0.02)	--	-0.024 (0.016)	-0.03 (0.02)	--	-0.16* (0.020)	-0.16* (0.02)	--	-0.084* (0.014)	-0.09* (0.01)	--	-0.091* (0.014)	-0.09* (0.01)
Age	--	0.03 (0.016)	0.04* (0.02)	--	0.0081 (0.016)	0.01 (0.02)	--	-0.0071 (0.02)	-0.01 (0.02)	--	-0.0082 (0.014)	-0.01 (0.02)	--	-0.070* (0.014)	-0.06* (0.01)
Personal net monthly income	--	0.0019 (0.015)	-0.01 (0.02)	--	0.040* (0.019)	0.03 (0.02)	--	0.02 (0.00)	0.03 (0.02)	--	0.03* (0.00)	0.02 (0.01)	--	0.011 (0.014)	0.00 (0.01)
Education level	--	-0.15* (0.015)	-0.14* (0.02)	--	-0.14* (0.016)	-0.13* (0.02)	--	-0.20* (0.020)	-0.20* (0.02)	--	0.12* (0.014)	0.12* (0.01)	--	-0.14* (0.014)	-0.13* (0.01)
N	--	4,378	4,036	--	3,715	3,438	--	2,406	2,249	--	5,094	4,682	--	4,953	4,566
R^2		0.026	0.024		0.020	0.017		0.067	0.065		0.025	0.025		0.028	0.027

Standardized coefficients; Standard errors in parentheses

* $p< 0.05$

[a] Number of observations, R^2, and standard errors and coefficients for Female, Age, Net Income and Education and not reported in original article

[b] In MET-minutes/week, see Crutzen and Görtiz (2011) for details

[c] Replication without supported cases

A.4 – Details of Replication of Kalmijn (2013)

Replication of the table of summary statistics for this article was entirely successful, see Table A4.1. Replication of the model coefficients for all eight models was nearly entirely successful. There is a small discrepancy in the estimated coefficient on Child's Age and the intercept in Model 3, and another on Child's Age in Model 5. See Tables A4.2 and A4.3 for details.

Table A4.1: Replication of Summary Statistics, Kalmijn (2013)

Variable	Mean	Std Dev	Minimum	Maximum	N
Contact frequency father (ln)	3.16	1.33	0	5.71	2,328
Quality of tie with father	0	1.00	-2.49	0.92	2,242
Index of support from father	1.93	0.49	1.00	3.00	2,328
Index of support to father	1.92	0.49	1.00	3.00	2,328
Contact frequency mother (ln)	3.32	1.26	0	5.71	2,328
Quality of tie with mother	0	1.00	-2.82	0.86	2,328
Index of support from mother	2.01	0.50	1.00	3.00	2,328
Index of support to mother	2.01	0.48	1.00	3.00	2,328
Father divorced	0.14	0.34	0	1.00	2,330
Father repartnered	0.09	0.29	0	1.00	2,330
Mother repartnered	0.05	0.23	0	1.00	2,328
Age of child at divorce < 18 years	0.34	0.47	0	1.00	318
Father’s age	64.72	8.62	43.00	95.00	2,221
Child higher educated	0.39	0.52	0	9.00	2,302
Child is female	0.58	0.49	0	1.00	2,330
Child’s age	35.68	7.26	18.00	49.00	2,330
Child is single and never married	0.14	0.35	0	1.00	2,310
Child is single and divorced	0.04	0.18	0	1.00	2,310
Child has children	0.61	0.49	0	1.00	2,330
Child has financial problems	0.38	0.72	0	5.00	2,330
Financial module missing	0.24	0.43	0	1.00	2,330
Child poor general health	0.09	0.28	0	1.00	2,330
Health module missing	0.18	0.38	0	1.00	2,330

Replicated values identical for all variables

Table A4.2: Comparison of Estimated Coefficients from Father Models:
Original; Replication; Replication without Supported Cases

