Freie Universität Berlin

On the Fade-away of an Initial Bias in Longitudinal Surveys

Ulrich Rendtel
Juha Alho

School of Business & Economics
Discussion Paper

Economics

2022/4

On the Fade-away of an Initial Bias in Longitudinal Surveys

Ulrich Rendtel*
Juha Alho†‡

February 15, 2022

Abstract

We propose a novel view of selection bias in longitudinal surveys. Such bias may arise from initial nonresponse in a probability sample, or it may be caused by self-selection in an internet survey. A contraction theorem from mathematical demography is used to show that an initial bias can "fade-away" in later panel waves, if the transition laws in the observed sample and the population are identical. Panel attrition is incorporated into the Markovian framework. Extensions to Markov chains of higher order are given, and the limitations of our approach under population heterogeneity are discussed. We use empirical data from a German Labour Market Panel to demonstrate the extend and speed of the fade-away effect. The implications of the new approach on the treatment of nonresponse, and attrition weighting, are discussed.

Keywords: Longitudinal survey, panel survey, internet recruitment, panel attrition, nonresponse bias, self-selection bias, Markov chain, Mover-Stayer model, weak ergodicity.

1 Introduction

In longitudinal surveys the members of an initial sample are observed over some time span. In a panel survey the initial sample members are interviewed in regular time intervals, for example. In many instances the initial sample

*FB Wirtschaftswissenschaft, Freie Universität Berlin, Germany
†Department of Social Research, University Helsinki, Finland
‡Alho's research was partially supported by the Academy of Finland grant 345218

is a probability sample with known inclusion probabilities. However, it may be also obtained from the internet, where users are invited to participate in a survey. For such "River Sampling" there exists no nonresponse by definition (cf. American Association for Public Opinion Research 2010). These online participants are then collected in access panels of persons who agree to answer questionnaires in the future.

The selection of the initial sample may be subject to nonresponse, if a probability sample is intended (cf. Särndal and Lundström 2005). Or, there may be informative self-selection (cf., Baker et al. 2013, Cornesse et al. 2020, Keiding and Louis 2016, Smyk, Tyrowicz and van der Velde 2021). In both cases the estimation of the prevalence of an interesting characteristic in the population may be biased. However, the characteristic of interest may change over time according to some transitions law. If the transition law in the population and in the observed initial sample is the same, the state distributions in the two populations tend to become similar. In particular, under a Markov chain model the eventual prevalences of different states may be deduced from observed state transition probabilities, if they persist over time (cf., Bishop, Fienberg and Holland 1975, ch. 7). But, Rendtel (2005) noticed that this has implications for the study of selection biases. Using register information he demonstrated that the results of the respondent sample and the initial gross sample could converge over time in terms of their income characteristics, even though the samples were subject to initial selection bias. The bias was *reduced* during the follow-up. In Rendtel (2013) the term "fade away" was coined for this phenomenon.

Here we expand these early findings in three directions. First, we consider non time-homogeneous transitions laws; this case is known as *weak ergodicity* in mathematical demography. Second, we discuss panel attrition, the losses of the initial sample after the start of the panel. Third, we investigate Markov chains of higher order, and the effect of population heterogeneity in terms of the Mover-Stayer model (cf., Singer and Spilerman 1974 and Heckman and Singer 1982).

We analyse data from the German Panel on Labour Market and Social Security (PASS) in 2006-2010. This panel was selected from a register of persons in a special labour market program, and we were able to compute the transition laws for the intended initial sample, the gross-sample, and the much smaller realized sample. Despite the considerable initial nonresponse rate of 71.4%, PASS did not exhibit a substantial bias for the labour market status of interest. But, we use the transition and attrition probabilities of PASS in a simulation to display the speed of the fade-away effect. We investigate the effect of attrition, weighting compensation of attrition, and the size of the fade-away effect under a Mover-Stayer model.

In Section 2 we present a contraction theorem for non time-homogeneous Markov chains. For ready perusal, a proof of the result is given in the Appendix. This provides an estimate of the speed of geometric convergence. Section 3 discusses the effect of population heterogeneity and chains of higher order. Section 4 incorporates attrition. We then present the PASS panel in Section 5. After discussing empirical results we display the results of simulation runs with a sizeable initial bias. We display the size and the speed of the fade-away effect with and without attrition and under a Mover-Stayer model. In the final discussion we draw conclusions for an alternative treatments of nonresponse in longitudinal surveys.

2 Dynamic Changes of State

2.1 Time-inhomogeneous Markov Chain

We consider a finite set of states $\mathscr{S} = \{1, \ldots, S\}$. The states could be defined in terms of social or economic characteristics, for example. In the application of Section 5, we will have $S = 2$ and the states relate to unemployment recipiency.

Let Y_t be the state of an individual at time $t = 0, 1, 2\ldots$ The state transitions of the individual are assumed to be Markovian, or

$$P(Y_t = j|Y_{t-1} = i, Y_{t-2} = s_{t-2}, \ldots, Y_0 = s_0) = P(Y_t = j|Y_{t-1} = i) \quad (1)$$

The $S \times S$ matrix of transition probabilities from time $t-1$ to time t is $\mathbf{P}(t) = (p_{ij}(t))$, where $p_{ij}(t) = P(Y_t = j|Y_{t-1} = i)$. In the case of panel surveys t indexes the panel waves. Transition probabilities from time 0 to time t are given by $\mathbf{P}^{(t)} = \mathbf{P}(1)\mathbf{P}(2)\ldots\mathbf{P}(t)$.

2.2 Contraction Theorem

Consider two populations. One population will correspond to our observed sample. Their members change states according to the *same* transition probabilities $p_{ij}(t)$. The initial state distributions of the two populations are the S-dimensional vectors $\boldsymbol{\nu}(0) = (\nu_1(0), \ldots, \nu_S(0))^T$, and $\boldsymbol{\pi}(0) = (\pi_1(0), \ldots, \pi_S(0))^T$. The subsequent *expected* state distributions satisfy the recursions $\boldsymbol{\nu}(t) = \mathbf{P}^T(t)\boldsymbol{\nu}(t-1)$, and $\boldsymbol{\pi}(t) = \mathbf{P}^T(t)\boldsymbol{\pi}(t-1)$ for $t = 1, 2, \ldots$

When all components of $\boldsymbol{\pi}(t)$ are strictly positive, we have the lower and upper bounds

$$m_t \equiv \min_i \frac{\nu_i(t)}{\pi_i(t)} \le \frac{\nu_j(t)}{\pi_j(t)} \le \max_i \frac{\nu_i(t)}{\pi_i(t)} \equiv M_t, \quad (2)$$

for all $j = 1, \ldots, S$. Then the following contraction theorem holds.

Theorem. *Suppose there is a lower bound* $0 < p_L \leq p_{ij}(t)$ *for* $t = 1, 2, \ldots$. *Then* $\boldsymbol{\nu}(t)$ *and* $\boldsymbol{\pi}(t)$ *converge uniformly in the sense that*

$$M_t - m_t \to 0, \quad \text{as} \quad t \to +\infty. \tag{3}$$

A proof can be found in the Appendix.

The result is sometimes called *weak ergodicity.* The regularity conditions can be stated in various ways. Essentially, the chain must be irreducible and aperiodic (see, e.g., Çinlar 1975, for the definitions of these concepts). More general statements are given, e.g., in Le Bras (1977) and Cohen (1979).

