Scharrer, Nina:

Impacts of the Total Survey Error with a Focus on the Nonresponse Error and its Handling

Masterarbeit, Wintersemester 2016

Fakultät für Mathematik, Informatik und Statistik

Ludwig-Maximilians-Universität München

https://doi.org/10.5282/ubm/epub.38413

LUDWIG-MAXIMILIANS-UNIVERSITY MUNICH

DEPARTMENT OF STATISTICS

Impacts of the Total Survey Error with a Focus on the Nonresponse Error and its Handling

MASTER'S THESIS

Author: Nina Scharrer

Advisor: M.Sc. Julia Plaß

Working Group: Foundations of Statistics and their Applications

Date: Munich, 15.11.2016

Abstract

The aim of survey statistics is to depict the public opinion by using representative samples of the whole population. Therefore, in this field of research it is focused on collecting data with a minimum of errors, which is an important prerequisite for the subsequent steps of data analysis. However, there are many sources of errors having different influence on parameter estimation. Those impacts can be quantified with the help of the Mean Squared Error (MSE), comprising both, systematic and random error. Viewing the different error sources and their impacts leads to the concept of the Total Survey Error (TSE). It contains the different components sampling error, specification error, coverage error, nonresponse error, measurement error, processing error and includes further constraints and theories. One aim of this thesis is to determine the impacts of the different error sources by simulating a realistic data set which orientates on the structure of the Allbus 2014 and applying constructed error models to this data set. This way, direction and magnitude of the different error sources can be evaluated. It is shown that nonresponse error is a major error component of the TSE. Hence, in the second part of this thesis, it is concentrated on different approaches for handling missing data as consequence of nonresponse error. Besides common weighting and imputation methods, likelihood-based approaches, either ignoring or explicitly modeling the missing process, will be introduced, applied and discussed critically especially with respect to their assumptions. The application of the correction methods again is based on simulated data sets including missing values following the different missing mechanisms Missing Completely at Random (MCAR), Missing At Random (MAR) and Missing Not At Random (MNAR). It results that the performances of missing data methods strongly depend on these underlying processes. While in the case of MCAR data simple ad-hoc procedures should be preferred, for MAR or MNAR data more advanced methods are required.

Contents

List of Figures

List of Tables

List of Abbreviations

Allbus	Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (The German General Social Survey)
EM	Expectation Maximization
FIML	Full Information Maximum Likelihood
MCAR	Missing Completely At Random
MAR	Missing At Random
MNAR	Missing Not At Random
MI	Multiple Imputation
ML	Maximum Likelihood
PI	Partial Identification
PMM	Pattern Mixture Model
RCAR	Rounding Completely At Random
RAR	Rounding At Random
RNAR	Rounding Not At Random
SEM	Structural Equation Model
SM	Selection Model
TSE	Total Survey Error

List of Notations

Symbols	Meaning
$\mathbb{E}$	expected value
Var	variance
Var_{within}	within-variance in context of MI
Var_{between}	between-imputation-variance in context of MI
σ	standard deviation
$Bias$	bias
$Bias_{\text{rel}}$	relative bias
MSE	mean squared error
L	likelihood
ℓ	log-likelihood
$\mathbb{P}$	probability distribution
Γ	set of possible probability distributions
	in context of sensitivity analysis
$\mid\ \mid$	absolute value
$\lfloor\ \rfloor$	floor function
$\lceil\ \rfloor$	rounding function

Indices and Superscripts[1]	Meaning
$i = 1, ..., N$	subject of population
$i = 1, ..., n$	subject of sample
$j = 1, ..., k$	variables
$t = 1, ..., T$	iterations

Parameters	Meaning
N	population size
n	sample size
m	number of missing units
k	number of variables or categories
θ	parameter (vector) of interest
$\hat{\theta}$	estimation of parameter (vector) of interest
μ	mean (vector)
β	vector of regression coefficients
b	vector of regression coefficients of selection equation
	in context of SM

ρ	correlation coefficient
ϵ	error term
e	error term of selection equation in context of SM
ξ	unknown parameter (vector) determining missing mechanism
X	data set
X^{obs}	observed part of data set
X^{mis}	missing part of data set
X_j	variable j
X_j^*	latent variable j
x_{ij}	value of unit i in variable X_j
M	missing data indicator (matrix)
Y	dependent variable of linear regression model
p	probability
w	weight
l	number of adjustment cells in context of weighting
h	rounding threshold
z	z-value of standard normal distribution
a	level of accuracy
α	significance level
λ	inverse Mills Ratio
γ	not identified parameter in context of PI
H	identification region in context of PI
η	set of values encompassed in identification region H

[1] superscripts in cases of iterations are written in round brackets

1 Introduction to the Total Survey Error Approach

1.1 Total Survey Error Theory

During the last decades the importance of public opinion has consistently increased in many different fields like politics, marketing or economy. Therefore, survey research, which gathers those desirable information with the help of polls and provides them to the customer, has attracted more and more attention and has been developed further. Not only technology and professionalization especially with respect to statistical understandings and methodology, have evolved, but also the approaches to this field have changed, as is summarized in Biemer and Lyberg (2003). Thus, standardization came to the fore which means that people strive after guidelines for good practice.

In this context, a new way of thinking about the different problems and sources of error in this discipline has developed which can be summarized as the Total Survey Error (TSE) approach. Weisberg (2005) denotes this framework as a paradigm which means a set of concepts including theories, basic assumptions, values and research methods that constitute a way of viewing reality and are commonly accepted by members of the scientific community. Basically, the idea of the TSE approach can be traced back to different literature sources. Neyman (1934) laid the foundation for sampling error in the field of survey methodology, whereas other error sources are ignored in this landmark paper. In contrast to this, Deming (1944) firstly outlined multiple error sources in sample surveys which affect the usefulness of surveys. Over the years, this concept has been developed further (cp. Groves and Lyberg (2010)), until it finally predominated survey research in the 1990s. Thus, TSE illustrates several possible sources of survey error which arise during the process of data collection and analysis and may diminish the accuracy of inferences derived from survey data due to deviation of survey responses from the underlying true value so that consequently the survey quality may be compromised. Survey research then seeks to minimize Total Survey Error which, according to Biemer (2010b), refers to the accumulation of all errors that may arise in the design, collection, processing, and analysis of survey data. Besides, those quantitative aspects, which form the central point of the TSE approach, also constraints like time, costs and ethics are considered that affect the minimization of these errors. Therefore, the TSE approach cannot only be viewed as a framework to understand the difficulties of survey research and to assess the quality of surveys, but also as a planning criterion. Hence, among a set of alternative designs for a given cost and timeliness situation, the one with the smallest Total Survey Error should be chosen, since it maximizes the accuracy of the results which is according to Biemer and Lyberg (2003) one dimension of the multidimensional concept of survey quality.

All those aspects show the important role of the TSE approach in the field of survey research and form the motivation for the TSE approach to be the basis of the following thesis. One main objective of this work is to outline the different problems and consequences that can arise in the different stages of a survey. For this purpose a realistic data set, resembling the Allbus 2014 data set regarding its main features, is simulated and analyses of substantive interest are conducted in **R** 3.0.1 (see R Core Team (2013)). Here, analyses mean the estimation of unknown population parameters, in general abbreviated as θ. This thesis concentrates on the estimation of the income mean μ_{income} and the regression coefficients β of a certain linear regression model. In the next step, different error models are constructed, consciously applied to this data set and consequences are illustrated by comparing the results of the corresponding statistical analyses to those of the underlying basic simulation data set. Researchers often put emphasis on sampling error (cp. Weisberg (2005)) which should be queried, when analyzing the impacts of the different error sources, since they partly have more serious impacts on the survey results compared to the sampling error. It should be made clear that statistics calculated on the collected survey data, can easily be biased and therefore lead to wrong conclusions when consequences of different survey errors are ignored. Therefore, this thesis should be a warning. Since the introduced error sources are commonly arising, when dealing with survey data, a certain expertise in this field is required to identify and handle those errors, which can be gained with the help of this work.

So besides the illustration of the impacts, another objective of the thesis is to provide different ways to handle those problems. Since it is too comprehensive to explain and challenge methods for handling all the different error sources, this part of the thesis focuses only on the nonresponse error as one of the main error sources of the TSE. First, a naive ad-hoc method is presented and applied to the simulated data sets with missing values following different mechanisms. Here it is focused on the aspects that are still useful about this approach. This critical presentation aims to make clear that there are better approaches to deal with missing data and to convince researchers, who still often go back to these older and less complicated methods in practice. Then, established approaches like weighting, imputation procedures and finally modern Maximum-Likelihood approaches are presented and evaluated due to their performances on the same data sets. However, also limitations and critical assumptions are discussed, in order to emphasize that even though there are quite good methods for correcting missing values in survey data in general there are restrictions and moreover uncertainty about the results remains. In order to close the circle back to the TSE approach, the second part of the thesis, concentrating on the nonresponse error, also contains a literature overview about possible attempts to minimize this error and their consequences for other error sources of the TSE.

Thus, the following structure of the thesis results: First, the different error sources are introduced in the following subsection 1.2 as well as a possible measure of quantification in section 1.3. Secondly, the impacts of the different errors are shown based on a simulation in chapter 2, before focusing on the main error component of the nonresponse error in chapter 3. Besides the current state of research concerning minimizing nonresponse (cp. section 3.1), the previously described methods for handling nonresponse error are introduced in section 3.2 and applied (cp. section 3.5) to the simulated data set (cp. section 3.4). The nonresponse chapter then is completed with section 3.3, containing a critical view concerning the assumptions of the applied nonresponse methods, resulting in the approach of partial identification. The following chapter 4 then points out limits of the simulation and presents a few modeling extensions in order to make further statements. All conclusions as well as an outlook for further research are summarized in the last chapter 5.

1.2 Types of Survey Error

In literature sources describing the Total Survey Error (cp. e.g. Weisberg (2005), Groves and Lyberg (2010), Biemer (2010b), Faulbaum (2014)), the different components are not consistently defined, but differ with respect to classification and naming. However, they all share a major division of sampling and nonsampling error, whereas the latter is decomposed in many components. In the following nonsampling error is determined as specification error, coverage error, nonresponse error, measurement error and processing error. Figure (1.1) illustrates this partition before introducing and considering the different sources of error in detail.

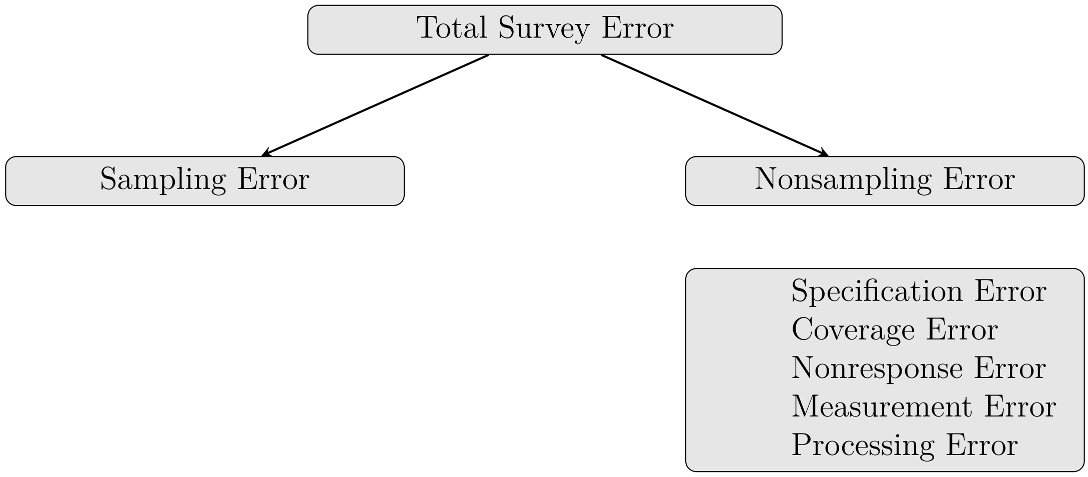

Figure 1.1: Components of the Total Survey Error (graphical representation based on Biemer (2010b))

Sampling Error

Indeed, the best-known and researched source of survey error is sampling error. It arises, when not the whole population can be included in a survey, but only a subset. The sample is called representative, if it is a smaller picture of the population of interest which is the case, when attributes have roughly the same distribution in population and sample. The aim is to generalize the results beyond the people who have been sampled and to be able to make statements which are applicable for the whole population. However, this approach is only possible, when the drawn sample is based on simple random sampling, which means that each observation of the population has the same known probability to be sampled. Since each sample gives a different estimation, which deviates from the true population value, the representativeness of a sample is not guaranteed. Sampling error is often expressed through standard errors, since estimates derived from any sample are subject to sampling variability, which is usually measured as the standard error. The standard error can be received by the following formula (cp. Kauermann and Küchenhoff (2010)):

$$\text{standard error} = \frac{\sigma}{\sqrt{n}} \tag{1}$$

It shows that in general a larger sample size n results in decreased standard errors, which is equivalent to a gain in precision. Thus, larger samples are in general more representative, but this relationship of equation (1) is a square root relationship, which means that quadrupling the sample size only halves the standard error, if the other component σ is held constant. For large populations σ, which describes the standard deviation in the population, can be determined by

$$\sigma \leq \sqrt{\frac{n \cdot a}{z^2_{1-\frac{\alpha}{2}}}}, \tag{2}$$

where $z_{1-\frac{\alpha}{2}}$ is the critical z-value for a two sided test resulting from the corresponding quantile of the standard normal distribution and a is the demanded level of accuracy. For closer interpretations and remarks see Kauermann and Küchenhoff (2010). Even though an increase of the sample size can reduce the uncertainty of the sampling error, it cannot be prevented, since in general the maximal sample size is limited by factors like time and money. Furthermore, there are other factors that have an influence on sampling error. Besides the sample design and the proportion of the sample size and the population size, the heterogeneity is crucial. If there is less variability in the variable within the population, the standard errors are smaller and consequently the sample is more representative. All those aspects show that sampling error plays an important role in the Total Survey Error approach.