Variables	Model 1: Contact frequency with father			Model 2: Perceived quality with father			Model 3: Support from father			Model 4: Support to father		
Variables	Orig.	Repl.	Repl.ns[b]	Orig.	Repl.	Repl.ns[b]	Orig.	Repl.	Repl.ns[b]	Orig.	Repl.	Repl.ns[b]
Father divorced (vs. first married)	- 0.961*	-0.961*	-1.005*	- 0.833*	- 0.833*	-0.848*	- 0.804*	- 0.804*	-0.868*	-0.668*	- 0.668*	-0.731*
	(0.118)	(0.118)	(0.122)	(0.113)	(0.113)	(0.118)	(0.104)	(0.104)	(0.105)	(0.115)	(0.115)	(0.118)
Father repartnered (vs. divorced)	- 0.382*	- 0.382*	-0.328*	- 0.155	- 0.155	-0.142	- 0.332*	- 0.332*	-0.255*	-0.534*	- 0.534*	-0.483*
	(0.125)	(0.125)	(0.128)	(0.125)	(0.125)	(0.129)	(0.119)	(0.119)	(0.122)	(0.123)	(0.123)	(0.127)
Mother repartnered (vs. divorced)	0.265*	0.265*	0.281*	0.135	0.135	0.124	0.308*	0.308*	0.302*	0.343*	0.343*	0.372*
	(0.116)	(0.116)	(0.119)	(0.121)	(0.121)	(0.123)	(0.117)	(0.117)	(0.120)	(0.113)	(0.113)	(0.117)
Child higher educated	- 0.231*	- 0.231*	-0.233*	0.017	0.017	0.007	0.122*	0.122*	0.115*	0.068	0.068	0.068
	(0.054)	(0.054)	(0.056)	(0.041)	(0.041)	(0.042)	(0.035)	(0.035)	(0.035)	(0.036)	(0.036)	(0.036)
Child is female	0.120*	0.120*	0.108*	0.042	0.042 (	0.063	0.102*	0.102*	0.109*	0.120*	0.120*	0.128*
	(0.037)	(0.037)	(0.038)	(0.040)	0.040)	(0.041)	(0.037)	(0.037)	(0.037)	(0.040)	(0.040)	(0.041)
Child's age	- 0.021*	- 0.021*	-0.021*	- 0.019*	- 0.019*	-0.019*	- 0.041*	- 0.040*	-0.040*	-0.009*	- 0.009*	-0.009*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Child single	0.173*	0.173*	0.167*	0.109	0.109	0.152*	0.390*	0.390*	0.408*	0.342*	0.342*	0.375*
	(0.065)	(0.065)	(0.067)	(0.065)	(0.065)	(0.066)	(0.064)	(0.064)	(0.066)	(0.067)	(0.067)	(0.068)
Child single divorced	- 0.093	- 0.093	-0.137	0.035	0.035	0.001	0.145	0.145	0.142	- 0.066	- 0.066	-0.125
	(0.106)	(0.106)	(0.108)	(0.116)	(0.116)	(0.119)	(0.101)	(0.101)	(0.105)	(0.118)	(0.118)	(0.123)
Child has children	0.200*	0.200*	0.187*	0.011	0.011	0.013	0.168*	0.168*	0.167*	0.004	0.004	0.007
	(0.046)	(0.046)	(0.047)	(0.050)	(0.050)	(0.050)	(0.047)	(0.047)	(0.048)	(0.049)	(0.049)	(0.50)
Child money problems	- 0.122*	- 0.122*	-0.125*	- 0.093*	- 0.093*	-0.098*	- 0.074*	- 0.074*	-0.075*	-0.071*	- 0.071*	-0.066*
	(0.029)	(0.029)	(0.030)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.030)	(0.026)	(0.026)	(0.026)
Child poor health	- 0.116	- 0.116	-0.113	- 0.237*	- 0.237*	-0.264*	- 0.127	- 0.127	-0.123	- 0.038	- 0.038	-0.045
	(0.074)	(0.074)	(0.076)	(0.074)	(0.074)	(0.076)	(0.071)	(0.071)	(0.074)	(0.074)	(0.074)	(0.077)
Money module missing	- 0.057	- 0.057	-0.057	- 0.052	- 0.052	-0.055	0.018	0.018	0.018	0.006	0.006	0.00
	(0.045)	(0.045)	(0.046)	(0.049)	(0.049)	(0.049)	(0.046)	(0.046)	(0.047)	(0.048)	(0.048)	(0.049)