An important feature is the speed of convergence to the joint distribution on the state space. In the case of a fixed transition matrix $\mathbf{P}$ there exists a steady state distribution π^*. Then the second largest eigenvalue λ_2 of $\mathbf{P}$ determines the speed of the convergence to the steady state distribution. One can prove $|p_{ij}^{(t)} - \pi_j^*| = O(|\lambda_2|^t)$ for all $i, j \in S$., see Seneta (1980, Theorem 4.2) for a proof.

Example. Let $S = 2$, and consider a Markov chain with transition probabilities of the form

$$\mathbf{P}(t) = \begin{bmatrix} 1 - a(t) & a(t) \\ b(t) & 1 - b(t) \end{bmatrix}, \tag{4}$$

where $0 < a(t) < 1$ and $0 < b(t) < 1$. The spectral decomposition of $\mathbf{P}(t)$ yields $\mathbf{P}(t) = \mathbf{V}(t)\boldsymbol{\Lambda}(t)\mathbf{U}(t)^T$. The dominant eigenvalue is $\lambda_1(t) = 1$, and the corresponding right eigenvector can be taken to be $\mathbf{V}_1(t) = (1, 1)^T$ for all $t = 1, 2, \ldots$. Corresponding to this choice we have $\mathbf{U}_1(t) = (\frac{b(t)}{a(t)+b(t)}, \frac{a(t)}{a(t)+b(t)})^T$. The second eigenvalue is $\lambda_2(t) = 1 - (a(t) + b(t))$. If it happens that $a(t) = c(t)a$ and $b(t) = c(t)b$ for some constants $a > 0, b > 0$, then the eigenvectors do not depend on time t, but $\lambda_2(t)$ does. Now if values $c(t)$ are bounded away from zero for two populations, then the two populations will have the asymptotic state distribution given by $\mathbf{U}_1(t) = (\frac{b}{a+b}, \frac{a}{a+b})^T$. This shows that *two populations with different time-varying* values $c(t)$ can have the same asymptotic state distributions. In the general case, no asymptotic state distribution need to exist, but the two the state distributions still become asymptotically equal.

3 Population Heterogeneity

A naive reading of the contraction theorem suggests that when dynamic changes of state occur, we should expect – contrary to everyday observations!

– similar state distributions in population sub-groups. By extending the setting, we show that the tendency has limitations.

3.1 Markov Case

A simple form of population heterogeneity can be formally described by assuming that, in addition to states, the individuals belong to classes with different transition probabilities. We assume that class membership is permanent, i.e., it does not change over time, but that it is unobservable, or *latent*.

If the contraction theorem applies within each class, then the distribution differences within the classes become negligible. However, it is no longer guaranteed that the marginal differences become negligible.

3.2 Movers and Stayers

An important case with two latent classes is the Mover-Stayer model. In social mobility studies it was observed early on (e.g., Singer and Spilerman 1974, and references therein) that state transitions do not always appear to be Markovian. In the simplest case we would have *movers* who change state as described in Section 2.2, and *stayers* who never change state at all. This is a limiting case of the model discussed above, when the probabilities of state changes converge to zero.

Suppose the fraction of stayers is $0 < q < 1$ of the target population. Then, the matrix of one-step transition probabilities can be written as

$$\mathbf{P}_q = q\mathbf{I} + (1-q)\mathbf{P}_M, \tag{5}$$

where $\mathbf{I}$ is an $S \times S$ identity matrix and $\mathbf{P}_M$ is the transition matrix of the movers (cf., Singer and Spilerman 1974, 372). Unlike in the Markovian case, the t-step transition probabilities are of the form

$$\mathbf{P}_q^{(t)} = q\mathbf{I} + (1-q)\mathbf{P}_M^{(t)}, \tag{6}$$

where $\mathbf{P}_M^{(t)}$ is as defined in Section 2.1. In order that the state distributions of two such populations were to converge, they should have (i) the same value q and (ii) equal state distributions of stayers at $t = 0$.

The presence of stayers or other forms of heterogeneity may be empirically detected, if information about time spent in different states is available on individual level (for an early discussion, see Alho 1990 , and for a more recent one Dudel 2021).

3.3 Markov chains of higher order

A special case of heterogeneity are Markov chains of higher order. Here the transition behaviour depends not only on the previous state but earlier states may also have an impact on the chances which state is reached in the next step. For example, the chances to get out of poverty may deteriorate the longer one has been in poverty before.

This case may be treated by an extension of the state space. For example, for a second order Markov chain the state space becomes $\mathscr{S}^* = \mathscr{S} \times \mathscr{S}$, where the first component indicates the state at time $t-1$ and the second component indicates the state at time t. Of course, it is logically impossible to reach all states $\mathscr{S}^*$ in one transition. For example, we cannot reach the state $(2,2)$ from $(1,1)$ in one step. However, we may step from $(1,1)$ to $(1,2)$ and then from $(1,2)$ to $(2,2)$. To show that the contraction theorem holds, the transition matrix $P_{\mathscr{S}^*}$ over the extended state space $\mathscr{S}^*$ must be strictly positive, or become strictly positive upon multiplication (cf., Appendix).

Note, that in this case the states correspond to vectors of consecutive states. But, this is merely a convention, and the states could just as well be labeled in some other way.

4 Nonresponse in Panel Studies

Let $R_{k,t}$ be the response indicator, or for individual k with $R_{k,t} = 1$ we observe the value of $Y_{k,t}$, but if $R_{k,t} = 0$ we don't. Assume that *attrition occurs independently of state transitions*, or the events $\{Y_{k,t} = j\}$ and $\{R_{k,t} = 1\}$ are independent given the event $\{Y_{k,t-1} = i, R_{k,t-1} = 1\}$. Furthermore, being a respondent is assumed to have no influence on state transitions. This means that $P(Y_{k,t} = j \mid Y_{k,t-1} = i, R_{k,t-1} = 1) = p_{ij}(t)$, where $p_{ij}(t)$ is the (i,j) element of the transition matrix $\mathbf{P}(t)$. Then, we have that

$$P(Y_{k,t} = j, R_{k,t} = 1 \mid Y_{k,t-1} = i, R_{k,t-1} = 1) = r_i(t)p_{ij}(t), \tag{7}$$

where $r_i(t) = P(R_{k,t} = 1 \mid Y_{k,t-1} = i, R_{k,t-1} = 1)$ does depend only on the previous state i . The corresponding *attrition probabilities* are $1 - r_i(t)$. Note that equation (7) is equivalent to assuming that attrition is Missing At Random in the sense of Rubin (1976).

We may estimate the state specific response probability $r_i(t)$ by the empirical response ratio $\hat{r}_i(t) = \sum_j n_{i,j}(t)/n_i(t-1)$, where $n_i(t-1)$ is the number responding in state i at $t-1$, and $n_{ij}(t)$ is the number of them responding in j at t. Let $w_{i,j}(t) = n_j(t-1)/\sum_j n_{i,j}(t-1)$. With these weights we can compute the attrition corrected number of units in state j at wave t

recursively, as

$$\tilde{Z}_j(t) = \sum_{i=1}^{S} \tilde{Z}_i(t-1) w_{i,j}(t)/\hat{r}_i(t). \tag{8}$$

5 Application to Unemployment Benefits

5.1 PASS: Aims and Data

The survey PASS started in 2006, with some 12,500 households interviewed (cf., Trappman et al. 2013). The acronym stands for the German Panel Arbeitsmarkt and Soziale Sicherung (Panel Labour market and social security). It had broad aims of providing income information for social security purposes in Germany. The specific subpanel we consider involves a follow-up of a cohort of individuals who were unemployment recipients in 2006. The recipiency rules had been changed in a major way in 2005 with the aim of activating the unemployed as job-seekers. This was part of the so-called *Hartz Reforms*, named after the chairman of the committee that made the proposals. A new means-tested benefit scheme, so-called *Unemployment Benefit II*, or $UBII$ was introduced. The primary goal of the follow-up was to see if the new rules worked, i.e., that the prevalence of the recipiency in the cohort would go down. Our focus will be on the changes in the sample estimates of the prevalence $\boldsymbol{\pi}(t)$, based on an initial sample of size 23,773.