Specification Error

The first component of the nonsampling error is called specification error. It describes the problem that information gathered through the survey differ from the concept that should originally have been measured. This definition is obviously closely related to the notation of validity (cp. Winker (2010)). Specification error can lead to invalid inferences, since wrong parameters are estimated by the survey, when the wrong construct is measured. Of course, the interpretation of invalid measurements is doubtful. Since basically useless results are obviously the worst case, specification error is often in the focus of researchers. However, specification error can be quite difficult to detect without the help of subject matter experts. Biemer and Lyberg (2003) provide a good example for this kind of error source. Besides expert knowledge, there are some statistical measures like cronbach's alpha which can hush these fears about specification error. Yet, in order to avoid this scenario in the first place, a good questionnaire design is necessary, wherefore an efficient communication between the researcher and the questionnaire designer is fundamental.

Coverage Error

The second nonsampling error, denoted as coverage error or also often as frame error, is closely related to the previously described sampling error. In the section of sampling error simple random sampling is assumed. Often this assumption is not fulfilled, because some members of the target population are not listed in the sampling frame, whereas others are even duplicated or have another divergent probability to be sampled. The sampling frame is the actual set of units from which the sample is drawn. Thus, in the optimal case, it contains all members of the target population without any duplicates and no elements that should not be included. Unfortunately, these ideals are rarely satisfied.

An example for such a sampling frame are telephone interviews, where households are randomly selected from white pages. Households without phone, cell phone-only households or households that are not listed cannot be taken into account. The elements in the sampling frame do not correspond correctly to the target population to which the researcher wants to make inferences, instead there is an over- or underrepresentation of certain subgroups. If the part of the population, not included in the list, is different on key features of interest, the results are biased. Statistics calculated for the underlying target population differ from the statistical estimates of the drawn sample. This phenomenon is referred to as coverage error. Besides the degree of difference between the observations of the sampling frame and those that are not included, the extent of coverage error obviously also depends on the proportion of those subpopulations. None of the different types of surveys, including telephone interviews, internet surveys and face-to-face interviews, are spared from this error.

There are several studies, trying to find distinctive characteristics of participants and non-participants, in order to draw conclusions about the overrepresentation of different groups. For instance the European Health Examination Survey provides an overview of literature concerning this topic (see Homepage of European Health Examination Survey (2016)). Findings differ, but nevertheless trends about who is more likely to be considered in surveys, exist. In general more educated, more affluent and younger people are often overrepresented in surveys, as well as women and white people (cp. Smith (2008)). This knowledge will be used in section 2.2 to simulate coverage error.

Nonresponse Error

After having obtained a correctly specified sample of the target population, another common source of error can occur that bears upon the refusal of responses of the selected units. In this case data are missing. This so called nonresponse error comprises both, unit and item nonresponse. Unit nonresponse describes the fact that

it is not possible to interview an intended respondent. This can be due to a complete refusal of the individual to take part in the interview or simply due to missing accessibility. In contrast to that, item nonresponse error occurs, when the questionnaire is only partially completed. This comprises interviews that are prematurely terminated and interviews with questions that are skipped or left blank. Latter case usually occurs, if a question is sensitive, like questions about income. But of course there are several other reasons why people refuse to give answers (cp. Graham (2012)). There can be respondent burdens, which means that the responders simply do not know the answer or do not understand the question. But data can also get lost during the collection stage or due to a mechanical breakdown.
Anyhow, in present times, characterized by a large number of polls and surveys, the presence of nonresponse error rises, which results in a mass of missing data and low response rates. However, response rates per se are not suitable indicators for representative and credible samples, since this measure does not give information about the degree of dissimilarity between the responders and nonresponders. Nonresponse error has its effect through two components, the nonresponse rate $\frac{n-m}{n}$ and the difference between nonresponders and responders in the survey concerning the variable of interest (cp. Groves (1998)). Thus, nonresponse becomes a problem, when the nonresponders differ from reponders systematically regarding key measures. When this happens it typically biases the findings of the study.
In order to appraise, whether nonresponse is a problem, it is advisable to consider the underlying nonresponse mechanism which models the reasons for the dropout. In this thesis dropout is not only defined as unit nonresponse, but also describes the absence of responses to certain questions. By viewing the underlying nonresponse mechanism, the aim is to identify the dependency structure between the observed and the missing data, thus to characterize a model describing the present situation a priori. This way at least some additional information about the data can be gathered. Ignoring the underlying missing mechanism in general leads to biased results, so does the assumption of wrong missing models. According to Rubin (1976), who has been the first to formalize these mechanism, it is differentiated between three of them:

- Missing Completely At Random (MCAR)
- Missing At Random (MAR)
- Missing Not At Random (MNAR)

MCAR describes the situation, when there is no systematic difference between people who answered the question and those who did not respond. The responders then can be considered as a subsample of the original sample which usually does not lead to any problems considering parameter estimation. MCAR only becomes a problem, if many variables have single missing values. Then, the loss of information is big, if those observations are simply ignored. Apart from that situation, data with MCAR structure is considered as ignorable, which means that estimates of the unknown parameters maintain their behavioral properties, even if the underlying missing mechanism is ignored during the process of estimation. Those properties for estimates include amongst others unbiasedness and consistency. Bias will be explained in chapter 1.3 and consistency, formally denoted as $\lim_{n\to\infty} p(|\hat{\theta}_n - \theta| > \epsilon) = 0, \quad \forall \epsilon > 0$ (cp. Spieß (2008)), describes the property that the probability of large deviations between the true and the estimated parameter value decreases for increasing sample sizes n. Ignorability is a strong condition that is fulfilled, when the probability of a dropout of a variable cannot be predicted with the help of any other available data for that unit. Thus, the probability of missing data of a variable is independent of the values of the affected variable and all the other variables.

If the latter assumption does not hold, we are dealing with MAR. In this case, the dropout of a considered variable can be explained through the observed data, but it is not related to the person's actual value on the missing variable. In other words, there is a systematic relationship between the propensity of missing values and the observed data. MAR is often referred to as ignorable which does not mean that the cause of missingness may be ignored, but the underlying missing data creation model. Thus, it is sufficient to include the causing variable in the analysis model, but it is not necessary, to know the precise probability distribution generating the dropout. When dealing with the missing data, in the case of MAR no information about the missing data itself has to be included.

The situation is different with MNAR. Here, the probability that a value is missing, depends, possibly besides other factors, on the true, unobservable, missing value itself. Thus, there is a correlation between the cause of missingness and the variable of interest. However, the variable, causing the dropout, is not completely observable and consequently cannot be included in the missing data analysis model. In this case, missing data is not ignorable, because the cause of missingness has not been measured and is therefore not accessible for analysis. Instead, the missing data mechanism itself has to be modeled as you deal with the missing data. It is important to note that in the situation of MNAR the obtained data set is not representative for the population of interest.

According to which of these causes of missingness is present, there are different ways to handle the situation. However, the missingness mechanism is not necessarily consistent for all units and moreover, in general these conditions cannot be tested directly (cp. Graham (2012)). It is only possible, to examine patterns in the data in order to get an idea of what is the most likely mechanism. Often researcher just assume a certain mechanism, in general MCAR or MAR, since for MNAR special nonignorable methods are necessary, but common techniques like ad-hoc methods or imputation assume MCAR or rather MAR. When this assumption is not fulfilled, biased results of population parameters can be produced (cp. Collins et al. (2001)). This thesis does not focus on the burden of trying to find out which of the nonresponse mechanisms is given. Instead, the different scenarios are simulated and considered as known. In the following those missing data mechanisms are introduced formally and regarded as probability models:

Let X be a $(n \times k)$ data set, where x_{ij} is the value of variable X_j for subject i and let further M be the missing data indicator matrix with components

$$m_{ij} = \begin{cases} 1, & \text{if } x_{ij} \text{ is missing} \\ 0, & \text{if } x_{ij} \text{ is present} \end{cases}$$

M can then be written as a function of the unknown parameter vector ξ and the data, consisting of observable and missing values $X = \{X^{\text{obs}}, X^{\text{mis}}\}$. Following Little and Rubin (1987) or rather Little and Rubin (2002), the dropout process is described with the help of the conditional probability and consequently the notation for the three previously described mechanism is constituted by

1. MCAR: $f(M|X,\xi) = f(M|\xi), \quad \forall X, \xi$
2. MAR: $f(M|X,\xi) = f(M|X^{\text{obs}},\xi), \quad \forall X^{\text{mis}}, \xi$
3. MNAR: $f(M|X,\xi) = f(M|X,\xi), \quad \forall X, \xi.$

Measurement Error

Measurement error is often viewed as one of the most damaging sources of error and has therefore been studied extensively in the survey methods literature (cp. e.g. Biemer and Lyberg (2003)). Measurement error refers to continuous variables, whereas in the case of categorical variables it is spoken of misclassification. There are different sources of measurement errors: the responders, the interviewer, the survey questionnaire and other interview factors. But the consequence is always that responders wittingly or unwittingly give incorrect information as response to survey questions. Deliberate, false statements are often based on the respondent's unwillingness to give an answer to a sensitive question. Again, income can be instanced which is often provided as a rounded value (cp. Hanisch (2005)). If this behavior is not on purpose, often the interviewer causes this inaccurate measures of the phenomena of interest due to his speech, appearance, subconscious manipulation or failure in the transcription of the responses. But also the questionnaire can be a major source of error, if it contains ambiguous questions or confusing instructions. In any case, the recorded survey statistic differs from its true value due to imperfections in the way the statistic is collected.

This concept can be represented with the help of a measurement error model, which is a regression model containing measurement errors in the variables. Ignoring those errors, when estimating the regression parameters, results in asymptotically biased estimates, meaning that the parameters do not tend to the true values, even if the sample size is very large. In simple linear models, which we are dealing with in this work, the effect of measurement errors express through systematically underestimated effects (cp. Schneeweiss and Augustin (2006)).

Usually measurement error models are described using latent variables. Thus, let X be the observable variable and X^* the latent variable that would be obtained, if there were no errors in measurement. The true variable X^* is related to the response variable X by a conditional distribution $f(X|X^*, \theta)$, where θ is a vector of unknown model parameters. X then can be understand as noisy observation of X^*, which gives

$$X = X^* + \epsilon.$$

The first differentiation in the context of measurement errors then is between systematic and stochastic errors. In the case of systematic errors, the observable variable X is linked to the true latent variable X^* by a fixed functional relationship. Examples are the shifting of a constant or a proportional error. When the structure of these kind of errors are known, a correction is straightforward, which is why focus should be set on random errors.

In this case, there are different models to describe the relationship of the measurement error ϵ and the corresponding variable. The most common assumption is that of classical measurement errors. It contains an additive or rather multiplicative random error ϵ with mean zero and independence of the true value, so $\mathbb{E}(\epsilon) = 0$ and

$\epsilon \perp X^*$. Typically, ϵ is assumed to be normally distributed with $\epsilon \sim N(0, \sigma_\epsilon^2)$. Thus, in the case of classical measurement errors, observed values are viewed as imperfect measurements of the true values.

Another relationship is described with the so called Berkson model which differs from the classical measurement model by the assumption of the measurement error's independence with the observed value X, instead of the true underlying value X^*, so $\epsilon \perp X$. Thus, in this case the magnitude of the measurement errors does not depend on the variable values being measured. Berkson errors arises in particular, when instead of individual measurements each unit is assigned a certain corresponding group's average value, so that the observed values obviously are less variable than the true underlying values.

Finally, besides the classical and the Berkson error there is also the measurement error of rounding, as has been mentioned before. In this case

$$X^* = h \cdot \lceil (\frac{X}{h}) \rfloor,$$

whereas h describes the threshold, the corresponding value is rounded to. This simple rounding results in a rounding model, where the corresponding rounding error ϵ then is defined as

$$\epsilon = X^* - X$$

which equals equation (1.2) of the classical measurement error model. But in the case of rounding, the error ϵ clearly is not independent of X^* and also not independent of X (cp. Schneeweiss et al. (2010)). Globally, the expectation of ϵ equals zero. Therefore, rounding is usually not a problem, when the focus is on the average of the variable of interest, since the expectation of rounded and not rounded values are approximately equal (cp. Schneeweiss et al. (2010)). However, the measurement error of rounding leads to an abnormal concentration of observations at certain numbers, which is why it is often spoken of heaped data. Consequently, this kind of error distorts the distribution of the variable of interest. Obviously, estimates can be biased, when these are dependent on the shape of the distribution, as it is for example the case, when focus lies on estimating variance. Under certain conditions, however, this distortion can be corrected by a term called Sheppard's correction (cp. Schneeweiss et al. (2010)), which shall not be expanded here. Anyway, due to the impacts on variance estimation, also regression coefficient estimation based on rounded data might be distorted, but according to Schneeweiss et al. (2010) only in the case of rounded explanatory variables. Later, in section 2 the impacts of rounded response values on regression coefficient estimation will be evaluated. But now, in order to initially get an impression of heaped data, the income data of the Allbus data set is viewed and presented with the help of a histogram (see figure (1.2)). For reasons of clarity, the x-axis is truncated after an income value of 3000, which means that 176 observations with higher income value are not plotted. This way, however, the peaks in the distribution can be seen even more distinctly. Income values that are divisible through 100 are reported very often. Hardly any respondent gave the exact income value, in general it is at least rounded to the next threshold of 10.