Health module missing	- 0.003	- 0.003	-0.013	- 0.051	- 0.051	-0.055	0.005	0.005	0.006	0.039	0.039	0.034
	(0.054)	(0.054)	(0.054)	(0.056)	(0.056)	(0.057)	(0.051)	(0.051)	(0.052)	(0.055)	(0.055)	(0.056)
Intercept	0.862*[a]	0.862*	0.877*	0.807*[a]	0.807 *	0.802*	1.137*[a]	1.317*	1.327*	0.326*[a]	0.326*	0.338*
		(0.113)	(0.115)		(0.115)	(0.117)		(0.108)	(0.111)		(0.114)	(0.116)
N	2,280	2,280	2,201	2,212	2,212	2,136	2,280	2,280	2,201	2,280	2,280	2,201
Adjusted R^2	0.185	0.185	0.186	0.113	0.113	0.117	0.188	0.188	0.190	0.119	0.119	0.123

All variables standardized, Standard errors in parentheses

* $p < 0.05$

[a] Standard errors not reported in original article

[b] Replication without supported cases

Table A4.3: Comparison of Estimated Coefficients from Mother Models:
Original; Replication; Replication without Supported Cases

Variables	Model 5: Contact frequency with mother			Model 6: Perceived quality with mother			Model 7: Support from mother			Model 8: Support to mother		
	Orig.	Repl.	Repl.ns[b]	Orig.	Repl.	Repl.ns[b]	Orig.	Repl.	Repl.ns[b]	Orig.	Repl.	Repl.ns[b]
Mother divorced (vs. first married)	- 0.365*	- 0.365*	-0.375*	- 0.414*	- 0.414*	-0.388*	- 0.217*	- 0.217*	-0.242*	- 0.001	- 0.001	-0.016
	(0.119)	(0.119)	(0.125)	(0.112)	(0.112)	(0.116)	(0.104)	(0.105)	(0.108)	(0.112)	(0.112)	(0.115))
Mother repartnered (vs. divorced)	- 0.027	- 0.027	-0.038	- 0.097	- 0.097	-0.084	- 0.071	- 0.071	-0.062	-0.284*	-0.284*	-0.243
	(0.135)	(0.135)	(0.134)	(0.132)	(0.132)	(0.134)	(0.119)	(0.128)	(0.130)	(0.128)	(0.128)	(0.130)
Father repartnered (vs. divorced)	0.157	0.157	-0.113	0.034	0.034	0.014	- 0.168	- 0.168	-0.129	- 0.191	- 0.191	-0.180
	(0.134)	(0.134)	(0.137)	(0.132)	(0.132)	(0.136)	(0.117)	(0.127)	(0.131)	(0.128)	(0.128)	(0.130)
Child higher educated	- 0.253*	- 0.253*	-0.260*	0.075*	0.075*	0.066	0.143*	0.143*	0.131*	0.057	0.058	0.053
	(0.058)	(0.058)	(0.060)	(0.037)	(0.037)	(0.038)	(0035)	(0.037)	(0.037)	(0.036)	(0.036)	(0.036)
Child is female	0.206*	0.206*	0.198*	0.074	0.074	0.081	0.207*	0.207*	0.207*	0.270*	0.270*	0.276*
	(0.040)	(0.040)	(0.040)	(0.041)	(0.041)	(0.042)	(0.037)	(0.038)	(0.038)	(0.041)	(0.041)	(0.042)
Child’s age	- 0.023*	- 0.022*	-0.022*	- 0.022*	- 0.022*	-0.023*	- 0.043*	- 0.043*	-0.044*	-0.010*	-0.010*	-0.010*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Child single	0.259*	0.259*	0.255*	0.080	0.080	0.097	0.462*	0.462*	0.460*	0.446*	0.446*	0.458*
	(0.069)	(0.069)	(0.071)	(0.065)	(0.065)	(0.066)	(0.064)	(0.063)	(0.066)	(0.068)	(0.068)	(0.070)
Child single divorced	- 0.081	- 0.081	-0.087	0.076	0.076	0.079	0.167	0.167	0.185	- 0.126	- 0.126	-0.158
	(0.116)	(0.116)	(0.115)	(0.118)	(0.118)	(0.122)	(0.101)	(0.116)	(0.119)	(0.116)	(0.116)	(0.119)
Child has children	0.261*	0.261*	0.245*	- 0.008	- 0.008	-0.011	0.228*	0.228*	0.223*	0.013	0.013	0.008
	(0.048)	(0.048)	(0.049)	(0.050)	(0.050)	(0.050)	(0.047)	(0.048)	(0.049)	(0.051)	(0.051)	(0.051)
Child money problems	- 0.115*	- 0.115*	-0.124*	- 0.082*	- 0.082*	-0.092*	- 0.033	- 0.033	-0.040	- 0.044	- 0.044	-0.044
	(0.032)	(0.032)	(0.033)	(0.032)	(0.033)	(0.033)	(0.029)	(0.031)	(0.031)	(0.029)	(0.029)	(0.030)
Child poor health	0.022	0.022	0.003	- 0.221*	- 0.221*	-0.247*	- 0.049	- 0.049	-0.067	- 0.049	- 0.049	-0.065
	(0.079)	(0.079)	(0.080)	(0.078)	(0.078)	(0.081)	(0.071)	(0.075)	(0.078)	(0.076)	(0.076)	(0.080)
Money module missing	- 0.006	- 0.006	0.006	- 0.001	- 0.001	0.000	0.042	0.043	0.057	0.064	0.064	0.070
	(0.048)	(0.048)	(0.049)	(0.049)	(0.049)	(0.050))	(0.046)	(0.048)	(0.049)	(0.049)	(0.049)	(0.049)
Health module missing	- 0.005	- 0.005	-0.014	- 0.099^{+}	- 0.099	-0.100	- 0.026	- 0.026	-0.031	- 0.011	- 0.011	-0.009
	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)	(0.051)	(0.053)	(0.054)	(0.056)	(0.056)	(0.057)