The key to our analyses is that the recipient cohort could be linked with the so-called Integrated Employment Biographies register maintained by the German Federal Employment Agency (http://fdz.iab.de/en.aspx). This second, independent source of unemployment recipiency information was available from January 2005 until December 2011, permitting an analysis for the five panel waves 2006-2010.

In PASS we will have benefit recipients and nonrecipients, or $S = 2$. In estimation, we take the point of view of a researcher who does not have access to the register information.

5.2 Initial nonresponse, Attrition and Transitions

Many persons in the target population of PASS had a low education or migration background, so numerous measures to reduce nonresponse were implemented. Despite these efforts, only 6,798 persons responded out of the 23,773 benefit recipients initially selected, or a disappointing 28.6 %. Attrition over time is detailed in Table 1. The increase from $t = 1$ to $t = 2$ was due to refusal conversion.

Table 1: Size of Responding Sample, and the Fraction of Initial Respondents Remaining in the Panel.

t	Sample Size	Fraction Remaining (%)
0	6,798	100.0
1	3,468	51.0
2	3,665	53.9
3	2,697	39.7
4	2,257	33.2

Top panel of Table 2 displays the estimated transition matrices for the respondents. This would be available for a researcher. The bottom panel has the same information, based on registry data, for the nonrespondents. There is little difference between the two groups. Or, in this case the transition probabilities of the sample and the target population are clearly approximately equal.

Focus on the top panel of Table 2. Let us take the transition matrices presented there as estimates $\hat{\mathbf{P}}(t), t = 1, 2, 3, 4$ of (4), with elements dependent on $\hat{a}(t)$ and $\hat{b}(t)$. The second eigenvalues of the estimated transition matrices are $\hat{\lambda}_2(t) = 1 - (\hat{a}(t) + \hat{b}(t))$, and the estimates of the asymptotic probability of $UBII$-recipiency are $\hat{\pi}_1(t) = \hat{b}(t)/(\hat{a}(t) + \hat{b}(t))$. Starting from the latter, for $\hat{\pi}_1(t)$ we get the values 0.53, 0.44, 0.45, 0.36. We note that for any substantive discussion of the success of the *Hartz Reforms* the outlook painted by the consecutive estimates change in a major way. This shows that the consideration of non time-homogeneity is important. The estimates of the second eigenvalues $\hat{\lambda}_2(t)$ are 0.62, 0.68, 0.71, 0.75. This says that the fade away effect is slowing down.

Table 3 indicates that $UBII$-recipients tend to have a bit lower attrition probability than nonrecipients. Or, there is some evidence that $r_i(t)$ depends on i. While the difference is small in waves 1, 3 and 4, the attrition behaviour between recipients and non-recipients amounts 7 percentage points for wave 2, where refusal conversion was practiced. As the topic of the survey was UBII recipience it seems that the refusal conversion was especially successful among the UBII recipients. As recipiency of UBII payments was the main focus of the survey, persons who are still in this state could be more interested in responding (cf., Groves et al. 2000).

Table 4 compares the fraction of $UBII$-recipients for the sample initially selected to the study (line $FULL$), the initial respondent sample (line $INITIAL$) and the responding sample in later panel waves (line $RESP$). Despite an initial nonresponse rate of 71.5 %, the initial bias was only 2.3 percentage points. Comparing the percentages for $FULL$ with those

Table 2: *UBII*-Recipient State (yes, no) and Transition Probabilities for Respondents and Nonrespondents in the Recipient Sample of PASS.

Respondents at	Transition $t \rightarrow t+1$							
$t = 0$	$0 \rightarrow 1$		$1 \rightarrow 2$		$2 \rightarrow 3$		$3 \rightarrow 4$	
	yes	no	yes	no	yes	no	yes	no
yes	0.82	0.18	0.82	0.18	0.84	0.16	0.84	0.16
no	0.20	0.80	0.14	0.86	0.13	0.87	0.09	0.91
Nonrespondents at	Transition $t \rightarrow t+1$							
$t = 0$	$0 \rightarrow 1$		$1 \rightarrow 2$		$2 \rightarrow 3$		$3 \rightarrow 4$	
	yes	no	yes	no	yes	no	yes	no
yes	0.82	0.18	0.83	0.17	0.85	0.15	0.83	0.17
no	0.19	0.81	0.15	0.85	0.14	0.86	0.10	0.90

Table 3: Attrition Probability at $t + 1$ Conditional on *UBII*-Participation at t.

Transition $t \rightarrow t+1$	*UBI*I at t	Attrition Probability
$0 \rightarrow 1$	yes	0.460
	no	0.475
$1 \rightarrow 2$	yes	0.342
	no	0.410
$2 \rightarrow 3$	yes	0.282
	no	0.284
$3 \rightarrow 4$	yes	0.286
	no	0.309

Table 4: *UBII*-Recipients Among the Complete Initial Sample (*FULL*), the Initial Responding Sample (*INITIAL*) and the Responding Sample (*RESP*).

Sample	t=0	t=1	t=2	t=3	t=4
FULL	0.790	0.689	0.616	0.576	0.523
INITIAL	0.813	0.706	0.619	0.572	0.519
RESP	0.813	0.732	0.636	0.602	0.548

for $INITIAL$ we see that the over-representation of persons with UBII-payments fell from 2.3 at the start of the panel to 0.4 four waves later. This is the pure fade-away effect without attrition. The reduction was, however, compensated by panel attrition, since UBII recipients had somewhat higher probability to stay in. Therefore, comparing $FULL$ to $RESP$, the bias remained at about 2.5 percentage points.

The initial bias in the PASS survey is too small to display the full power of the contraction theorem. Therefore we demonstrate below the fade-away effect by means of a simulation study with the same transition and attrition probabilities as in the PASS.

5.3 Longitudinal profiles

It is plausible that persons with longer periods in UBII have a lower chance to escape this state. Therefore Table 5 displays the empirical transition probabilities of the PASS sample for longitudinal profiles of length two. As before, transition of the respondents and nonrespondents are similar, and omitted here for brevity. From Table 5 we read for the transition 0/1 to 1/2 that the probability to stay in the UBII state is 0.83 if one has been the last two timepoints in UBII while it is only 0.70, if one comes from the state (no,yes). This difference is displayed also for the transitions 1/2 to 2/3 and 2/3 to 3/4. If we compute the second eigenvalue for these transition matrices we get the values 0.75, 0.79 and 0.81. Like in the Markov first order case the speed of convergence slows down.

5.4 A simulation study in the PASS framework

We start our simulations with a sample size of 6,798, like the PASS at time $t = 0$. However, we assume a substantial (counterfactual) over-representation of UBII recipients at level 0.95, which is 16 percentage points over the true level. As the receipt of UBII payments is an important topic of the survey, persons who are still in this state at the start of the survey may be more motivated to participate than persons who have already escaped this state.