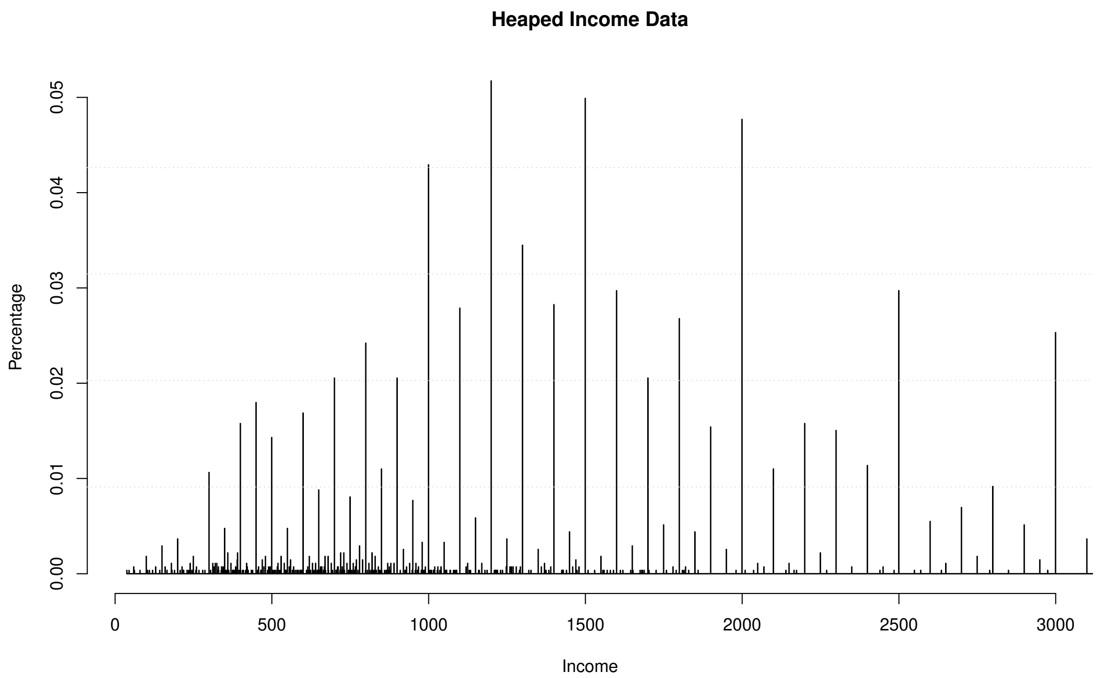

Figure 1.2: Truncated Histogram of income in the Allbus data set

Processing Error

Finally, the raw data that is gathered in a research study typically needs to be processed before it can be analyzed. Also in this step of the survey, crucial errors can occur. Errors in editing, data entry, coding, treatment of outliers or assignment of survey weights are only some of them. Biemer and Lyberg (2003) give a close overview of different data-processing errors, their effects on survey estimates and ways to control them.

All those errors of the TSE are illustrated in the following figure (1.3), which integrated their possible appearance in the different stages of a survey.

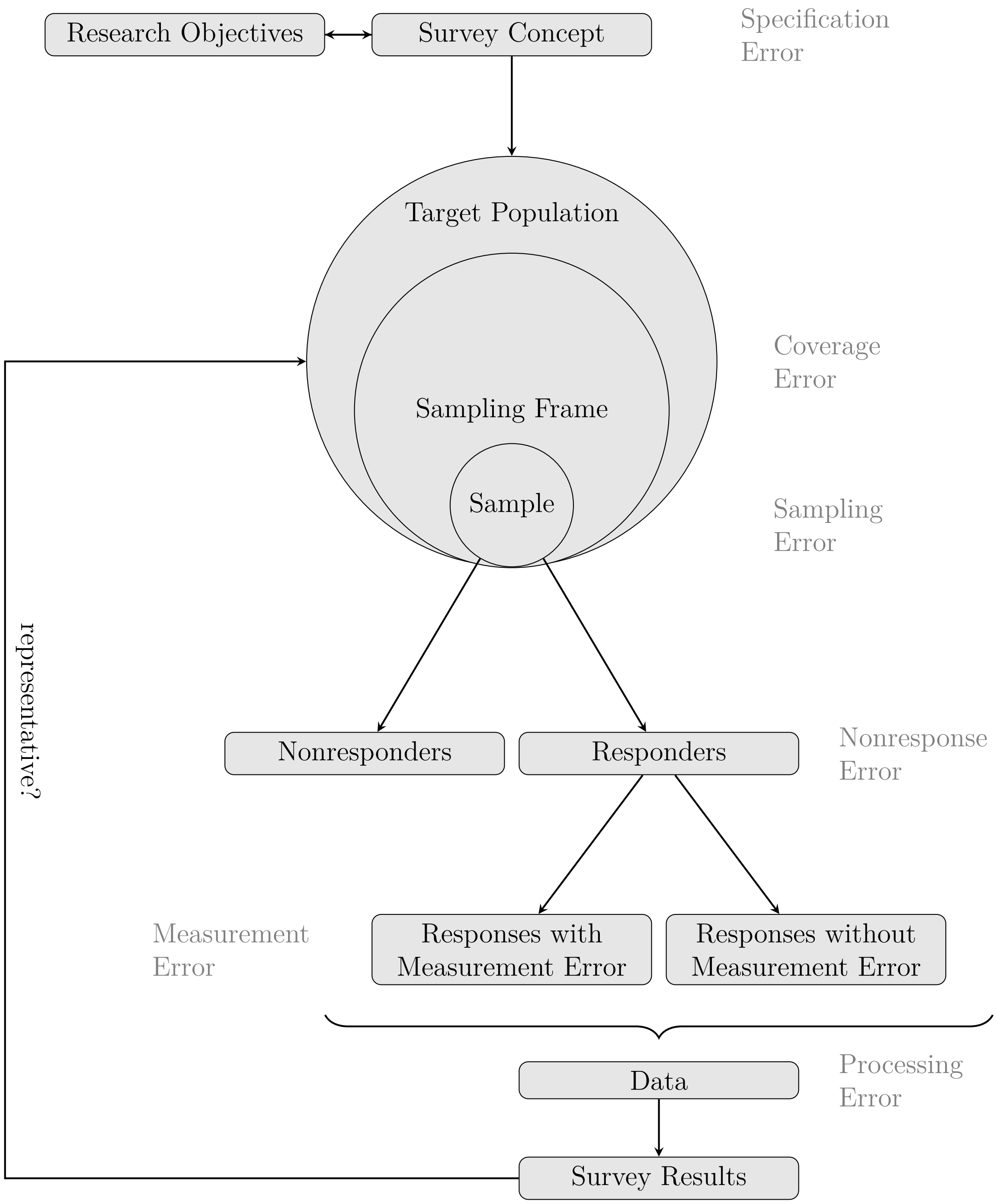

Figure 1.3: Illustration of main error sources of the Total Survey Error

In order to guarantee optimal quality of a survey, all described components have to be considered carefully and minimized with available sources. That way, alternative survey designs, satisfying the specified quality and cost constraints can be compared, using TSE as criterion. Obviously, errors and their consequences on survey results cannot be completely avoided. Besides, there is also the fact of survey-related effects (cp. Weisberg (2005)), which arise due to decisions during the survey. These include effects which are related to the questions, to the order or to the mode of the survey. More precisely, factors like the question wording, the order of questions or response options and the interviewer as an error source itself compared to self-administered surveys can also affect survey results. This shows clearly that survey results are vulnerable to distortion, which makes it necessary, to study the different sources, in order to obtain reliable results. For this purpose the TSE is a good tool. It recognizes distinct ways that survey statistics can depart from some target parameters and it consequently has the potential of protecting against various errors of inference. By separating the types of errors, it is possible to learn about their impacts and handling. This way, the different components can be viewed as a causal, connected system that has to be understand and studied.

1.3 Quantifying the Total Survey Error

Since the main objective of the next section is to evaluate, whether the estimates of the data sets containing errors, are close to the true population parameters, an acceptable metric for quantification of those differences between true and estimated value is needed. This way, the statistical impacts of the different types of errors can be assessed. For this purpose the Mean Squared Error (MSE), which is a measurement of the accuracy of survey data, is appropriate (cp. Biemer (2010a)), as the following details confirm.

Let θ be the true parameter of interest which is in the case of this analyses the income mean μ_{income} and the regression coefficients β. Each corresponding estimate $\hat{\theta}$ that is computed from the survey data, has a corresponding MSE that reflects the effects of the underlying sources of error on the estimates. Thus, the MSE gauges the magnitude of those effects. A small MSE then indicates that the TSE is small and under control. The MSE is defined as the average of the square of the error, with the error being the amount by which the estimate $\hat{\theta}$ differs from the quantity to be estimated, denoted with θ. In practical situations the true, error-free parameter is usually not known. However, this thesis is based on a simulated data basis (cp. section 2.1), which means that this problem is not present. Reformulating the definition of the MSE (cp. appendix A) results in the sum of the squared bias and the observed variance of the estimate:

$$MSE(\hat{\theta}) = \mathbb{E}(\hat{\theta} - \theta)^2 = Bias(\hat{\theta})^2 + Var(\hat{\theta}) \tag{3}$$

As equation (3) then shows, the MSE considers both, systematic and random error. The bias of an estimate is an appropriate measurement for systematic errors. It is defined as the difference between this estimate's expected value $\mathbb{E}(\hat{\theta})$ and θ, the true value of the parameter, being estimated, so $Bias(\theta) = \mathbb{E}(\hat{\theta}) - \theta$. Instead of the absolute bias, often the relative bias $Bias_{\text{rel}}(\theta) = \frac{Bias(\theta)}{|\theta|}$ is considered for reasons of easy interpretation, since the bias is set in relation to the true population value. A negative or positive relative bias can be interpreted as an under- or overestimation of the true value by the received estimate, whereas a relative bias of 0 indicates that the estimation is not biased. Besides the bias, the second components of the MSE is the variance of the corresponding survey estimate. It reflects the impacts of random errors and consequently gives information about the goodness of an estimated parameter by reflecting its variation.

Let now the focus be on a certain characteristic X_j, which is the basic for a statistic of interest like the mean. Its value x_{ij} for a particular unit i in the survey is higher or lower than its true value x^*_{ij}, since there are various error sources in the survey that have a cumulative effect on the responses.

$$x^*_{ij} = x_{ij} + \epsilon_i$$

is received. In the case of systematic error, the sum of the errors is not zero, because either positive or negative errors are dominant. $\mathbb{E}(\epsilon_{ij}) \neq 0$ directly affects the average value of a variable, $\hat{\mu}_j = \sum_{i=1}^{n} \frac{1}{n} x_{ij}$, so that the estimated mean is either negatively or positively biased. In contrast to that, random errors do not add bias to an observation, which means they have a mean of zero. Due to $\mathbb{E}(\epsilon_{ij}) = 0$, random errors do not affect the estimated mean of the variable, but they only increase the variance of the estimated values, which is the second component in equation (3). Since $Var(\epsilon_{ij}) \neq 0$, $Var(x^*_{ij}) = Var(x_{ij}) + Var(\epsilon_{ij})$ increases with constant $Var(x_{ij})$. Increased variance has for example a weakening effect on regression coefficients of the independent variables.

Thus, in the following section 2.3, which deals with the analysis of direction and magnitude of the different TSE effects on selected parameter estimation, it is focused on those parameters, mean μ and regression coefficients β. More precisely, μ_{income} is chosen as well as all regression coefficients of the predictors from the linear regression model (4) presented later on. This model contains the estimation of the intercept, as well as 21 further regression parameter β belonging to one continuous, one binary and 4 categorical predictors.

To sum up, each discussed error source contributes a systematic error, a random error or both to the TSE. Thus, Biemer (2010a) shows a further fragmentation of the MSE, differentiating between the contributions of the diverse error sources. In the following analyses, the impacts of the different errors are studied separately, which is why it is only referred to this separation. However, this chapter should have pointed out that the MSE is an appropriate measure to quantify the TSE, since it considers both components, systematic and random errors. Therefore, it will be used in the following chapter 2, to quantify the impacts. Figure (1.4) finally illustrates, what has been described before.