Intercept	0.691*[a]	0.691*	0.720*	0.832*[a]	0.832*	0.0859*	1.207*[a]	1.207*	1.255*	0.159 [a]	0.159	0.183
		(0.119)	(0.121)		(0.115)	(0.117)		(0.109)	(0.110)		(0.114)	(0.117)
N	2,280	2,280	2,201	2,280	2,280	2,201	2,280	2,280	2,201	2,280	2,280	2,201
Adjusted R^2	0.081	0.081	0.082	0.053	0.053	0.056	0.136	0.136	0.138	0.057	0.057	0.058

All variables standardized, Standard errors in parentheses

* $p < 0.05$

[a] Standard errors not reported in original article

[b] Replication without supported cases

A.5 – Details of Replication of Karpinska et al (2013)

Replication of this paper's reported summary statistics encountered a few difficulties, as shown in Table A5.1. However, not all of these variables appear in the models, which replicated well, with only small discrepancies in some coefficients, as shown in Table A5.2.

Table A5.1: Replication of Summary Statistics, Karpinska et al (2013)

Variable		Mean	Std Dev	Minimum	Maximum	N
Retention recommendation	Total sample	5.09	2.52	1	11	238
Dispositions						
Stereotypes – relative scores	Hard qualities – total sample	0.68	0.27	0.33 *0.29*	3.5 *2.3*	238
	Soft qualities – total sample	1.16	0.41	0.28	2.33	238
Age norm	Total sample	64.5	6.04	50	80	238
Characteristics of respondents						
Managerial position	Higher supervisory positions	41.5		0	1	-- *99*
	Intermediate supervisory or commercial positions	41.5		0	1	-- *99*
	Supervisory manual workers	16.8		0	1	-- *40*
Age (years)		45.3	9.4	24	65	238
Male		76.4	0.42	0	1	238
Education (years)		15.5	2.55	8	18	238
		Unable to find appropriate education variable				
Job level of subordinates	Low-skilled	18.9		0	1	238 *45*
	High-skilled	81.1		0	1	238 *193*
Size of organization		375.8 *377.3*	1,048.54 *1,048.76*	10 *1*	10,000	238 *237*
Sector	Industry	26.4		0	1	238 *63*
	Services	38.8		0	1	238 *92*
	Public	34.8		0	1	238 *83*

Replicated values identical, unless shown *in italics*

Table A5.2: Comparison of Estimated Model Coefficients:
Original; Replication; Replication without Supported Cases