We use the transition matrices of Table 2 and the attrition probabilities of Table 3 to generate the states and response indicators at times $t = 1, 2, 3, 4$. The UBII percentages on the basis of the original intended sample (line $FULL$ in Table 4) are displayed as "True" prevalence. The simulation run is repeated $R = 100$ times to display the variance due to the stochastic nature of the transitions and the attrition. The displayed means are the means over all replications.

Start UBII	Transition 0/1 to 1/2 UBII yes,yes	yes,no	no,yes	no,no
yes,yes	0.83	0.17	0.00	0.00
yes,no	0.00	0.00	0.19	0.81
no,yes	0.70	0.30	0.00	0.00
no,no	0.00	0.00	0.11	0.89
	Transition 1/2 to 2/3			
Start	UBII			
UBII	yes,yes	yes,no	no,yes	no,no
yes,yes	0.86	0.14	0.00	0.00
yes,no	0.00	0.00	0.21	0.79
no,yes	0.74	0.26	0.00	0.00
no,no	0.00	0.00	0.10	0.90
	Transition 2/3 to 3/4			
Start	UBII			
UBII	yes,yes	yes,no	no,yes	no,no
yes,yes	0.85	0.15	0.00	0.00
yes,no	0.00	0.00	0.20	0.80
no,yes	0.69	0.31	0.00	0.00
no,no	0.00	0.00	0.07	0.93

Table 5: Second order transition matrices on UBII profiles

Figure 1 compares three estimates of the prevalence of UBII with the true prevalence until time $t = 4$. By design, the initial bias at $t = 0$ is 16 %. The upmost line (dotted line with stars) results from the simulated samples with attrition. No correction for losses due to attrition was done here. For this version the initial bias decreased substantially from 16 to 3 percentage points, or by 4/5. The next line is from the sample with no attrition after $t = 0$. It is located somewhat below the first line. Their difference displays the additional bias due to attrition. These differences are small: $t = 4$ they amount only 1.1 percentage points. However, the losses with respect to sample size are substantial: The sample size decreases from 6,798 to 1,197. The difference to the PASS sample size in Table 1 are due to conversions of earlier nonrespondents which is not reflected in the simulation design.

The last of the three lines is given by a sample with attrition, in which the estimated prevalence is obtained by weighting. The weights work almost perfectly in the simulation. There is no practical difference to the prevalence values of the no-attrition case. Here the final bias at $t = 4$ amounts only 1.9 percentage points, which is 1/10 of the initial bias.

Figure 2 displays the 90 percent confidence bands under a binomial model for the initial UBII state. Here we generated UBII at $t = 0$ from a binomial distribution with parameters $n = 6,799$ and $p = 0.95$. For the different UBII starting values we generated the subsequent UBII states and attrition indicators as in the previous simulation runs. For each of the 100 replications with attrition we computed a separate weighting variable. The interval is given by the 5 percent and the 95 percent point over the replications. The length of the intervals matches those from the binomial model. For example, at $t = 4$ we obtain for sample size $n = 1,197$ and $p = 0.55$. The length $2 \times 1.65 \times \sqrt{p(1-p)/n} = 0.046$ can be compared with 0.044 in Figure 2. Despite the substantial (counterfactual) initial bias of 16 percentage points, the true value is already covered by confidence interval at time $t = 4$. This demonstrates again the strength of the fade-away effect.

Above, we discussed the Mover-Stayer model as a departure from the Markov model, for which the contraction theorem holds. In the case of UBII recipients it seems plausible that a certain fraction of this population remains permanently in the state UBII. We will adopt this scenario in our simulation setting in order to study the decline of an initial nonresponse bias under this setting.

The transition matrices for the PASS indicate steady state distributions in UBII in the range of 30 to 40 percent. Thus, a proportion as high as $q = 0.3$ of the stayers might be plausible. As the percentage of UBII recipients at $t = 0$ is 95 percent we assume that the stayer group consists entirely of UBII recipients. For the mover group we have to select transition proba-

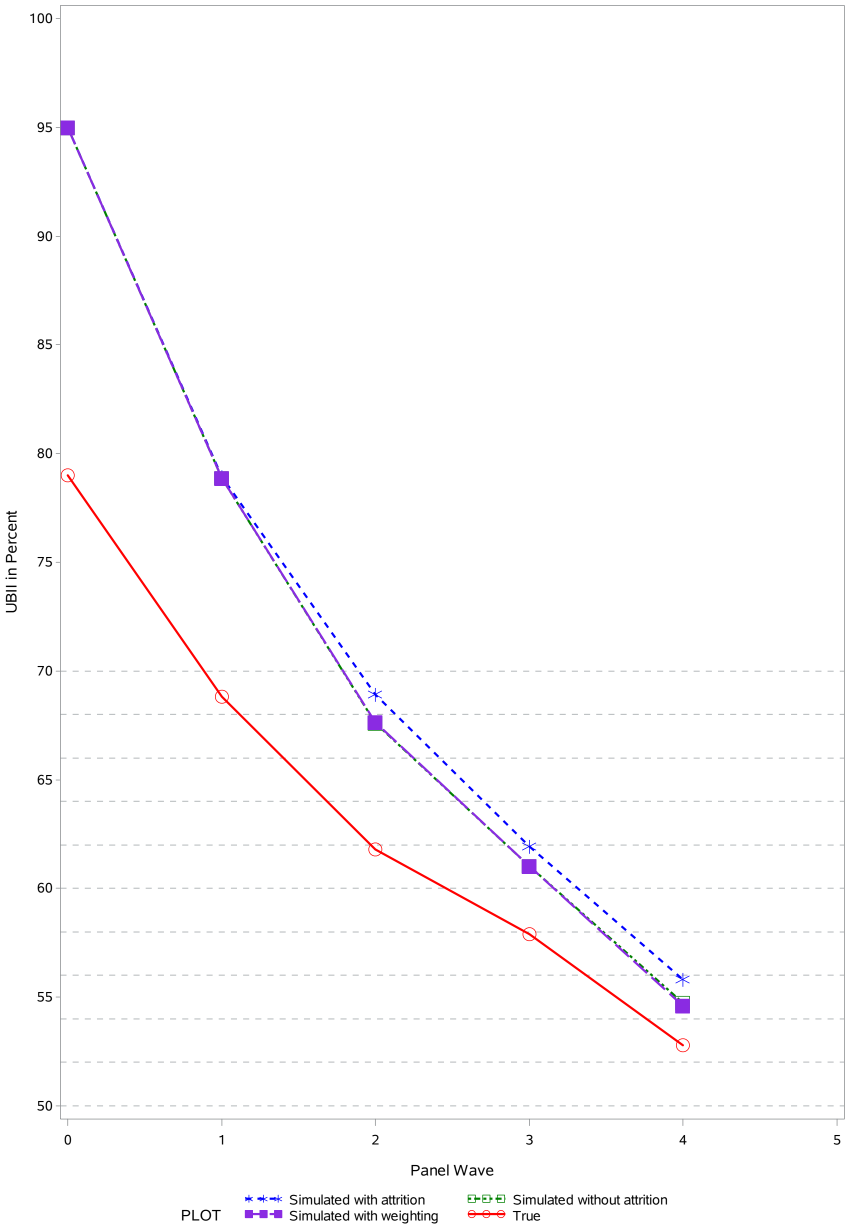

Figure 1: Reduction of the initial bias: simulated prevalences with attrition (Stars), without attrition (Filled Squares) and with attrition and weighting correction (Open Squares) compared with true prevalence (Circles)