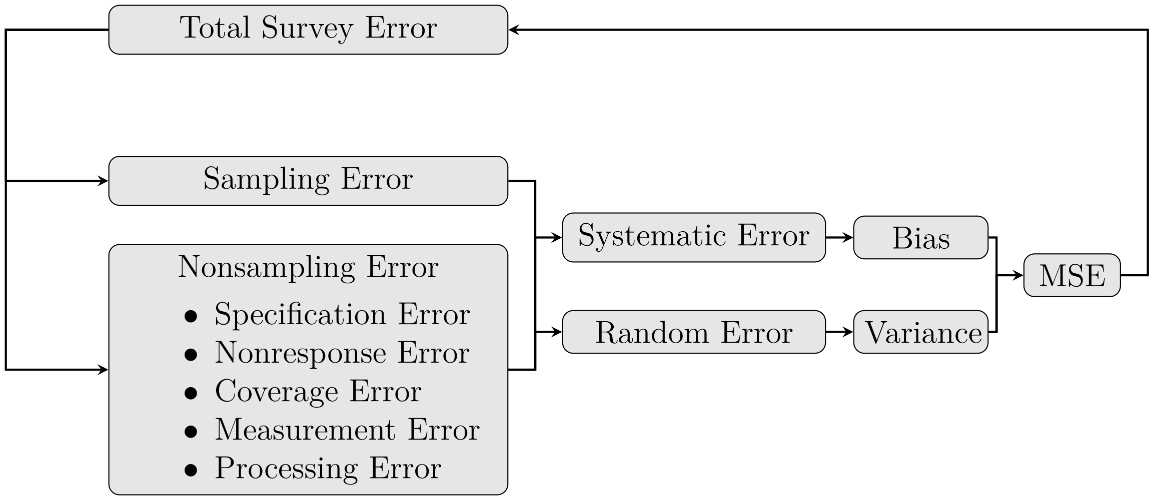

Figure 1.4: Quantification of the TSE (graphical representation based on Biemer (2010))

2 Effects of the different Types of Survey Error

2.1 Simulated Data Basis

In order to demonstrate, whether and, if applicable, how much the different types of survey error affect the considered parameter estimation, a data basis is needed which is regarded as gold standard. This implies that the parameters of the data basis are known and are considered as true values. That way, deviations can be detected, when error model are applied and traced back to the different sources of the TSE. For this approach it is reasonable to simulate data, since then the data generating processes are known and consequently also their parameters. Available data sets, however, are not suitable, since they already contain different errors like missing data or measurement error, therefore it is not justifiable to consider those values as true. Nevertheless, it is desirable to have a data set that is close to reality. That means distributions and dependence structures between different variables should not be pure invention and furthermore the focus should be on attributes that are often considered in the field of survey research. From these considerations it appears that the structure of the simulated data should orientate on an established and reputable national Social Survey, in order to fulfill the previously mentioned criteria.
For this purpose the German General Social Survey (Allbus) 2014 is chosen. Since the Allbus gives a representative cross section of attitudes, behavior and social structure of the German population every two years, this version is the recently published one at the time, when the analysis within the framework of the thesis started. For further information to this study it is referred to the Allbus-Homepage (cp. GESIS (2016)). Here, also the underlying data set, consisting of 3471 observations and 861 variables, is freely accessible.

After reduction of the Allbus 2014 data set to the selected, relevant variables X_{income}, X_{age}, X_{gender}, $X_{\text{education}}$, $X_{\text{professional activity}}$, $X_{\text{family status}}$, $X_{\text{election intention}}$ and $X_{\text{willingness}}$ and afterward combination of several factor levels with only few observations, the pattern of this data set is analyzed in order to re-simulate it. X_{income} is the variable of interest and will form the dependent variable of the analysis model. The remaining variables will serve as predictors. Their influence on the values of X_{income} should obviously differ according to the Allbus.

The construction process of the categorical covariates is based on the idea of sampling the possible values from the appropriate Allbus contingency tables. This approach is equivalent of drawing balls from an urn with replacement which implicates that the distributions of the constructed variables are known.
In the first step the characteristic gender is simulated. Therefore, 3500 observations, which is determined as size of the new simulated data set, are drawn randomly with replacement, whereas the probability of 0.508 for a male outcome, or rather the converse probability 0.492 for a female outcome, are given, based on the relative frequencies in the Allbus data set. Thus, the simulated variable gender then follows a binomial model with parameters $n = 3500$, $p = 0.508$ and $1 - p = 0.492$, abbreviated as $X_{\text{gender}} \sim Bin(3500, 0.508, 0.492)$ (cp. Fahrmeir et al. (2007)).
In order to roughly maintain the dependency structure between attributes, the next categorical variable education is constructed the same way, but conditional on the simulated variable gender. Thus, the distribution of education for male observations differs from that for females which results in two multinomial models with given parameters. As before, those are determined by joint contingency tables of the variables gender and education in the Allbus data set. So, if an observation has the value male, a corresponding value of the appropriate distribution for education is attached. The multinomial distribution, abbreviated as $X \sim Multin(n, p_1, \cdots, p_k)$, is a generalization of the binomial distribution with regard to the number of categories k. In the case of education it is distinguished between $k = 6$ categories, namely "no graduation", "volks-, hauptschule", "mittlere reife", "fachhochschulreife", "hochschulreife" and "other graduation". Thus, $X_{\text{education}}|X_{\text{gender}} \sim Multin(n, p_1, \cdots, p_6)$ with different probabilities p for men and women. It is received:
$X_{\text{education}}|(X_{\text{gender}} = \mathit{male}) \sim Multin(1746, 0.018, 0.307, 0.305, 0.085, 0.276, 0.010)$,
$X_{\text{education}}|(X_{\text{gender}} = \mathit{female}) \sim Multin(1754, 0.019, 0.254, 0.356, 0.071, 0.287, 0.012)$
This chain of dependency between the simulated covariates is then carried on in the same way, so that the next categorical covariate depends on all those variables that have been simulated before. This way, the construction of the third covariate $X_{\text{professional activity}}$ equals drawing values from 12 different distributions with replacement, whereas it is differed between 4 possible outcomes for the job variable, namely "full-time", "half-time", "part-time" and "not employed". The 12 urns arise from the different possible combinations of gender and education. Finally, this approach is pursued, when focusing on the outstanding categorical covariates family status and election intention. Those can take the values "married living together", "married living apart", "single", "divorced", "widowed" and "CDU-CSU", "SPD", "die gruenen", "die linke", "extreme right-wing", "FDP", "Other Party", "would not vote".

Obviuosly, the order of constructing the categorical variables has an influence on the resulted variable values. Since the dependence structure of the Allbus data set shall only roughly be maintained, the determination due to careful consideration is suitable and the outcome satisfies claims.

In order to simulate the continuous variable age, its distribution in the Allbus data set has to be analyzed. As the following plots (2.1) and (2.2) show, the variable X_{age} in the Allbus data set does not exactly follow a normal distribution. The bars of the histogram (2.1) illustrate the true age distribution in the Allbus data set, whereas the drawn blue line indicates the corresponding normal distribution. Obviously, there is a deviation from the normal distribution especially in lower age categories. Here, we have more observations with small age values than it would have been expected, if age follows a normal distribution.

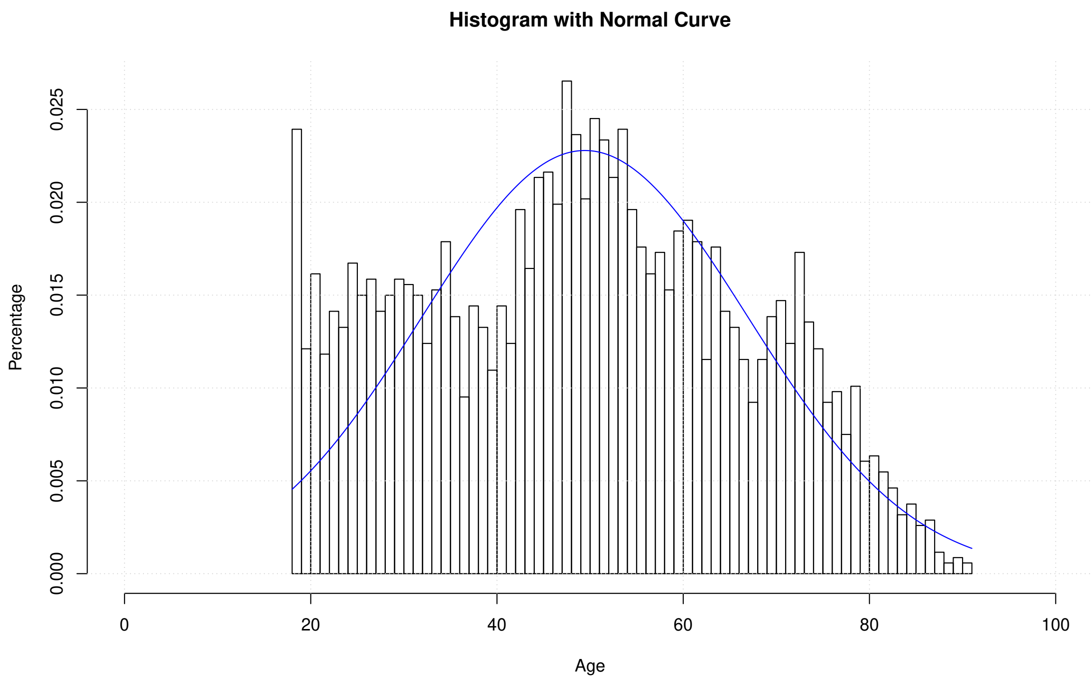

Figure 2.1: Histogram of age in the Allbus data set containing a normal curve

The Normal Quantile-Quantile-Plot (2.2) confirms these deviations. By plotting the quantiles of the observed age distribution against the theoretical quantiles of the normal distributions, those two distributions are compared. The closer all points lie to the drawn, blue line, the closer the distribution of the sample comes to the normal distribution. It can be seen that for middle aged people of the sample, the normal distribution provides a quite good approximation, whereas extreme values obviously deviate from the line. However, the assumption of normal distributed simulated age values is not completely devious, when viewing those results.

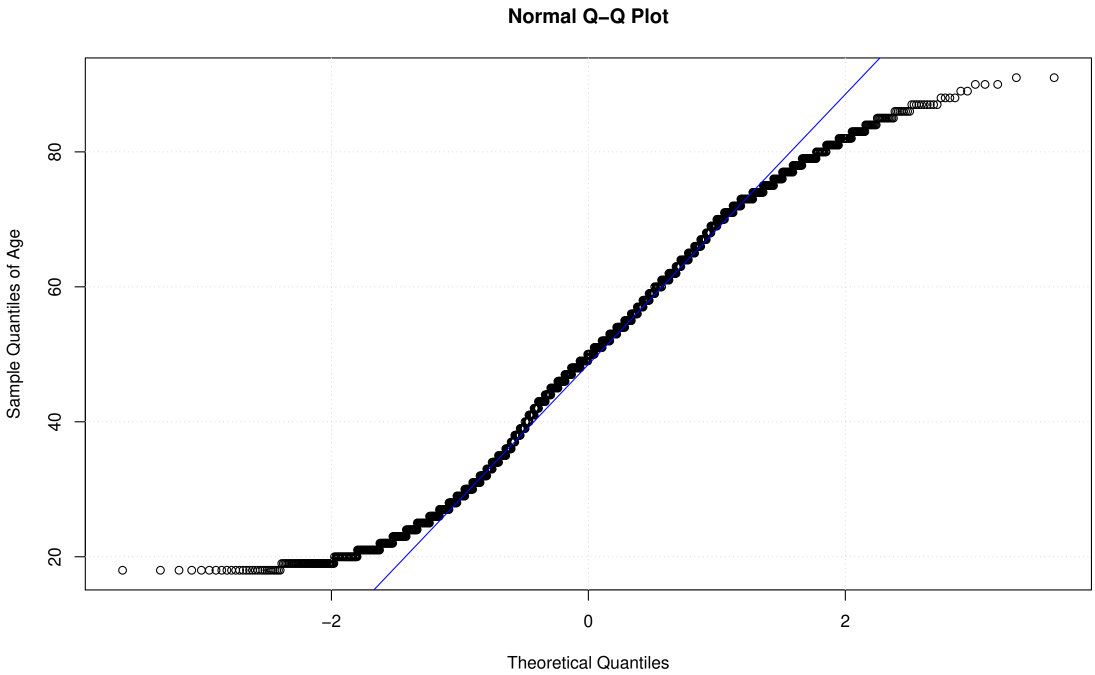

Figure 2.2: Normal Quantile-Quantile-Plot of age in the Allbus data set

As a consequence, X_{age} will be simulated as a truncated normal distribution, which means that the probability distribution of the normally distributed random variable age is bounded below and above. This way, the simulated age variable is closer to the given distribution in the Allbus data set and especially satisfies the condition that only full-aged people are regarded. Furthermore, it has been revealed that the distribution of age differs between the different categories of the other variables, except gender. That means, comparing the distribution of age between men and women, a roughly identical distribution is obtained, whereas it becomes apparent that the subsamples of the other variables education, professional activity, family status and election intention, do not originate from the same distribution regarding age. Therefore, the idea is to simulate the distributions of age for the different combinations of those categorical variables according to minimum, maximum, mean and variance in the Allbus data set. Since considering all those categorical variables would expend too much effort and the aim is just to roughly maintain the dependence structure of the Allbus 2014, the approach is restricted to the covari-

ates $X_{\text{education}}$ and $X_{\text{family status}}$. Thus, for each combination of those two variables, the distribution of the Allbus data set is reconstructed via a truncated normal distribution. Then, according to which value combination is given in the previously simulated data, a value for age is drawn from the appropriate distribution and is assigned to the corresponding observation. Thus, also the continuous independent variable X_{age} is constructed with a certain dependence structure to the other variables, namely by
$X_{\text{age}} \mid X_{\text{education}}, X_{\text{family status}} \sim trunc\,N(\mu_{\text{age(Allbus)|education(Allbus),family status(Allbus)}},$
$\sigma^2_{\text{age(Allbus)|education(Allbus),family status(Allbus)}})$.