Variables	Total workers			Low-skilled			Highly-skilled		
	original	repl.	repl.ns[a]	original	repl.	repl.ns[a]	original	repl.	repl.ns[a]
Organizational context									
Structural labour force shortage	1.154*	1.154*	1.127*	1.396*	1.396*	1.384*	1.082*	1.082*	1.058*
	(7.32)	(7.32)	(6.91)	(4.04)	(4.04)	(3.93)	(6.12)	(6.12)	(5.77)
Incidental labour force shortage	0.970*	0.971*	0.964*	0.849*	0.849*	0.725*	1.003*	1.003*	1.026*
	(6.22)	(6.22)	(6.03)	(2.40)	(2.40)	(2.00)	(5.79)	(5.79)	(5.79)
No labour force shortage	0.254	0.255	0.265	-0.124	-0.124	-0.150	0.335	0.334*	0.362*
	(1.61)	(1.61)	(1.6)	(-0.34)	(-0.35)	(-0.41)	(1.91)	(1.91)	(2.02)
Need for downsizing	-	-	-	-	-	-	-	-	-
Employees' attributes									
Knowledge and experience									
Difficult to replace	1.521*	1.522*	1.524*	1.140*	1.139*	1.227*	1.614*	1.614*	1.600*
	(13.51)	(13.51)	(13.17)	(4.61)	(4.61)	(4.82)	(12.70)	(12.70)	(12.26)
Not difficult to replace	-	-	-	-	-	-	-	-	-
Employee's age (years)									
65	-0.839*	-0.839*	-0.884*	-0.393	-0.394	-0.425	-0.924*	-0.924*	-0.966
	(-5.31)	(-5.32)	(-5.41)	(-1.16)	(-1.16)	(-1.20)	(-5.22)	(-5.22)	(-5.28)
63	-0.187	-0.185	-0.211	0.031	0.032	-0.026	-0.218	-0.218	-0.246
	(-1.17)	(-1.16)	(-1.29)	(0.08)	(0.09)	(-0.07)	(-1.23)	(-1.23)	(-1.35)
61	-0.111	-0.112	-0.145	-0.186	-0.186	-0.231	-0.084	-0.085	-0.116
	(-0.75)	(-0.75)	(-0.94)	(-0.56)	(-0.56)	(-0.68)	(-0.51)	(-0.51)	(-0.68)
59	-	-	-	-	-	-	-	-	-
Occupationally flexible	0.464*	0.464*	0.510*	0.630*	0.630*	0.679*	0.426*	0.426*	0.470*
(reference number)	(4.16)	(4.16)	(4.44)	(2.56)	(2.55)	(2.66)	(3.41)	(3.41)	(3.66)
Attitude towards retirement									

Looking forward	-0.721*	-0.721*	-0.677*	-1.006*	-1.006*	-1.033*	-0.669*	-0.669*	-0.605*
	(-6.52)	(-6.52)	(-5.94)	(-4.14)	(-4.14)	(-4.14)	(-5.41)	(-5.41)	(-4.76)
Not looking forward	-	-	-	-	-	-	-	-	
Good health (reference frail)	0.917*	0.917*	0.951*	0.825*	0.826*	0.870*	0.911*	0.911*	0.945*
	(8.19)	(8.19)	(8.26)	(3.38)	(3.38)	(3.48)	(7.23)	(7.23)	(7.30)
Willingness to participate in Training									
High	0.686*	0.687*	0.680*	0.998*	0.997*	0.916*	0.614*	0.614*	0.617*
	(5.02)	(5.02)	(4.85)	(3.36)	(3.35)	(2.98)	(4.01)	(4.01)	(3.93)
Moderate	0.444*	0.444*	0.507*	0.680*	0.678*	0.702*	0.401*	0.400*	0.468*
	(3.29)	(3.29)	(3.63)	(2.29)	(2.29)	(2.28)	(2.67)	(2.67)	(3.01)
Low	-	-		-	-	-	-	-	-
Managing an employee									
Easy to manage	0.778*	0.778*	0.775*	0.640*	0.641*	0.727*	0.800*	0.800*	0.780*
	(5.66)	(5.67)	(5.47)	(2.05)	(2.05)	(2.27)	(5.23)	(5.23)	(4.94)
Occasionally difficult to manage	0.624*	0.625*	0.602*	0.983*	0.984*	1.113*	0.517*	0.517*	0.461*
	(4.54)	(4.55)	(4.25)	(3.38)	(3.39)	(3.69)	(3.33)	(3.33)	(2.89)
Difficult to manage	-	-		-	-	-	-	-	
Managers' characteristics									
Age norm	0.063*	0.063*	0.0650*	0.053	0.053	0.054	0.066*	0.066*	0.068*
	(3.94)	(3.94)	(4.00)	(1.70)	(1.69)	(1.67)	(3.61)	(3.61)	(3.69)
Hard stereotypes	-0.466	-0.465	-0.486	-1.184	-1.156	-1.146	-0.346	-0.345	-0.375
	(-1.27)	(-1.26)	(-1.30)	(-1.19)	(-1.17)	(-1.12)	(-0.87)	(-0.87)	(-0.94)
Soft stereotypes	0.561*	0.562*	0.516*	0.393	0.394	0.422	0.675*	0.676*	0.611*
	(2.35)	(2.35)	(2.10)	(0.76)	(0.77)	(0.77)	(2.51)	(2.51)	(2.23)
Age	0.006	0.006	0.006	0.037	0.036	0.033	0.000	0.000	0.002
	(0.57)	(0.54)	(0.53)	(1.63)	(1.59)	(1.34)	(0.02)	(0.00)	(0.13)
Male (reference female)	0.312	-0.312	-0.261	0.614	-0.613	-0.552	0.251	-0.251	-0.202
	(1.37)	(-1.37)	(-1-10)	(1.16)	(-1.15)	(-1.00)	(0.99)	(-0.99)	(-0.76)
Low-skilled workers (ref. high-skilled workers)	0.304	0.305	0.349						
	(1.24)	(1.24)	(1.38)						