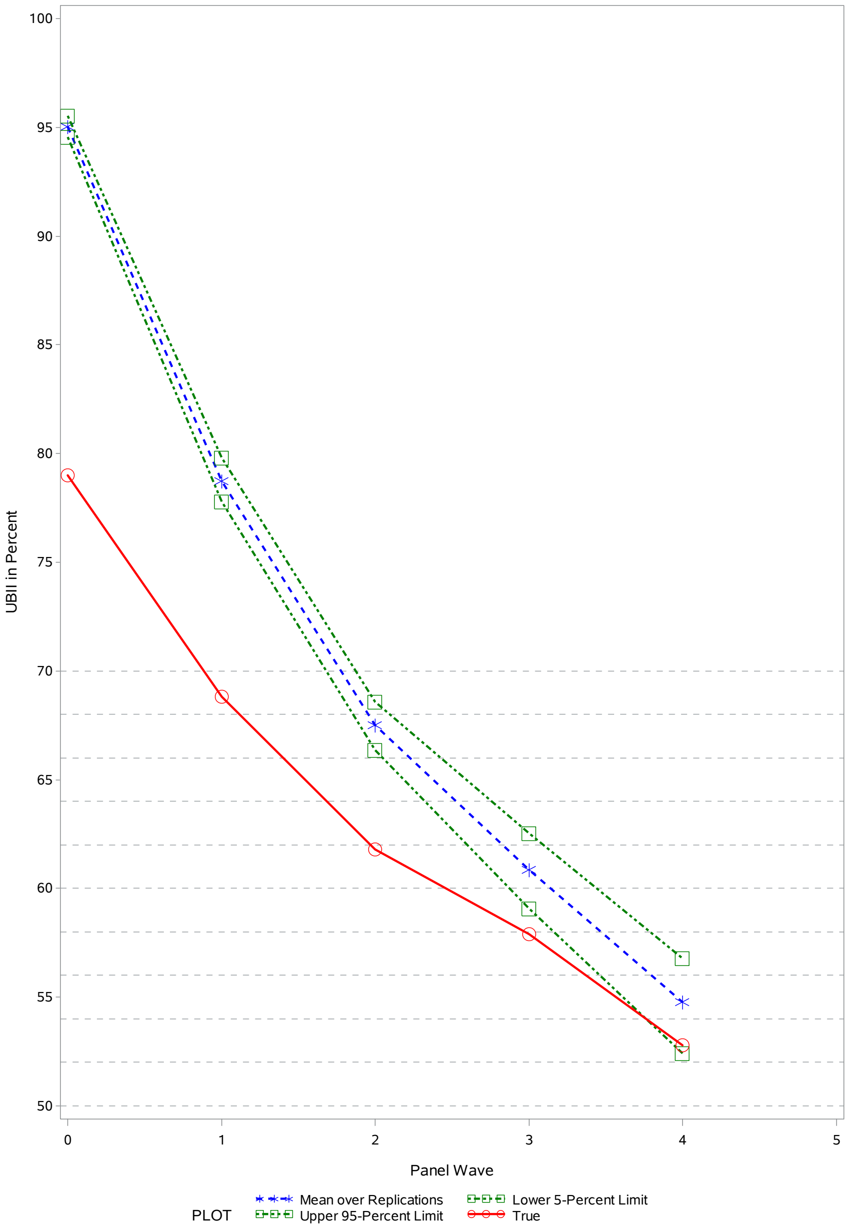

Figure 2: Confidence intervals of prevalence estimates. Simulation runs with attrition and use of weights. Dotted Lines: lower (5 percent point) and upper (95 percent point) limits of 100 simulation runs.

Table 6: $UBII$ State (yes, no) and transition probabilities for movers in the simulation runs.

Respondents at	Transition $t \to t+1$							
$t = 0$	$0 \to 1$		$1 \to 2$		$2 \to 3$		$3 \to 4$	
	yes	no	yes	no	yes	no	yes	no
yes	0.743	0.257	0.743	0.257	0.771	0.229	0.771	0.229
no	0.20	0.80	0.14	0.86	0.13	0.87	0.09	0.91

bilities in equations (5) and (6). These values should be compatible with the marginal transitions from PASS, which fix the left side of equations (5) and (6). The resulting transition probabilities are given in Table 6. The attrition probabilities of Table 3 remain unchanged.

Figure 3 compares the resulting prevalence of both cases: the homogeneous case (label Triangle) and the Mover-Stayer case (label Square). The mover part of the sample, which comprises 70 percent of the initial sample, still leads to a substantive reduction of the initial bias. Note, that the transition probabilities in Table 6 imply a higher transition intensity for the movers than the transitions in the PASS. Until $t = 4$ the initial bias has decreased from 16 percentage points to 8.2 percentage points (Sample with attrition). This is a reduction by 1/2. However, after this point in time there is little further bias reduction, as the lines of the true values and the mover stayer values are almost parallel.

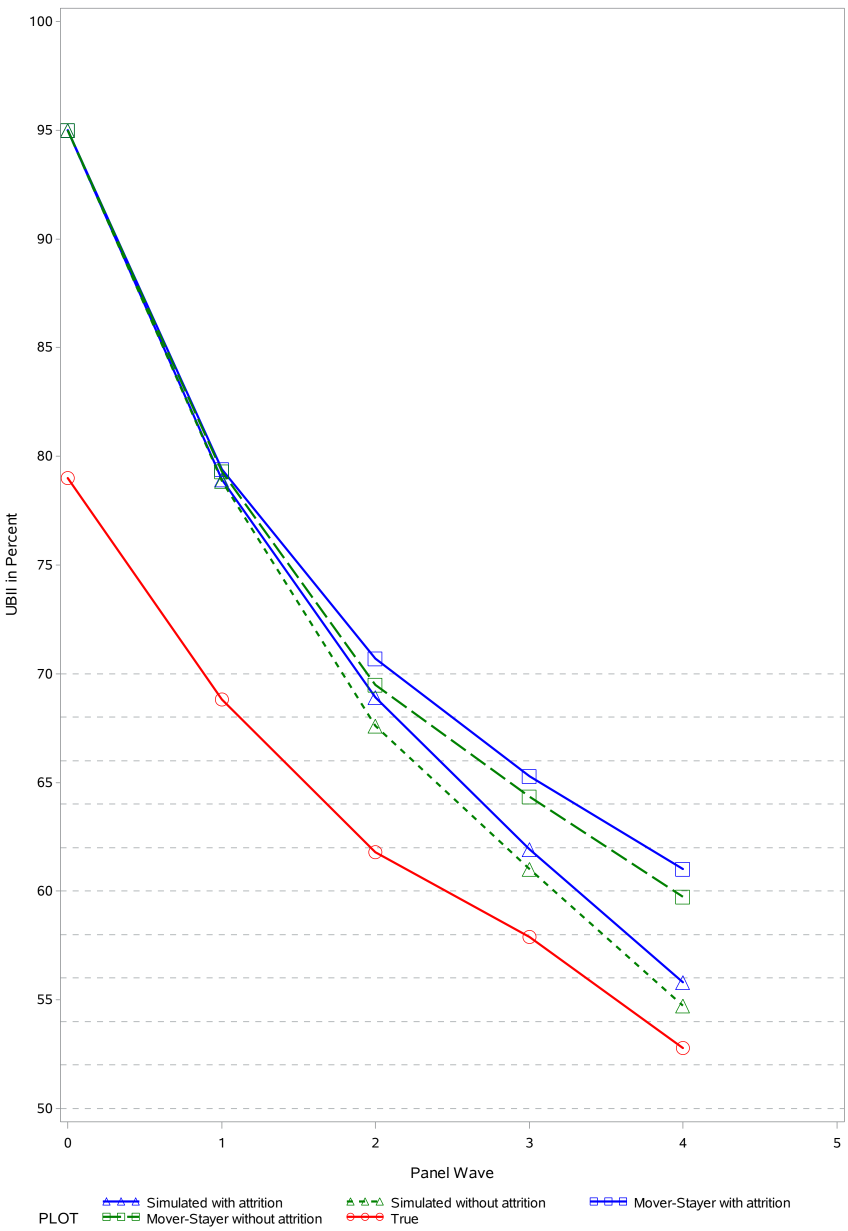

Figure 3: Comparison of prevalence in the homogeneous case (Label Triangle) and the Mover-Stayer model (Label Square)