Besides those covariates, the simulated data set should contain a dependent variable whose dependency with all the other covariates should be defined with the help of a linear regression model. In this case X_{income} is chosen as the dependent variable. First, the considered regression model

$$\begin{aligned}\log(X_{\text{income}}) = \beta_0 + \beta_{\text{age}}X_{\text{age}} + \beta_{\text{gender}}X_{\text{gender}} + \beta_{\text{education}}X_{\text{education}} + \\ + \beta_{\text{professional activity}}X_{\text{professional activity}} + \\ + \beta_{\text{family status}}X_{\text{family status}} + \\ + \beta_{\text{election intention}}X_{\text{election intention}} + \epsilon \end{aligned} \tag{4}$$

is fit on the Allbus data. At this point it is important to note that the output of the fitted regression model contains estimators for each category of a categorical variables, except the reference category, but equation (4) summarizes the regression coefficients of each categorical variable in a β-vector, in order to maintain clarity. Thus, $\beta_{\text{education}}$ for example comprises $\beta_{\text{no graduation}}$, $\beta_{\text{volks-, hauptschule}}$, $\beta_{\text{mittlere reife}}$, $\beta_{\text{fachhochschulreife}}$ and $\beta_{\text{hochschulreife}}$. The received, estimated regression coefficients $\hat{\beta}$ are then treated as true values and are also assumed for the simulated data set. With those beta coefficients, the previously simulated variable values and a normally distributed error term $\epsilon \sim N(0, 0.5)$, which also orientates on the regression output of the Allbus data set, predictions for the income variable of the simulated data set are made. Since the model contains the logarithmized dependent variable, which prevents from negative response values, the result has to be transformed afterwards in order to receive the interesting income values. Due to the normally distributed error terms, it follows then that $\log(X_{\text{income}})$, conditional on the other covariates X, is also normally distributed.

As last step, another categorical covariate, describing the willingness to take part in the interview, is constructed. In the Allbus data set this ordinal variable takes four possible values, evaluating the cooperation of the participant, lasting from “very easy”, over “rather easy” and “rather difficult” to “very difficult”. The simulation procedure of willingness then is analogous to the approaches of the other categorical variables before. Hereby, $X_{\text{willingness}}$ is simulated conditional on X_{gender} and $X_{\text{education}}$. The other covariates are omitted for the sake of convenience. The property of the variable $X_{\text{willingness}}$ is that it is not part of the linear regression model (4), which means, it is not constructed with a conscious influence on the dependent variable income. However, it cannot be excluded that there is any relationship between those two variables in the simulated data set. This variable will later play a role in the Heckman Selection Model in chapter 3.2, dealing with likelihood-based approaches, modeling the missing data process. There, a covariate is desirable that, among others, explains the dropout of a variable, but not the values of the affected variable itself. This aspect, however, will be discussed more precisely later on.

Now the data basis is received which contains the income as the dependent variable of the linear regression model, the corresponding predictors and an additional variable describing the willingness of the individuals to take part in the survey. For a better overview of the simulated data set view appendix C. Here, the frequencies of the categorical covariates, as well as the summaries of the constructed continuous variables are presented. In order to be able to compare the results of the simulated data set and the underlying Allbus 2014, also the summaries of the shortened Allbus 2014 data set are listed there.
It shows that in the simulated data set a few more women are included, whereas in the Allbus 2014 the male participants slightly dominate. More important, however, is the fact that the distribution of income in the Allbus 2014 contains more outliers compared to the simulated data set which could lead to overoptimistic results in the analyses to come. Apart from these two points, the two data sets are not deviating remarkably. Nevertheless, despite the similarity, it is noteworthy, to be careful with interpretations and inferences to the underlying Allbus population. The distributions and dependencies are only approximated and furthermore, the Allbus 2014 data set itself contains errors like missing data or measurement errors. Therefore, it is quite possible that the simulated data set is not representable for the underlying population of the Allbus, which is the whole adult population of Germany. Moreover, it is not possible, to test the goodness of the error models,

which are consciously build in chapter 2.2, in order to demonstrate the impacts of the different errors of the Total Survey Error. They are orientated as effectively as possible on the errors of the Allbus data set, however, they are very specific and cannot be tested, as mentioned before.
Nevertheless, this simulation approach results in a realistic data set which will be viewed as gold standard in the following analyses. As a result, the parameters of this data set are considered as true population values. As mentioned before, the following analyses of the parameters focus on the mean of the income variable and the regression coefficients of the independent variables in the linear model (4). Both, the income mean μ_{income} of 1492.951 and the exact values of the regression coefficients β can be seen in appendix C.
For a better understanding of those values, some of the parameters shall be interpreted briefly before proceeding with the creation of the different error models. The intercept estimate indicates the income value that is predicted for a person belonging to the reference category for all variables and having value 0 for all continuous predictor variables. In this case those category comprises divorced men of age 0 with other education, full-time work, who vote other parties. Obviously, the value of this reference group cannot be interpreted meaningfully. Considering then the parameter value of the only continuous variable age, the corresponding beta value of $\beta_{\text{age}} = 0.0113720951$ represents the difference in the predicted income value for each one-unit difference in age, if all the other variable values remain constant. This means that if the age increased by one unit, and all the other variables stayed the same, the income rises by $\exp(0.0113720951) = 1.0114370032$ units, on average. Interpreting then exemplary β_{female}, the regression coefficient for females, $\beta_{\text{female}} = -0.3783040793$ is the average difference in the logarithmized income between the reference group of men and the category, for which $X_{\text{gender}} = 1$, namely women. So compared to man with the same attributes, represented by the other variables, we would expect the income of a woman to be smaller by $\exp(-0.3783040793) = 0.6850221679$, on average. The interpretation of the other categorical variables is then straightforward and also has to be conducted with respect to the reference category. Besides the intercept estimate, also those parameters of age, females, mittlere reife, no graduation, volks-, hauptschule, half-time, part-time, no employment, single, married living apart, married living together, die gruenen, FDP and would not vote, display a p-value smaller than the alpha level of 0.1 in the simulated data basis. Thus, those effects differ statistically significant at a significance level of 0.1 from the respective reference category.

2.2 Creation of Data Sets with selected Error Models

In the following chapter it is described, how the data basis is deliberately changed, in order to reconstruct the different errors of the TSE approach for analyzing their impacts on parameter estimation. The sources of specification and processing error are omitted in those analyses, since on the one hand specification error is based on a particular research question which is not given in the present situation and process error arises in the step after data collection which is not focused on at this point. Therefore, it is concentrated on sampling error, coverage error, nonresponse error and measurement error. In these cases, data sets with simulated errors are constructed, whereas this approach is repeated $T = 100$ times, respectively, so that a set of data sets containing the same error model is obtained. Obviously, the main objective by iteratively creation of several data sets with the same error model is to avoid randomness of the results. These data sets are then the basis for the parameter estimation of interest. In each of those 100 data sets income mean and regression coefficients are calculated and these estimates will then be compared to the true population values in section 2.3. This way a raw bias is received and the relative bias can be easily calculated. Furthermore, a sampling distribution for those estimated parameters of the different samples is received. The variance of those estimates gives information about the goodness of the average parameters, so does the corresponding MSE.

Sampling Error

In order to demonstrate the impacts of this error source, random samples of size n are repeatedly and independently drawn from the total data set of $N = 3500$ observations. n is varied and takes the values $n = 500, 1000, 2000, 3000$. This way, differences between comparatively small, medium-sized and large sample sizes can be shown.

Coverage Error

As it has been the case in the previous section of sampling error, also the creation of coverage error is based on random samples of different sizes n which are repeatedly and independently drawn. In doing so, the different sample sizes of $n = 500, 1000, 2000, 3000$ stay the same. However, coverage error arises, when some subpopulations are more likely to be overrepresented in the sample than others.

This can be simulated by not attributing each of the $N = 3500$ observations the same probability of $\frac{1}{N}$ to reach the sample. Instead those probabilities differ and are dependent of certain variable values. Due to literature (cp. Smith (2008)) and careful considerations, the variables X_{gender} and $X_{\text{education}}$ are chosen for this purpose. First, women shall be overrepresented, since they tend to have less working hours compared to men which has an influence on their reachability. Consequently, it is assumed that it is more likely for a female individual to take part in a survey than a man who just differs from her with respect to gender. Thus, the probability for a woman to be sampled shall be doubled, whereas the probability for men is halved. Furthermore, it is assumed that the educational achievement is determining for the probability to be sampled. Well-educated people, including the variable values “hochschulreife” and “fachhochschulreife”, shall be overrepresented and are assigned the doubled probability of $\frac{2}{N}$. Observations with value “mittlere reife” and “volks-, hauptschule” still have the same probability of $\frac{1}{N}$ and the probability of low educated individuals (“no graduation”), as well as other educated individuals, is halved. From above the following joint sampling probabilities arise:

	male ($\frac{1}{2N}$)	female ($\frac{2}{N}$)
hochschulreife ($\frac{2}{N}$)	$\frac{1}{N^2}$	$\frac{4}{N^2}$
fachhochschulreife ($\frac{2}{N}$)	$\frac{1}{N^2}$	$\frac{4}{N^2}$
mittlere reife ($\frac{1}{N}$)	$\frac{1}{2N^2}$	$\frac{2}{N^2}$
volks-, hauptschule ($\frac{1}{N}$)	$\frac{1}{2N^2}$	$\frac{2}{N^2}$
no graduation ($\frac{1}{2N}$)	$\frac{1}{4N^2}$	$\frac{1}{N^2}$
other graduation ($\frac{1}{2N}$)	$\frac{1}{4N^2}$	$\frac{1}{N^2}$

Table 2.1: Joint sampling probabilities of coverage error model

Those probabilities, which obviously have a strong influence on the degree of representativeness of the sample and consequently on the magnitude of the impact, seem to be chosen arbitrarily. But since they have to be determined some how, to conduct the simulation, it is the best way to orientate them on literature knowledge. However, their choice can of course still be seen as critical. Because of the different drawing probabilities of observations, it is possible that in some samples with small sample sizes certain variable combinations are not only underrepresentated, but even not present. In order to receive fairly representative regression coefficient estimates, only samples with at least one observation of each gender-education combination are used for further analyses, while others are discarded and sampling is repeated until 100 data sets are received.

Nonresponse Error

Showing the impacts of nonresponse error requires the differentiation of the three nonresponse mechanism (see section 1.2). Therefore, the different error models are regarded separately. In all three situations, however, it is confined to the case, where just one variable, namely X_{income}, has missing data, while all the other predictors of the analysis model are always observable. Yet, this case could be extended to the case in which explanatory variable values are also sometimes missing. For simulating MAR the Allbus data set, which has missing values in the income variable itself, can serve as a model again by orientating the dropout-probabilities on the given variable values. In the Allbus data set the dropout depends amongst others on the variables X_{gender}, $X_{\text{education}}$ and $X_{\text{willingness}}$, as corresponding χ^2-tests showed. Consequently, with those variables a logit model is built on the Allbus data, having a binary "missing" indicator M as dependent variable that shows, whether the observation i has a missing value in the income variable ($m_i = 1$) or not ($m_i = 0$).

$$\begin{aligned} &\mathbb{P}(M=1) = \\ &\frac{\exp(\beta_0 + \beta_{\text{gender}} X_{\text{gender}} + \beta_{\text{education}} X_{\text{education}} + \beta_{\text{willingness}} X_{\text{willingness}})}{1 + \exp(\beta_0 + \beta_{\text{gender}} X_{\text{gender}} + \beta_{\text{education}} X_{\text{education}} + \beta_{\text{willingness}} X_{\text{willingness}})} = \\ &\frac{1}{1 + \exp(-(\beta_0 + \beta_{\text{gender}} X_{\text{gender}} + \beta_{\text{education}} X_{\text{education}} + \beta_{\text{willingness}} X_{\text{willingness}}))} \end{aligned} \tag{5}$$

With the resulted regression coefficients β and the previously simulated covariates, the dropout in the simulated data set is predicted, since for each observation the probability for a dropout is obtained. Thus, the individual dropout of the income variable is following a bernoulli distribution with given probability. As a result, a data set of length $N = 3500$ is received that contains a not predefined number of missing values in the income variable. As it has been the case when focusing on the other error sources, this procedure of creating missing values with a given probability is repeated $T = 100$ times. Here, the calculated probabilities resulting from the logit model remain the same for all 100 iterations, but the random process, deciding, whether the income value drops out or not, is performed again. That is why the number of missing values in those $T = 100$ data sets differ, even though the underlying error model is the same. Again, those resulting data sets are then the basis for further analysis.