Constant	-2.286 (-1.87)	-1.642 (-1.44)	-1.793 (-1.55)	-2.496 (-1.00)	-1.228 (-0.55)	-1.226 (-0.53)	-2.324 (-1.77)	-1.82 (-1.39)	-2.034 (-1.54)
Random effects									
Variance level 2	1.54 (0.201)	1.54 (0.202)	1.57 (0.210)	1.37 (0.41)	1.38 (0.42)	1.47 (0.45)	1.54 (0.22)	1.54 (0.22)	1.56 (0.23)
Variance level 1	3.12 (0.143)	3.13 (0.144)	3.14 (0.148)	2.66 (0.28)	2.66 (0.28)	2.66 (0.29)	3.18 (0.16)	3.18 (0.16)	3.19 (0.17)
Model Fit (degrees of freedom)	-2,515.88 (20)	-2,515.14 (20)	-2391.5 (20)	-458.03 (19)	-458.09 (19)	-438.7 (19)	-2,046.62 (19)	-2,046.63 (19)	-1942.1 (19)
No. of vignettes (No. of respondents)	1,190 (238)	1,190 (238)	1,130 (226)	225 (45)	225 (45)	215 (43)	965 (193)	965 (193)	915 (183)

t-scores in parentheses

* $p < 0.05$

[a] Replication without supported cases

Online Appendix B: Bias in Means

The following tables report the means for all variables used in the specified models. Means for all cases used in the replication are given, along with means for the cases in unsupported households only. The estimate of bias is the difference in these two means (see Equation 3 in the main body of the paper). The estimates of absolute relative bias are the absolute value of the bias divided by the mean calculated on all cases (see Equation 4 in the paper).