6 Discussion

The driving force behind the contraction theorem is the frequency of changes between the states. As discussed in Rendtel (2013), transitions between poverty states and income quintiles are examples of situations, in which the effects predicted by the contraction theorem are likely to appear. In our simulations we could show, that an initial bias of a much higher magnitude than a standard error can disappear in a few transitions, and the true value is covered by the confidence interval.

An early hint of the possibility of the fade away phenomenon can be found in Fitzgerald, Gottschalk and Moffitt (1998), who reported that distributions in the Panel Study on Income Dynamics (PSID) became increasingly similar to those from the U.S. microcensus without any additional weighting. The contraction theorem applies, under regularity conditions, to any series of transitions on a finite state space. By extending the state space to Markov chains of higher order we may also show the fade-away of initial bias for longitudinal profiles.

Temporal variability of the characteristic of interest is a relevant aspect in the analysis of nonresponse. Traditionally, one has tried to estimate non-response probabilities (see Särndal and Lundström 2005) in a design-based setting, or to use predictions for the non-observed units in a model-based approach (see Little and Rubin 2002 and Rao and Molina 2015). Here we call upon the users of longitudinal data to have a look on the transition matrices and their second eigenvalues. If these eigenvalues are small, then there may be a substantial fade-away effect of a potential nonresponse bias, which is independent of the mechanism generating the nonresponse at the start of a panel. Besides, it is easy to estimate transition matrices from the observed sample, while it is difficult to derive estimates of response probabilities for the start of a panel. This is due to the simple fact that often, for the nonrespondents, only limited information may be available.

The empirical results from PASS suggest that the receipt of social benefits has only a minor impact on the attrition behaviour. As a consequence, the impact of attrition on the nonresponse bias was small, and could be controlled by weighting with inverse state-specific response rates. This result is in line with the review of literature on panel attrition by Watson and Wooden (2008). They concluded that "income is relatively unimportant for attrition". Instead, variables linked to field work are often good predictors for attrition (Behr et al. 2005). For example, in a panel with personal interview, a change of the interviewer is a risk factor for drop-out (cf., Basic and Rendtel 2007), previous interview experience, such as providing a contact name at the first wave, matters (Laurie, Smith and Scott, 1999), and staying in contact with panel members can be important (Iacovou and Lynn, 2017). In as much as such characteristics are known, they can be accommodated using propensities (Kreuter and Olsen 2011, Trappmann, Gramlich and Mosthaf 2015).

Attrition may decrease case numbers substantially; in our example from 6,798 to 1,197. This is often regarded as an argument to increase the sample size by a so-called refreshment sample, see Deng et al. (2013) and the literature therein. Often the refreshment samples are simply taken to boost the case numbers. But one should keep in mind, that each fresh new sample could incur a fresh bias into the joint sample. Thus it may happen that the estimates of the old panel and the refreshment sample differ systematically. Often such a difference is interpreted as an attrition bias of the old panel. However, there is also the possibility that a bias in the old panel fades away while the refreshment samples suffer from a substantial initial wave bias.

Rotation schemes for panel surveys are an automated way of replacing systematically parts of the old panel by a newly sampled rotation groups. Often there are systematic differences between estimates from different rotation groups, called rotation group bias (see, e.g., Solon 1986). These differences may have different causes like panel conditioning, or selective attrition (cf., Bailar 1975). In our context, these differences may also arise as a consequence of the fade away of bias among the earlier respondents.

Inhomogeneity of transitions rates, which is not explicitly controlled by observed variables can be a drawback of the approach discussed here. The Mover-Stayer model is a worst case scenario in this context. We checked the numerical effect of such a model deviation in our simulations. Because there is no mobility in the stayer group, the mobility in the mover group is to be higher than the marginal mobility. In our simulation setting the contraction within the mover group turned out to be considerably high and resulted in a reduction of the initial bias by 1/2 within the first four waves. However, after this initial period there seems to be no further bias reduction, as the Stayers limit the chances of change.

In order to reduce the effect of inhomogeneous transition laws we should look for observed stable variables which stratify the transition laws. Within these strata, the transitions should be almost homogeneous, so that the contraction theorem holds within each stratum. If we know the stratum sizes in the population then we can estimate the prevalence of the characteristic of interest within the strata using the fade-away effect. In the second stage, we can weight the stratum estimators by stratum sizes. Note, that the argument for the selection of stratum variables is different from classical approaches of design-based sampling. There, the strata variables are chosen in order to reduce the variance of the variable of interest within strata, see, for example, Särndal et al. (1992). Here, we look for variables which homogenize the transition behaviour.

Internet surveys are a fast and cheap means to collect data from the population of internet users. This is especially useful in domains, in which the willingness to respond is low (e.g., Legleye 2018). The range of sampling strategies is broad: selection from a frame of email addresses, stratified sampling from a pool of self-selected volunteers, called access panel, or just River sampling, which asks for participation by an easy to use voting gadget in the internet, see AAPOR (2010), Cornesse et al. (2020), Lehdonvirta et al. (2020) and Valliant and Dever (2011) for examples. While they may involve elements of probability sampling, internet surveys are prone to self-selection bias (Bethlehem 2010). Elliott and Valliant (2017) provide a formal overview of inference for such samples. Our results add a temporal perspective which suggests that for variables with turnover a bias may be reduced, over time. There are numerous evaluation studies which compare the results of online panels with figures from sample surveys or registers, see, for example, Cornesse et al. (2020), Lehdonvirta et al. (2020). However, the comparisons we have seen, have been cross-sectional. It would seem to be fruitful to have assessments that include a longitudinal perspective, for example, from access panels.

As we should safeguard against inhomogeneity in transition behaviour we should use the above mentioned stratification approach with variables whose's population totals are known, such as gender, birth cohort, or parental family status. This should be feasible for many access panels.

The approach we have introduced does not help with the initial bias at the time of study. To address this, one would have to estimate initial response probabilities, or use calibration information in a design-based setting, (cf. Särndal and Lundström 2005 and Särndal 2007). In a model-based setting we can use predictions for the missing observations under the Missing At Random assumption or, if it does not hold, we have to switch to sample selection or pattern mixture models (cf. Little and Rubin 2002). All these approaches rely heavily on assumptions on the response process or the probability law of the outcome variables, or both. Typically, these assumptions cannot be verified from the observed data alone. However, the application of these approaches has implications for the estimation results in later panel waves. For example, a weighting scheme of the initial panel wave should be updated sequentially by attrition weights. Or, in a Bayesian framework we may postulate a prior distribution on the prevalence parameter at time $t = 0$, which is then updated for the next panel waves. If the underlying response model for initial participation is not correctly specified, the resulting prevalence estimates are biased, see Enderle et al. (2013) for the case of an access panel. Here a new issue arises: is it better to trust in the fade-away effect of the initial nonresponse without a model for initial nonresponse ,or should one use a potentially wrong model for nonresponse compensation, which may incur also biases in later panel waves? This question is addressed to future research.