If the underlying dropout process follows MNAR, the missingness depends not only on the values of the independent variables, but also on the values of the focused variable with missing values itself. In this case the income. The modeling approach then is equivalent to (5), except for the income variable which is added as covariate:

$$\mathbb{P}(M=1) = \tag{6}$$
$$\frac{\exp(\beta_0 + \beta_{\text{gender}} X_{\text{gender}} + \beta_{\text{education}} X_{\text{education}} + \beta_{\text{will.}} X_{\text{will.}} + \beta_{\text{income}} X_{\text{income}})}{1 + \exp(\beta_0 + \beta_{\text{gender}} X_{\text{gender}} + \beta_{\text{education}} X_{\text{education}} + \beta_{\text{will.}} X_{\text{will.}} + \beta_{\text{income}} X_{\text{income}})} =$$
$$\frac{1}{1 + \exp(-(\beta_0 + \beta_{\text{gender}} X_{\text{gender}} + \beta_{\text{education}} X_{\text{education}} + \beta_{\text{will.}} X_{\text{will.}} + \beta_{\text{income}} X_{\text{income}}))}$$

For $\beta_0, \beta_{\text{gender}}, \beta_{\text{education}}$ and $\beta_{\text{willingness}}$ the preassigned values from model (5) are maintained. However, since the true values of the missing income values in the Allbus data set are not known, β_{income} can not be calculated in the same way. Instead the beta value for the income value has to be determined in a plausible way. It is set to $\beta_{\text{income}} = 0.00005$ which can be interpreted as an increase of the odds ratio by $\exp(0.00005) = 1.00005$, if the income rises of one unit. The odds ratio is defined as ratio between the probability of a dropout and the probability of no dropout (cp. Fahrmeir et al. (2007)). Due to this determination, it can be proceeded in the same way as in the MAR case. However, it is important to note that this determination has an influence on further analysis. Choosing a larger value for β_{income} would implicate a higher odds ratio and consequently a higher probability for a dropout. Nevertheless, this value has to be set to a realistic, fix value. Choosing 0.00002 for β_{income} would give nearly the same data sets, as have resulted, when applying the MAR mechanism. Since in this case, nearly the same observations have a dropout, the income of a person would not have a great influence on the missingness and the differences between those two models could not be analyzed. Therefore, a larger value for β_{income} is necessary. However, choosing β_{income} even larger than 0.00005 would have resulted in dropout-probabilities close to 1. Thus, this parameter choice of $\beta_{\text{income}} = 0.00005$ is quite reasonable.
Finally, the last missingness process MCAR is based on randomness. Consequently, a certain, determined number of missings, which orientates on the number of missing values in the income variable of the Allbus data set, is created with some completely random process. In the income variable of the Allbus data set a number of 746 values are missing, which constitutes a proportion of 0.215. Transferring this percentage to the simulated data set, finally results in $T = 100$ data sets of size $N = 3500$ with 750 missing values in the income variable, respectively.

Measurement Error

As it has been described before, measurement error comprises many different sources. Consequently, the analysis of this error source has to be specified. While a lot of research has already been done in this field of application (see e.g. Schneeweiß and Mittag (1986), Buzas et al. (2005), Chesher (1991)), it is now only focused on one specific type of measurement error, namely the measurement error of rounding which is a typical behavior of people having to answer questions about the own income. Rounding is a subcategory of coarse data which is defined as data that are neither entirely missing nor are perfectly present (cp. Heitjan and Rubin (1991)). Coarse data also encompasses heaping, censoring and missing data which connects this error source with that of nonresponse error. In order to find out about the impacts of this error model, rounded values in the data set have to be created. At this, rounding to the next multiple of 10 is considered, which results in the rounding function $f_{10}(x) = 10 \cdot \lfloor \frac{x}{10} + \frac{1}{2} \rfloor$. This rounding threshold is chosen, since the resulting number of income values that are divisible through 10, 100 and 1000 this way is most similar to the corresponsing percentage of the Allbus data set.
Based on the approach of the nonresponse error, different models for the creation of rounding behavior are chosen. It is differentiated between rounding behavior that is completely at random, rounding behavior that depends on covariates of the data set and finally rounding behavior that cannot only be traced back to certain covariates, but also to the values of the interested, rounded variable itself, here the income variable. Thus, those models are similar to those of MCAR, MAR and MNAR and are consequently in the following abbreviated as RCAR, RAR and RNAR. In order to avoid confusion, even the same covariates, namely X_{gender}, X_{Election} and $X_{\text{ willingness}}$ are chosen, to have an influence on the rounding behavior, when focusing on the appropriate rounding mechanisms.
The simulation of the two rounding models RAR and RNAR is then proceeded analogous to the nonresponse case, except that the dummy variable of interest M, having indicated the dropout of an income value before, is now replaced by a dummy variable X_{rounding} that indicates, whether the income value is rounded ($X_{\text{rounding}} = 1$) or not ($X_{\text{rounding}} = 0$). Based on the previously introduced rounding function, an income value is regarded as rounded, if it is divisible through 10. In this case, it is assumed that the individual has not stated its exact income. Even though these numbers appeal to be rounded, it is important to note that they could nevertheless be the true values and consequently have no measurement error.
For the last rounding model RCAR again the Allbus data set serves as a role model. Analyzing the rounding behavior of the income variable in the Allbus data set gives

that 71.795% of the listed values are divisible through 10. However, this percentages is transferred to the simulated data set, so that 100 data sets are created containing 2200 observations which are rounded to the next threshold of 10, respectively. Which observations are rounded is chosen randomly. With the already existing number of 313 income values of the simulated data set that have already been divisible through 10 before simulating the rounding error consciously, there are then 2513 income values divisible through 10 in each of the 100 data sets, which corresponds to the given percentage of the Allbus data set. This way, data sets with different rounding errors are received and the impacts of this kind of measurement error can be analyzed in the next step.

2.3 Direction and Magnitude of Effects on Parameter Estimation

Sampling Error

Starting with the impacts of the sampling error, the following tables and figures give an overview of the results. First, it is focused on the estimates for the income mean μ_{income}. Table (2.2) comprises the expected value, the corresponding bias and relative bias, the variance and the MSE of the estimates for the different sample sizes. The listed values are the averages of the 100 data sets, respectively.

	$n = 500$	$n = 1000$	$n = 2000$	$n = 3000$
μ_{income}	1492.951	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1490.602	1492.278	1492.281	1493.053
$Bias(\hat{\mu}_{\text{income}})$	-2.34902	-0.67283	-0.67011	0.10241
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00157	-0.00045	-0.00045	0.00007
$Var(\hat{\mu}_{\text{income}})$	2006.50326	1094.93459	275.11373	61.83687
$MSE(\hat{\mu}_{\text{income}})$	2012.02115	1095.38729	275.56277	61.84736

Table 2.2: Expected value, variance, bias and MSE of mean estimation of income in order to quantify sampling error

It shows that for all regarded sample sizes, $\mathbb{E}(\hat{\mu}_{\text{income}})$ is close to the true population mean of 1492.951, whereas, as expected, the estimates are better for larger samples. The bias takes the highest value of -2.34902 in the case of a small sample of $n = 500$. Here, the result is a negative bias, which means that the parameter is underestimated. However, this is not the case for all sample sizes. It should be noticed that besides the decreasing bias for larger sample sizes, also the variances of

the estimates decrease distinctly which results in smaller MSEs and a higher representativeness.
In order to make sure that these deviations from the true mean income values are statistically not significant, a two-sided t-test can be conducted. The null hypothesis contains the statement that the true income mean of 1492.951 could also be the mean value of the sample containing a certain error, here the sampling error. This means that the resulting bias might be by chance, whereas the alternative hypotheses states that there are significant differences. These deviations between the true mean income and each of the 100 resulting mean income estimates per error model are tested on a significance level of $\alpha = 0.05$. As a result, none of the average mean estimates for $n = 500$, $n = 1000$, $n = 2000$, $n = 3000$ differ significantly from the true income mean, since in the sample scenario of $n = 500$ in 98 of 100 times the null hypothesis may not be rejected, for $n = 1000$ it is in 97% of the cases, for $n = 2000$ in 99% and for $n = 3000$ in all 100% of the cases. Consequently, mean estimation here is not noticeably affected by sampling error which can also be seen in the following figure, showing for the different sample sizes the corresponding boxplots of the different mean estimates of the $T = 100$ samples, respectively.

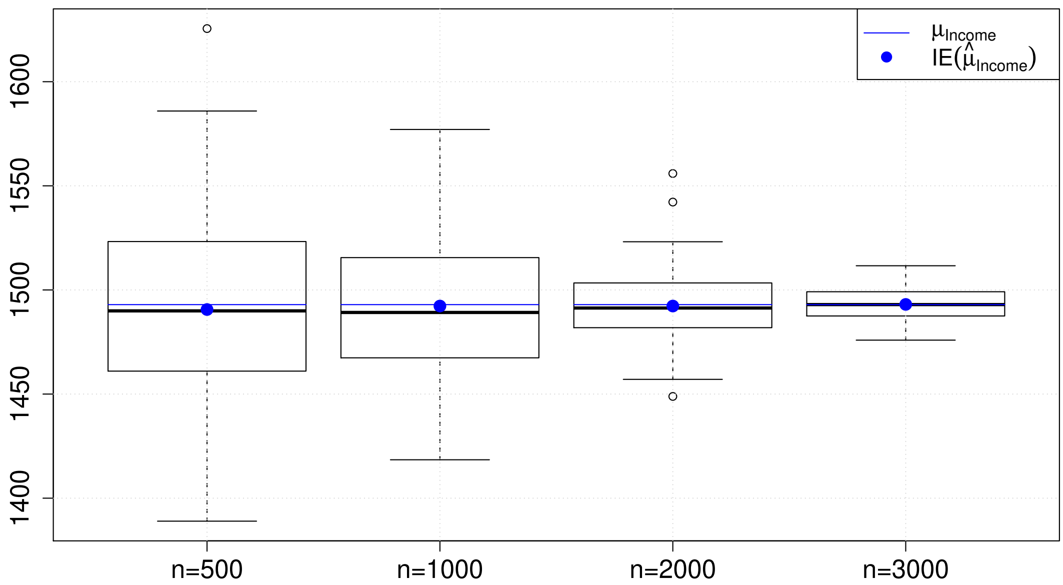

Figure 2.3: Boxplots of income mean estimates for different sample sizes

Whereas the drawn black line within the box represents the median of the sample, the blue line indicates the true value μ_{income} of the target population. Consequently, those two values should not be compared directly, since they indicate different statistical measurements. Instead the difference between the blue line and the blue dot, noting the estimated value, should be assessed. For increasing sample sizes, the boxes, displaying the variation in those samples, obviously get smaller and the prediction of the true value more precise.
In order to make also statements about the estimation of the regression coefficients, we consider table (2.3), which is confined to the representation of the resulting MSEs for reasons of clarity. The complete table (C.6) with expected values, variances and biases can be found in the appendix.

	$n = 500$	$n = 1000$	$n = 2000$	$n = 3000$
$MSE(\hat{\beta}_0)$	0.0825926243	0.0290543619	0.0098488046	0.0020103048
$MSE(\hat{\beta}_{\text{age}})$	0.0000025652	0.0000015689	0.0000004719	0.0000001000
$MSE(\hat{\beta}_{\text{female}})$	0.0021497401	0.0007941525	0.0002463387	0.0000420196
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0741067867	0.0365600113	0.0105487393	0.0019236736
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0601953276	0.0203148217	0.0075737900	0.0012522573
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0649202739	0.0219705149	0.0077103756	0.0012620555
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0693787401	0.0229147785	0.0078471397	0.0012432106
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0635400125	0.0225464891	0.0077629780	0.0014164809
$MSE(\hat{\beta}_{\text{half-time}})$	0.0043497917	0.0022160056	0.0006491382	0.0001424458
$MSE(\hat{\beta}_{\text{part-time}})$	0.0068955666	0.0026568048	0.0010621736	0.0002127282
$MSE(\hat{\beta}_{\text{not employed}})$	0.0017520438	0.0009080421	0.0002696788	0.0000709569
$MSE(\hat{\beta}_{\text{married living together}})$	0.0054620948	0.0024377131	0.0006165484	0.0001583126
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0359318615	0.0115208465	0.0032639055	0.0008976778
$MSE(\hat{\beta}_{\text{widowed}})$	0.0118854566	0.0047788168	0.0015703984	0.0003322932
$MSE(\hat{\beta}_{\text{single}})$	0.0063128172	0.0026525454	0.0008721015	0.0002047679
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0095949533	0.0048627868	0.0012408221	0.0002995102
$MSE(\hat{\beta}_{\text{SPD}})$	0.0117203007	0.0051882874	0.0013938861	0.0002574235
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0097938571	0.0054435135	0.0017723805	0.0003103283
$MSE(\hat{\beta}_{\text{die linke}})$	0.0110373558	0.0068738564	0.0017512232	0.0003130554
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0149759706	0.0070284133	0.0019085982	0.0004188738
$MSE(\hat{\beta}_{\text{FDP}})$	0.0157995717	0.0078989014	0.0022400439	0.0004134289
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0140761716	0.0061615106	0.0021898871	0.0003425453

Table 2.3: MSE of regression coefficient estimates in order to quantify sampling error

With increasing sample size the MSEs of the regression coefficient estimates become smaller which corresponds to the findings of table (2.2) and confirms that with larger samples we can have more confidence in the sample's representativeness. Viewing the average values of the beta coefficients resulting from the different samples gives that they slightly differ from the true values. In theory, the expected values of the coefficients are not affected by changes in sample size, but actually the coefficients differ because of the sampling variation which comes along with different samples. This can be seen in the following figure, representing exemplary the distribution of the estimated intercepts for the different sample sizes.