Table B1: Estimates of Means and Bias in Variables Used in Model in Paper 1

	Mean		n			absolute
	All Cases	Unsupp. Cases	All Cases	Unsupp. Cases	bias	relative bias
Dependent Variable						
No threat	0.9273	0.9263	6370	5997	-0.0010	0.1%
Traditional threat only	0.0509	0.0515	6370	5997	0.0007	1.3%
Digital threat only	0.0091	0.0095	6370	5997	0.0004*	4.4%
Both threats	0.0127	0.0127	6370	5997	0.0000	0.3%
Independent Variables						
Women	0.5312	0.5306	6370	5997	-0.0006	0.1%
Age	45.4339	44.702	6370	5997	-0.7322*	1.6%
Education (6 categories)	3.4791	3.5279	6370	5997	0.0488*	1.4%
Living with partner	0.8061	0.8226	6370	5997	0.0165*	2.0%
HH size	2.7911	2.8496	6370	5997	0.0585*	2.1%
HH income (1000 €)	3.2721	3.3035	6370	5997	0.0314	1.0%
Urbanization (5 categories)	3.0133	3.014	6370	5997	0.0007	0.0%
How often to visit a restaurant	2.2677	2.2801	6370	5997	0.0125*	0.6%
How often to go to a pub	2.0826	2.1021	6370	5997	0.0195*	0.9%
How often to meet friends	3.0421	3.058	6370	5997	0.0160*	0.5%
How often to go shopping	3.1152	3.1034	6370	5997	-0.0118*	0.4%
How often sports activities	3.1463	3.1868	6370	5997	0.0404*	1.3%
Work/ School (hrs/wk)	23.6374	24.303	6370	5997	0.6660*	2.8%
Commuting (hrs/wk)	3.2575	3.3575	6370	5997	0.1001*	3.1%
E-Mail (hrs/wk)	3.4777	3.5979	6370	5997	0.1202*	3.5%
Online search for information (hrs/wk)	2.1351	2.1759	6370	5997	0.0408*	1.9%
Shopping online (hrs/wk)	0.3792	0.3962	6370	5997	0.0170*	4.5%
Chatting (hrs/wk)	0.935	0.9567	6370	5997	0.0216*	2.3%
Online forum (hrs/wk)	0.3296	0.3402	6370	5997	0.0106*	3.2%
Hyves profile	0.254	0.264	6370	5997	0.0100*	3.9%
Using webcam	0.1526	0.1601	6370	5997	0.0075*	4.9%
Low self-control score	1.1757	1.175	6370	5997	-0.0007*	0.1%
Missing value "Low self-control"	0.0154	0.015	6370	5997	-0.0004	2.5%
Respondent is offender of digital crime	0.0072	0.0067	6370	5997	-0.0006	7.6%
Missing value "Offender of digital crime"	0.07	0.0574	6370	5997	-0.0127*	18.1%

* $p < 0.05$

Table B2: Estimates of Means and Bias in Variables Used in Full Model in Paper 2

	Mean		n			absolute
	All Cases	Unsupp. Cases	All Cases	Unsupp. Cases	bias	relative bias
Dependent Variable						
Fraud victim	0.0247	0.0255	6157	5879	0.0008*	3.4%
Independent Variables						
Women	0.529	0.5293	6157	5879	0.0004	0.1%
Age	44.3882	44.008	6157	5879	-0.3807*	0.9%
Education (6 categories)	3.5473	3.5778	6157	5879	0.0305*	0.9%
HH size	2.8269	2.8694	6157	5879	0.0425*	1.5%
HH income (1000 €)	3.1776	3.2033	6157	5879	0.0257	0.8%
Living with partner	0.8043	0.8182	6157	5879	0.0139	1.7%
Urbanization (5 categories)	3.0026	3.0051	6157	5879	0.0025	0.1%
Shopping online (hrs/wk)	0.4066	0.4188	6157	5879	0.0122*	3.0%
Online forum (hrs/wk)	0.3541	0.3618	6157	5879	0.0076*	2.2%
Hyves profile	0.2682	0.275	6157	5879	0.0069*	2.6%
Facebook	0.0153	0.0156	6157	5879	0.0004	2.5%
Other Networking Site	0.0424	0.0432	6157	5879	0.0008	1.9%
Using webcam	0.1624	0.1681	6157	5879	0.0056*	3.5%
Low self-control score	1.1174	1.1171	6157	5879	-0.0004	0.0%

* $p < 0.05$

Table B3: Estimates of Means and Bias in Variables Used in Sedentary Model in Paper 3

	Mean		n			absolute
	All Cases	Unsupp. Cases	All Cases	Unsupp. Cases	bias	relative bias
Dependent Variable						
Sedentary behavior	2547.24	2552.2	5094	4682	4.9210	0.2%
Independent Variables						
Sex : Female	1.5406	1.5406	5094	4682	-0.0001	0.0%
Age	47.2864	46.287	5094	4682	-0.9994*	2.1%
Personal net monthly income	1630.2	1628.3	5094	4682	-1.9143	0.1%
Social desirability (Marlowe-Crowne)	5.8799	5.8921	5094	4682	0.0123*	0.2%
Education (5-point scale)	3.4621	3.5269	5094	4682	0.0648*	1.9%