References

American Association for Public Opinion Research AAPOR) (2010), *Report on Online Panels* https://www.aapor.org/Education-Resources/Reports/Report-on-Online-Panels.aspx (Accessed 26.1.2022)

Alho, J. (1990), "Estimating Exposure Time Distributions." *Demography*, 27, 313-321.

Alho, J., and B.D. Spencer (2005), " *Statistical Demography and Forecasting*," Springer Verlag, New-York.

Bailar, B. (1975),"The Effects of Rotation Group Bias on Estimates from Panel Surveys." *Journal of the American Statistical Association* , 70, 23-30.

Baker, R., J. M. Brick, N. A. Bates, M. Battaglia, M. P. Couper, J. A. Dever, K. J. Gile, and R. Tourangeau (2013)," *Report of the AAPOR Task Force on Non-Probability Sampling*." http://www.aapor.org/AM/Template.cfm?Section=Reports1&Template=/CM/ContentDisplay.cfm&ContentID=5963 (Accessed 26.1.2022)

Basic, E., and U. Rendtel (2007),"Assessing the Bias Due to Non-coverage of Residential Movers in the German Microcensus Panel: An Evaluation Using Data from the Socio-Economic Panel," *ASTA Advances in Statistical Analysis*, 91, 311-334.

Behr, A., E. Bellgardt, and U. Rendtel (2005),"Extent and Determinants of Panel Attrition in the European Community Household Panel," *European Sociological Review*, 21, 489-512.

Bethlehem, J. (2010),"Selection Bias in Web Surveys," *International Statistical Review*, 78, 161-188.

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland (1975)," *Discrete Multivariate Analysis*," The MIT Press, Cambridge.

Çinlar, E. (1975)," *Introduction to Stochastic Processes*," Prentice Hall, Englewood Cliffs.

Cohen, J.E. (1979),"Ergodic Theorems in Demography," *Bulletin of the American Mathematical Society*, 1, 275–295.

Cornesse, C., A. Blom, D. Dutwin, J. Krosnick, E. De Leeuw, et al. (2020),"A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research," *Journal Survey Statistics and Methodology* 8, 4-36. https://doi.org/10.1093/jssam/smz041

Deng,Y., D. Hillygus, J. Reiter, Y. Si, and S. Zheng (2013),"Handling Attrition in Longitudinal Studies: The Case for Refreshment Samples," *Statistical Science* 28, 238-256. https://doi.org/10.1214/13-STS414

Dudel, C. (2021),"Expanding the Markov Chain Toolbox: Distributions of Occupation Times and Waiting Times," *Sociological Methods & Research*, 50, 401-428.

Elliott, M.R., and R. Valliant (2017),"Inference for Nonprobability Samples," *Statistical Science*, 32, 249-264.

Enderle, T., R. Münnich, and Ch. Bruch (2013), "On the Impact of Response Patterns on Survey Estimates from Access Panels," *Survey Research Methods* 7, 91-101.

Fitzgerald, J., P. Gottschalk, and R. Moffitt (1998),"An Analysis of Sample Attrition in Panel Data - The Michigan Panel Study of Income Dynamics, " *Journal of Human Resources* 33, 251-299.

Gary, S. (1986), "Effects of Rotation Group Bias on Estimation of Unemployment," *Journal of Business & Economic Statistics*, 4, 105-109. https://doi.org/10.2307/1391392

Groves, R.M., E. Singer, and A. Corning (2000), "Leverage-saliency theory of survey participation," *Public Opinion Quarterly* 64, 299-308.

Heckathorn, D.D. (2007),"Extensions of Respondent-driven Sampling: Analyzing Continuous Variables and Controlling for Differential Recruitment," *Sociological Methodology* 37,151-207.

Iacovou, M., and P. Lynn (2017),"Design and Implementation Issues to Improve the Research Value of the Longitudinal Component of EU-SILC," in: Atkinson A.C., Guio A. and Marlier E. (eds) *Monitoring*

Social Inclusion in Europe. Eurostat, Luxembourg, DOI:10 2785/60152 Pp. 479-495.

Keiding, N., and T.A. Louis (2016), "Perils and Potentials of Self-selected Entry to Epidemiological Studies and Surveys," *Journal of the Royal Statistical Society, A*, 179, 319-376.

Keyfitz, N. (1977), "*Introduction to the Mathematics of Population, with revisions*, Addison-Wesley, Reading, MA.

Kreuter, F., and K. Olson (2011), "Multiple Auxiliary Variables in Nonresponse Adjustment," *Sociological Methods & Research* 40, 311–332.

Laurie, H., R. Smith, and L. Scott (1999), "Strategies for Reducing Nonresponse in a Longitudinal Panel Survey," *Journal of Official Statistics* 15, 269-282.

Legleye, S., G. Charrance, N. Razafindratsima, N. Bajos, A. Bohet, C. Moreau, and the FECOND research Team (2018),"The Use of a Nonprobability Internet Panel to Monitor Sexual and Reproductive Health in the General Population," *Sociological Methods & Research*, 47, 314-348.

Little, R., and D. Rubin (2002),"*Statistical Analysis with Missing Data. Second Edition*," Wiley, New York.

Lehdonvirta, V., A. Oksanen, P. Räsänen, and G. Blank (2020),"Social Media, Web, and Panel Surveys: Using Non-Probability Samples in Social and Policy Research," *Policy & Internet*, 13, 134 - 155. `https://doi.org/10.1002/poi3.238`

Rao, J., and I. Molina (2015), *Small Area Estimation, 2nd Edition*, Wiley, New York.

Rendtel U. (2005), "A Markov Model for the Decline of Initial Nonresponse Bias in a Panel Survey," *Bulletin of the ISI 55th Session, Contributed Papers* Volume LXI, Book 2.

Rendtel U. (2013), "The Fade Away Effect of Initial Nonresponse in Panel Surveys: Empirical Results for EU-SILC," *Eurostat Methodologies and Workingpapers (ISSN 1977-0375) Edition 2013*, doi:10.2785/21863

Rubin, D. (1976), "Inference and Missing Data," *Biometrika* 63, 581- 592.

Särndal, C.-E. (2007), " The Calibration Approach in Survey Theory and Practice, " *Survey Methodology*, 33, 99–119.

Särndal, C.-E., and S. Lundström (2005), *Estimation in Surveys with Nonresponse* . Wiley, New York.

Särndal, C.-E., B. Swensson, and S. Wretman (1992), *Model Assisted Survey Sampling.* Springer, New York.

Seneta, E. (1973), *Non-negative Matrices* Wiley, New York.

Singer, B., and S. Spilerman (1974), "Social Mobility Models for Heterogeneous Populations," Ch. 12 in *Sociological Methodology, 1973-1974*, Jossey-Bass, San Francisco.

Smyk, M., J. Tyrowicz, and van der Velde (2021), "A Cautionary Note on the Reliability of the Online Survey Data: The Case of Wage Indicator, " *Sociological Methods & Research*, 50, 429-464.

Trappmann, M., J. Beste, A. Bethmann, and G. Müller (2013), " The PASS Panel Survey After Six Waves," *Journal for Labour Market Research*, 46, 275–281.