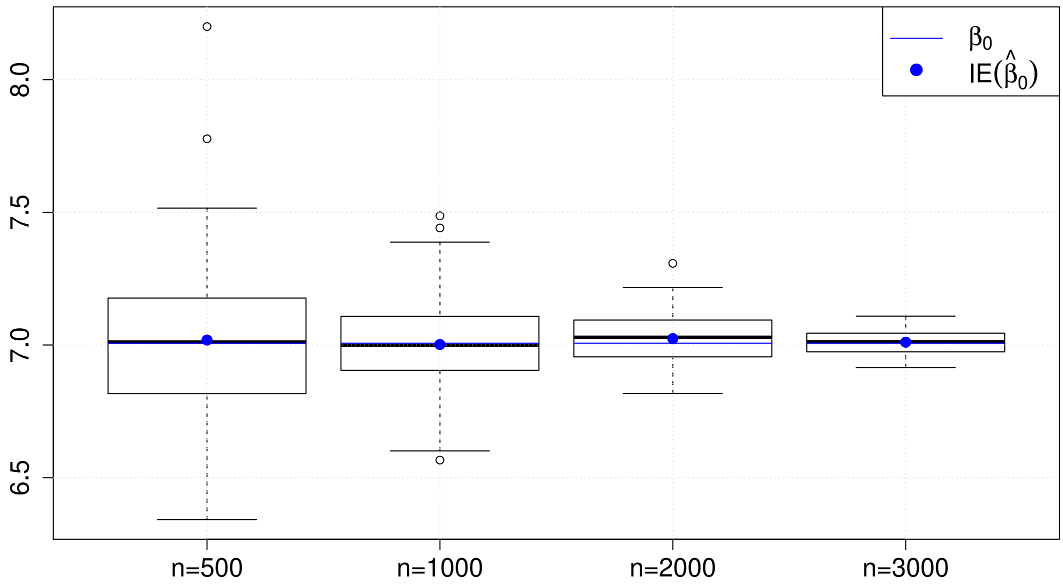

Figure 2.4: Boxplots of intercept estimates for different sample sizes

Variation is obviously systematically influenced by the sample size, which means that it becomes smaller for increased n. There is not such a systematic trend in the estimated average of the intercept. Thus, for certain sample sizes $\mathbb{E}(\hat{\beta}_0)$ the true value of $\beta_0 = 7.0070608279$ is slightly underestimated, in others cases this value is overestimated. However, the bias of estimating β_0 is small for all sample sizes and takes a maximal value of 0.0171504762 for $n = 2000$. These findings can also be assigned to the other beta coefficients, whereas the sample size of $n = 2000$ does not always yield the worst result concerning bias. For reasons of clarity the boxplots of the remaining coefficient estimates are placed in appendix B.

As a last step of interpretation, the change in significance of statistical effects due to sampling error is analyzed. Conclusions based on significance can be ascribed to a test with null hypothesis implying that the corresponding regression coefficient is equal to zero under the assumption of normally distributed error terms. A test statistic can be calculated which is based on the value of the estimated regression coefficient, as well as its variance. The null hypothesis is then rejected at a certain significance level α, if the test statistic has a larger value than the corresponding $(1-\frac{\alpha}{2})$-quantile of the t-distribution. The corresponding degrees of freedom are calculated by the difference of sample size and number of estimated parameters. In theory, the crucial test statistic becomes smaller, if the variance of the estimate increases. This is the case, when sample size decreases. Then there is a tendency to that effect that the null hypothesis, indicating that there is no significant effect, cannot be rejected. All in all, when sample size decreases, the corresponding p-values tend to become larger and effects tend to become non significant.

Viewing the average p-values, resulting from the regression models of the different sample sizes, which are presented in table (C.7), this continuous increase of the p-values can be recognized for all statistical significant effects. While for example the group of participants that are “married living together” in the true underlying sample have a significant higher logarithmized income than the reference group of “divorced” participants, this effect remains at the same level of significance for the sample size of $n = 3000$, namely smaller 0.01, but then changes for smaller sample sizes. For $n = 2000$ the corresponding p-value is only smaller 0.1 and for $n = 1000$ and $n = 500$ this effect is not significant anymore. In contrast to that, the highly significant effects of the intercept, “age”, “female”, “half-time”, “part-time” and “not employed” remain highly significant for all sample sizes. Also the non significant effects of “fachhochschulreife”, “hochschulreife”, “widowed”, “CDU-CSU”, “die linke”, “extreme right-wing” and “SPD” do not become significant, when drawing samples. In these cases, the corresponding p-values increase for larger sample sizes and thus approach to the true underlying p-values of those regression coefficients.

In conclusion, sampling error can not only affect the values of the estimated regression coefficients, but also the significance of statistical effects, which are attenuating with decreasing sample size.

Coverage Error

As a next step, focus is set on the impacts of coverage error and as before, it first of all lies on the estimates for the income mean μ_{income}:

	$n = 500$	$n = 1000$	$n = 2000$	$n = 3000$
μ_{income}	1492.951	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1289.939	1304.254	1352.104	1438.092
$Bias(\hat{\mu}_{\text{income}})$	-203.01198	-188.69738	-140.84656	-54.85946
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.13598	-0.12639	-0.09434	-0.03675
$Var(\hat{\mu}_{\text{income}})$	1880.28475	596.53409	194.77684	54.30150
$MSE(\hat{\mu}_{\text{income}})$	43094.14877	36203.23531	20032.53030	3063.86149

Table 2.4: Expected value, variance, bias and MSE of mean estimation of income in order to quantify coverage error

In all four cases of n, the true population income mean of 1492.951 is clearly underestimated. The averages of the samples deviate up to an absolute value of 203.01198 from the true value in the case of $n = 500$. Those deviations can be traced back to the different average incomes of men and women in the target population. Whereas the average female income in the simulated data set comes to 1118.066, those of men reaches a value of 1869.554. Since women are purposely overrepresented in these samples, the average income is underestimated. Apparently, a larger sample size decreases both, bias and variance of the estimations, which leads to considerably smaller MSEs. With increased sample size the corresponding MSEs decrease remarkably from a value of 43094.14877 in the case of $n = 500$ to 3063.86149 in the case of the largest sample size.

Before viewing the boxplots, which complete what was said before, a short statement about the corresponding t-tests that have already been conducted in the case of sampling error, is made. Here, the null hypothesis of equality between true and estimated mean income value could be rejected in the majority of the 100 cases for $n = 500, n = 1000, n = 2000$. Therefore, in these scenarios we can assume significant deviations. Concentrating then on the corresponding boxplots shows that all four boxes are located markedly below the blue line, denoting the true population mean. Due to the larger variation in smaller samples the whiskers of the boxplot in the case of $n = 500$ include the true value. However, the estimated sample means, denoted by the blue dots, are far from the true value, but approaching the true population mean with increasing n.

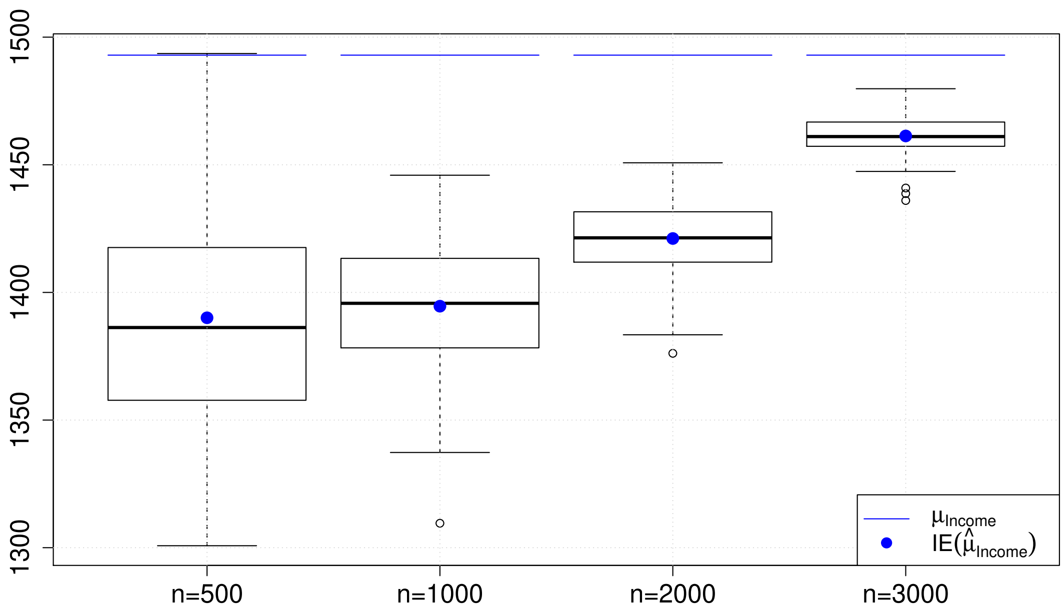

Figure 2.5: Boxplots of income mean estimates for different sample sizes containing coverage error

In order to get an overview of the impacts on regression parameter estimation, first the abbreviated table of MSEs is discussed:

	$n = 500$	$n = 1000$	$n = 2000$	$n = 3000$
$MSE(\hat{\beta}_0)$	0.1746848145	0.0759675696	0.0241422798	0.0043063296
$MSE(\hat{\beta}_{\text{Age}})$	0.0000032080	0.0000014495	0.0000004963	0.0000000899
$MSE(\hat{\beta}_{\text{Female}})$	0.0030964218	0.0010819041	0.0002826789	0.0000618589
$MSE(\hat{\beta}_{\text{No Graduation}})$	0.2175205198	0.0954216295	0.0233456802	0.0050660246
$MSE(\hat{\beta}_{\text{Volks-, Hauptschule}})$	0.1419779418	0.0691553314	0.0158626008	0.0031987587
$MSE(\hat{\beta}_{\text{Mittlere Reife}})$	0.1328235799	0.0654806901	0.0163393849	0.0031121287
$MSE(\hat{\beta}_{\text{Fachhochschulreife}})$	0.1384112213	0.0673351515	0.0162798288	0.0032720538
$MSE(\hat{\beta}_{\text{Hochschulreife}})$	0.1381689489	0.0659550599	0.0167554420	0.0033723480
$MSE(\hat{\beta}_{\text{half-time}})$	0.0030096672	0.0010350722	0.0003176226	0.0000465588
$MSE(\hat{\beta}_{\text{part-time}})$	0.0059792038	0.0020730085	0.0005645777	0.0001427316
$MSE(\hat{\beta}_{\text{not employed}})$	0.0020576189	0.0008079692	0.0002343462	0.0000777910
$MSE(\hat{\beta}_{\text{married living together}})$	0.0059394476	0.0027431138	0.0008551152	0.0001597583
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0314244963	0.0158825772	0.0050882673	0.0005327739
$MSE(\hat{\beta}_{\text{widowed}})$	0.0114055652	0.0065996384	0.0021758343	0.0003362787
$MSE(\hat{\beta}_{\text{single}})$	0.0066799580	0.0030803743	0.0009767531	0.0001433286
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0087138946	0.0034270689	0.0014494486	0.0004083341

$MSE(\hat{\beta}_{\text{SPD}})$	0.0099441068	0.0039823616	0.0014282208	0.0004098157
$MSE(\hat{\beta}_{\text{Die Gruenen}})$	0.0096541134	0.0043957662	0.0017913425	0.0004791308
$MSE(\hat{\beta}_{\text{Die Linke}})$	0.0128880915	0.0045721882	0.0021619391	0.0005436597
$MSE(\hat{\beta}_{\text{Extreme Right-Wing}})$	0.0145632891	0.0058501670	0.0016206209	0.0007477504
$MSE(\hat{\beta}_{\text{FDP}})$	0.0180285383	0.0077475888	0.0026027662	0.0005778129
$MSE(\hat{\beta}_{\text{Would not vote}})$	0.0140062628	0.0048118330	0.0025309676	0.0006097294

Table 2.5: MSE of regression coefficient estimates in order to quantify coverage error

Obviously, the estimates for all regression coefficients become more representative with increased sample size, which is reflected through smaller MSEs. The corresponding expected values and variances of the estimated regression coefficients are listed in the full table (C.8) in the appendix and are illustrated with the help of boxplots (see (B.5), (B.6), (B.7), (B.8)). Due to the underrepresentation of certain groups of participants, regression coefficient estimation should be viewed critically. In the samples of size 500 for example, we only have on average 2.620 observations with education category "other graduation", which is the reference category of this variable in the regression model. This obviously cannot result in good estimates and comparability. Interpreting table (C.8) further gives that in contrast to the MSEs, the corresponding biases of regression coefficients do not decrease consistently with increased sample size. But in nearly all cases of regression coefficient estimation, the true value is consistently over- or underestimated through the average expected value of the different sample sizes. This statement is for example valid for the intercept estimation. As the boxplots in figure (2.6) show, the expected values of the estimates in all samples are close to the true population value, which is only slightly overestimated. In all four cases, the true value lies within the boxes, in other words between the 25%- and the 75%-quantile. Thus, the box contains 50% of the data, which here means half of the mean estimates, respectively.
Finally, regarding the significances of the different effects reveals that highly significant results stay highly significant independent of the size of the sample and so do effects which are not significant in the target population. In the other cases there can be seen a tendency of decreasing p-values (cp. table (C.9)) and consequently a trend towards less or rather not significant effects for smaller sample sizes. Summarizing those findings gives that coverage error has more serious impacts on parameter estimation than sampling error, if the sampling frame noticeably differs from the target population, which is obviously the case in the present situation. Then both, mean and regression coefficient estimation are affected.