* $p < 0.05$

Table B4: Estimates of Means and Bias in Variables Used in Model 1 in Paper 4

	Mean		n			absolute
	All Cases	Unsupp. Cases	All Cases	Unsupp. Cases	bias	relative bias
Dependent Variable						
Contact frequency with father	0.0099	0.013	2280	2201	0.0031	31.4%
Independent Variables						
Father divorced	0.1346	0.1327	2280	2201	-0.0020	1.5%
Father repartnered	0.0553	0.055	2280	2201	-0.0003	0.5%
Mother repartnered	0.0886	0.0877	2280	2201	-0.0009	1.0%
Child higher educated	0.3921	0.3921	2280	2201	0.0000	0.0%
Child female	0.5763	0.5766	2280	2201	0.0002	0.0%
Child's age	35.6456	35.724	2280	2201	0.0786*	0.2%
Child single	0.1434	0.1377	2280	2201	-0.0058*	4.0%
Child single / divorced	0.0342	0.0332	2280	2201	-0.0010	3.1%
Child has children	0.6088	0.6124	2280	2201	0.0037	0.6%
Child has money problems	0.3645	0.3612	2280	2201	-0.0033	0.9%
Child in poor health	0.0868	0.0836	2280	2201	-0.0032*	3.7%
Money module missing	0.2355	0.2367	2280	2201	0.0012	0.5%
Health module missing	0.1706	0.169	2280	2201	-0.0016	0.9%

* $p < 0.05$

Table B5: Estimates of Means and Bias in Variables Used in Model 1 in Paper 5

	Mean		n			absolute
	All Cases	Unsupp. Cases	All Cases	Unsupp. Cases	bias	relative bias
Dependent Variable						
Retention recommendation	5.0924	5.085	1190	1130	-0.0075	0.1%
Independent Variables						
Structural labor force shortage	0.2513	0.246	1190	1130	-0.0052	2.1%
Incidental labor force shortage	0.2563	0.2593	1190	1130	0.0030	1.2%
No labor force shortage	0.2454	0.2469	1190	1130	0.0015	0.6%
Need for downsizing	0.2471	0.2478	1190	1130	0.0007	0.3%
Knowledge and experience difficult to replace	0.4941	0.4982	1190	1130	0.0041	0.8%
Knowledge and experience easy to replace	0.5059	0.5018	1190	1130	-0.0041	0.8%
Employee’s age 65 years	0.2286	0.2248	1190	1130	-0.0038	1.7%
Employee’s age 63 years	0.2202	0.2195	1190	1130	-0.0007	0.3%
Employee’s age 61 years	0.2731	0.2761	1190	1130	0.0030	1.1%
Employee’s age 59 years	0.2782	0.2796	1190	1130	0.0015	0.5%
Occupationally flexible	0.5126	0.5088	1190	1130	-0.0038	0.7%
Occupationally not flexible	0.4874	0.4912	1190	1130	0.0038	0.8%
Looking forward to retirement	0.5017	0.5044	1190	1130	0.0027	0.5%
Not looking forward to retirement	0.4983	0.4956	1190	1130	-0.0027	0.6%
Good health	0.4882	0.4903	1190	1130	0.0020	0.4%
Not so good health	0.5118	0.5097	1190	1130	-0.0020	0.4%
High willingness to participate in training	0.3143	0.3195	1190	1130	0.0052	1.6%
Moderate willingness	0.3311	0.3283	1190	1130	-0.0028	0.8%
Low willingness	0.3546	0.3522	1190	1130	-0.0024	0.7%
Easy to manage an employee	0.3336	0.3336	1190	1130	0.0000	0.0%
Occasionally difficult to manage	0.3403	0.3425	1190	1130	0.0021	0.6%
Difficult to manage	0.3261	0.3239	1190	1130	-0.0022	0.7%
Age norm of manager	64.5336	64.491	1190	1130	-0.0425	0.1%
Hard stereotypes	0.6802	0.678	1190	1130	-0.0022	0.3%
Soft stereotypes	1.1637	1.1548	1190	1130	-0.0089*	0.8%
Age of manager	45.2731	45.389	1190	1130	0.1163	0.3%
Male manager	0.7647	0.7743	1190	1130	0.0096*	1.3%
Supervising low-skilled workers	0.1891	0.1903	1190	1130	0.0012	0.6%

* $p < 0.05$