Trappmann, M., T. Gramlich, and A. Mosthaf (2015),"The effect of events between waves on panel attrition," *Survey Research Methods* 9, 31-43.

Valliant, R. and J. Dever (2011), "Estimating Propensity Adjustments for Volunteer Web Surveys," *Sociological Methods and Research* 40, 105 - 137. https://doi.org/10.1177/0049124110392533

Watson, N., and M. Wooden M. (2008), "Identifying Factors Affecting Longitudinal Survey Nonresponse," In: *P. Lynn (ed.) "Methodology of Longitudinal Surveys"*, Wiley, New York. Pp. 157-181.

Appendix

Proof of the Contraction Theorem

The proof below provides details for the Markov chain case, based on sketches in Alho and Spencer (2005), and Le Bras (1977).

Let $\pi_{\mathbf{F}}(\mathbf{t})$ be the distribution on the state space at wave t for the first Markov chain which started as the $FULL$-sample, i.e. the gross sample including respondents and nonrespondents. It's components are denoted by

$\pi_{F,i}(t)$. Similarly $\pi_{\mathbf{R}}(\mathbf{t})$ denotes the distribution on the state space for the second Markov chain for the sample of initial responders.

We assume that $P(t) > 0$, $\pi_{\mathbf{F}}(\mathbf{0}) > 0$, and $\pi_{\mathbf{R}}(\mathbf{0}) > 0$, so for all t the minima m_t and the maxima M_t of the ratios $\pi_{F,i}(t)/\pi_{R,i}(t)$ are well defined. Moreover, we assume that the elements of the transition matrices $P(t)$ are bounded from below by $p_{ij}(t) > p_L > 0$, and as probabilities they satisfy $p_{ij}(t) \leq 1$. All these assumptions can be relaxed but those mathematical details are not central for our applications, and we omit them.

Step 1. The ratios of the vector elements *contract* over time, i.e. for all t we have:

$$m_t \leq \pi_{F,i}(t+1)/\pi_{R,i}(t+1) \leq M_t, \quad i = 1, \ldots, I. \tag{9}$$

Because of $\pi_{F,j}(t+1) = \sum_i p_{i,j}(t)\pi_{F,i}(t)$ we have

$$\frac{\pi_{F,j}(t+1)}{\pi_{R,j}(t+1)} = \sum_i \frac{p_{i,j}(t)\pi_{R,i}(t)}{\sum_h p_{h,j}(t)\pi_{R,h}(t)} \quad \frac{\pi_{F,i}(t)}{\pi_{R,i}(t)}$$

Thus, for all j, the ratios $\pi_{F,j}(t+1)/\pi_{R,j}(t+1)$ are convex combinations of the ratios $\pi_{F,i}(t)/\pi_{R,i}(t)$ with weights

$$w_{i,j}(t) = \frac{p_{i,j}(t)\pi_{R,i}(t)}{\sum_h p_{h,j}(t)\pi_{R,h}(t)}.$$

Therefore it follows that

$$\begin{aligned}
\frac{\pi_{F,j}(t+1)}{\pi_{R,j}(t+1)} &= \sum_i w_{i,j} \frac{\pi_{F,i}(t)}{\pi_{R,i}(t)} \\
&\leq \sum_i w_{i,j} \max_h \frac{\pi_{F,h}(t)}{\pi_{R,h}(t)} \\
&= \max_h \frac{\pi_{F,h}(t)}{\pi_{R,h}(t)} \sum_i w_{i,j} \\
&= M_t
\end{aligned}$$

A similar argument holds for the minima. This proves Equation (9). As a consequence we have $m_t \uparrow$ and $M_t \downarrow$, and the task is to show that the limits are the same.

Step 2. As an intermediate step, we show that the weights are bounded from below,

$$w_{ij}(t) \geq (p_L)^2/I \quad i, j = 1, \ldots, I. \tag{10}$$

We first show that the numerator of the weights is bounded from below:

$$\begin{aligned}
p_{ij}(t)\pi_{R,j}(t) &\geq p_L \pi_{R,j}(t) \\
&= p_L \sum_h p_{hj}(t-1)\pi_{R,h}(t-1) \\
&\geq (p_L)^2 \sum_h \pi_{R,h}(t-1) \\
&= (p_L)^2
\end{aligned}$$

The summands of the denominator can be bounded from above by 1. Hence the sum is smaller than $I \times 1$. From both inequalities equation (10) is proven.

Step 3. To bound the difference $M_t - m_t$, define first adjusted weights $w^*_{ij}(t) = w_{ij}(t) - (p_L)^2/I \geq 0$, according to step 2. Their sum over i is less than 1. We define adjusted ratios for time $t+1$ which use these adjusted weights. The minima and maxima with respect of the weighted ratios are defined as:

$$m^*_{t+1} = \min_j \{\sum_{i=1}^{I} \frac{\pi_{F,i}(t)}{\pi_{R,i}(t)} w^*_{ij}(t)\}$$

and

$$M^*_{t+1} = \max_j \{\sum_{i=1}^{I} \frac{\pi_{F,i}(t)}{\pi_{R,i}(t)} w^*_{ij}(t)\}.$$

Nevertheless, $M_t - m_t = M^*_t - m^*_t$ as the same constant is subtracted from both M_t and m_t. Now we obtain,

$$\begin{aligned}
m^*_{t+1} &= \min_j \{\sum_{i=1}^{I} \frac{\pi_{F,i}(t)}{\pi_{R,i}(t)} w^*_{ij}(t)\} \\
&= \min_j \{\sum_{i=1}^{I} \frac{\pi_{F,i}(t)}{\pi_{R,i}(t)} (w_{ij}(t) - (p_L)^2/I)\} \\
&\geq \sum_{i=1}^{I} \min_j \{\frac{\pi_{F,j}(t)}{\pi_{R,j}(t)}\} (w_{ij}(t) - (p_L)^2/I) \\
&= m_t(1 - (p_L)^2)
\end{aligned}$$

Similarly it is shown that

$$M^*_{t+1} \leq M_t(1 - (p_L)^2))$$

Therefore we obtain the inequalities

$$\begin{aligned} 0 < M_{t+1} - m_{t+1} &\leq (M_t - m_t)(1 - (p_L)^2) \\ &\vdots \\ &\leq (M_0 - m_0)(1 - (p_L)^2)^t \end{aligned}$$

As $(1 - (p_l)^2)^t$ converges to 0 as $t \to +\infty$ we have completed the proof. This also shows that the convergence is geometric, determined by the constant $1 - (p_L)^2 < 1$.

More generally a contraction theorem may be proven, if the matrices $P(t)$ have strictly positive elements in the same positions, there is some $t_0 \geq 1$ such that the product $P(1) \cdots P(t_0)$ is strictly positive, and the positive elements of matrices $P(t)$ are uniformly bounded away from zero.

Diskussionsbeiträge - Fachbereich Wirtschaftswissenschaft - Freie Universität Berlin
Discussion Paper - School of Business & Economics - Freie Universität Berlin

2022 erschienen:

2022/1 AHRENS, Steffen; Ciril BOSCH-ROSA und Thomas MEISSNER: Intertemporal Consumption and Debt Aversion: A Replication and Extension
Economics

2022/2 WOLF, Elias: Estimating Growth at Risk with Skewed Stochastic Volatility Models
Economics

2022/3 GLAUBITZ, Rick; Astrid HARNACK-EBER und Miriam WETTER: The Gender Gap in Lifetime Earnings: the Role of Parenthood
Economics