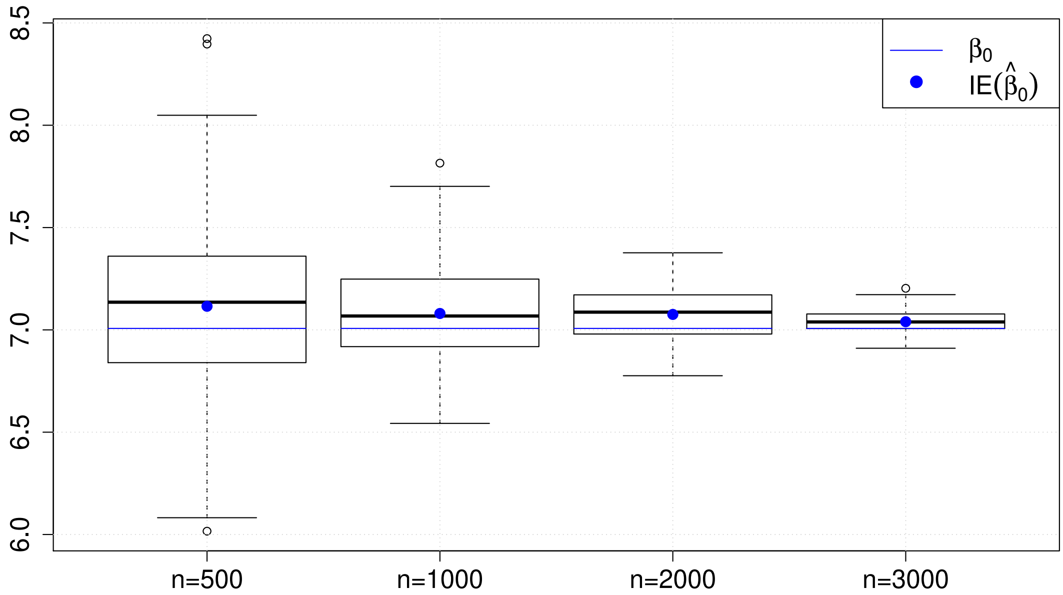

Figure 2.6: Boxplots of intercept estimates for different sample sizes containing coverage error

Nonresponse Error

In the case of nonresponse error it is not distinguished between different sample sizes, but between the three dropout mechanisms. First, the impacts of MCAR, MAR and MNAR nonresponse errors on the estimation of the income mean are presented:

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1492.019	1544.401	1529.321
$Bias(\hat{\mu}_{\text{income}})$	-0.93241	51.44947	36.36994
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00062	0.03446	0.02436
$Var(\hat{\mu}_{\text{income}})$	124.98964	232.82946	198.20299
$MSE(\hat{\mu}_{\text{income}})$	125.85902	2879.87698	1520.97589

Table 2.6: Expected value, variance, bias and MSE of mean estimation of income in order to quantify nonresponse error

In the case of MCAR the average mean of 1492.019 is a very good estimate for the true population mean of 1492.951. Consequently, the bias is small, so are the variance of the estimates and the resulting MSE of 125.85902 in this case. Thus, if there exists a nonresponse error, with a dropout that is completely at random, the situation can be compared to a random sample, since we are only observing some of the values. As a result, mean estimation is not influenced and the outcomes can be transferred to the target population. This is not equally the case in the second scenario, where the MAR mechanism is present. The estimated income mean of 1544.401 obviously is worse compared to the MCAR mean. It differs from the true value by an absolute amount of 51.44947 which brings along a noticeable overestimation of the true population mean. The MSE of 2879.87698 quantifies this large difference between the estimate and what is estimated. Consequently, the mean estimate in this situation tends not to be very representative. Finally, the impacts of the MNAR nonresponse mechanism can be classified close to that of MAR, but not that bad. Its mean estimation of 1529.321 is closer to the true population income mean, since now people with higher income value have a tendency to dropout more likely. Therefore, the income mean of the remaining responders is pushed down which leads to a smaller overestimation compared to the MAR scenario. Moreover, in the presence of MNAR the variance of the estimate is smaller which brings along that the MSE becomes smaller compared to the MAR case and reaches a value of 1520.97589. All those findings are again illustrated with the help of boxplots.

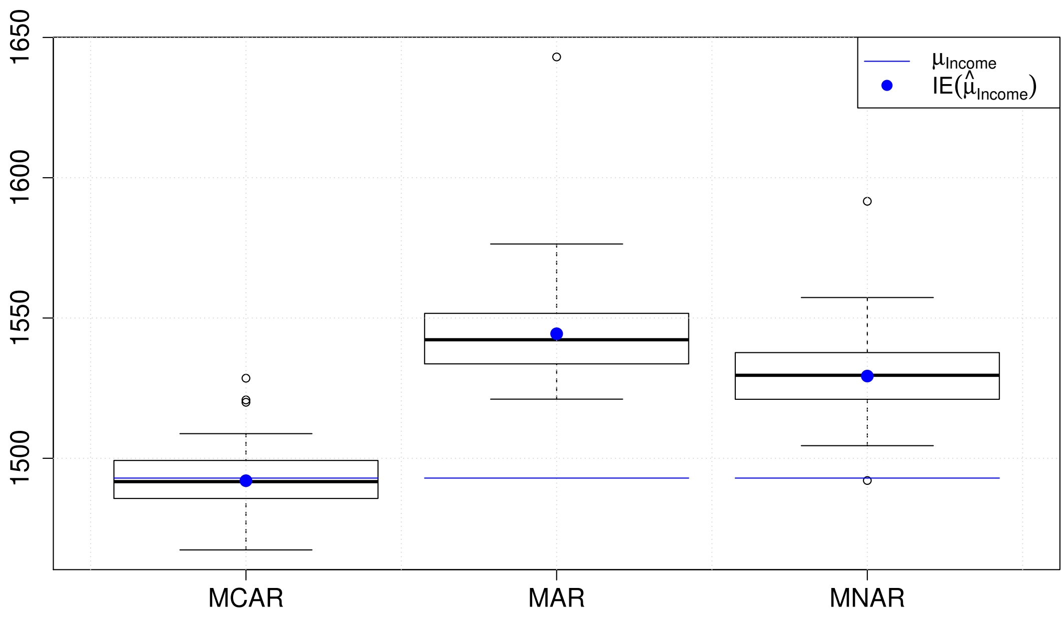

Figure 2.7: Boxplots of income mean estimates for different nonresponse mechanisms

Those statistical key figures show that the dropouts which are MAR or MNAR obviously affect mean estimation, even though these deviations from the true income mean value are in the majority of cases not significant according to t-tests. Nevertheless, the average range of the corresponding bias is not small, which is why the dropouts may not be ignored, but should be handled in a suitable way.
In order to find out, whether the different missing data mechanisms have a similar influence on regression parameters as they had on mean estimation, focus is first set on the statistical measures corresponding to the regression intercept. Here we see very similar estimates for MCAR, MAR and MNAR. All three average estimates are close to the true value of 7.0070608279 (see figure (2.8)), whereas it is noticeable that in the MCAR case we still have the smallest bias, which is represented by the distance between the true value line and the estimated value dot in color blue, the smallest variance, represented by the width of the box and consequently the smallest MSE, consisting of those two components.

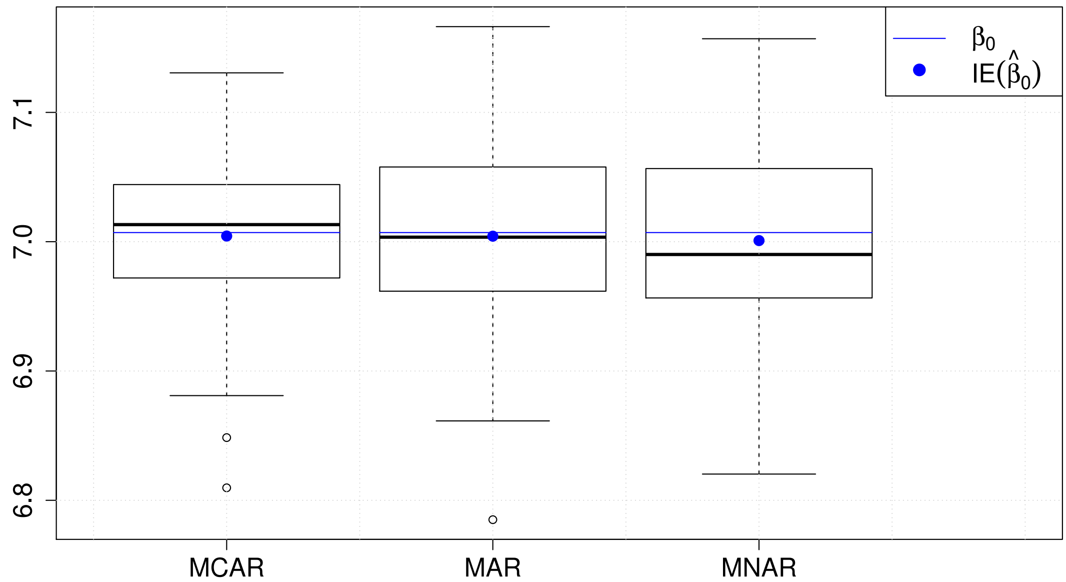

Figure 2.8: Boxplots of intercept estimates for different nonresponse mechanisms

Considering the remaining regression coefficients, it eventuates that in all the cases, except $\beta_{\text{hochschulreife}}$, the MSE is smallest, when the nonresponse dropout is completely at random (see table (2.7)). This confirms the findings that MCAR dropout does not have a strong impact on parameter estimation and attention should be focused on MAR and MNAR. However, the more detailed table (C.10) in the appendix reveals that for all three nonresponse mechanisms the bias of the estimated regression parameters is small. Thus, the significances do not change heavily, as table (C.11) shows. Strongly significant effects stay strongly significant for all nonresponse mechanisms and not significant effects do not become significant. However, effects that are significant at a higher significance level in the true underlying data situation are obviously weakened, when considering the regression outputs of the data with nonresponse error. However, except in the case of "die gruenen" all significant effects are also recognized as statistically significant in the presence of nonresponse. While the p-values in the MAR and MNAR scenarios make not much of a difference, the one of MCAR in general are smaller and therefore closer to the true p-values, if effects are significant.

	MCAR	MAR	MNAR
$MSE(\hat{\beta}_0)$	0.0042048848	0.0050811213	0.0044771267
$MSE(\hat{\beta}_{\text{age}})$	0.0000001380	0.0000002427	0.0000002409
$MSE(\hat{\beta}_{\text{female}})$	0.0000938447	0.0001526323	0.0001763245
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0039788141	0.0048166650	0.0050870244
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0029842714	0.0034675059	0.0033753390
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0030929637	0.0032046833	0.0031922267
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0033685208	0.0034415262	0.0035441220
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0031050359	0.0031316799	0.0030277848
$MSE(\hat{\beta}_{\text{half-time}})$	0.0002298237	0.0002937056	0.0003132581
$MSE(\hat{\beta}_{\text{part-time}})$	0.0003782038	0.0004203327	0.0005391832
$MSE(\hat{\beta}_{\text{not employed}})$	0.0000947949	0.0001578360	0.0001497775
$MSE(\hat{\beta}_{\text{married living together}})$	0.0002629782	0.0003101631	0.0004059300
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0013381430	0.0014205419	0.0015812653
$MSE(\hat{\beta}_{\text{widowed}})$	0.0004501837	0.0008222669	0.0010445201
$MSE(\hat{\beta}_{\text{single}})$	0.0003023645	0.0005154555	0.0005888173
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0004781217	0.0006275838	0.0006547944
$MSE(\hat{\beta}_{\text{SPD}})$	0.0005395279	0.0006281904	0.0006324777
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0004821276	0.0006507621	0.0007183523
$MSE(\hat{\beta}_{\text{die linke}})$	0.0006105415	0.0007766607	0.0008841787
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0005917100	0.0009249487	0.0011073786
$MSE(\hat{\beta}_{\text{FDP}})$	0.0009082497	0.0009917881	0.0010519991
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0006573288	0.0011008500	0.0011391278

Table 2.7: MSE of regression coefficient estimates in order to quantify nonresponse error

Finally it is important to note that the quality of those estimates is massively influenced by the number of missing values in the data set. The more missing values a variable has, the worse is its representativeness. The number of missing values in the income variable does not coincide in the three cases, since the probability for a dropout is different for a certain observation, respectively. Whereas in the situation of MAR the number of missing values within the 100 iterations varies between a minimum of 962 and a maximum of 1094, the respective values of the MNAR process are 1017 and 1138. The corresponding median of the MAR process is given by 1045, those of the MNAR process by 1091. The number of completely random missings has been orientated on the Allbus data set and is with a total number of 750 much smaller. The higher this number is chosen, the more converges the corresponding MSE to those MSEs of MAR and MNAR. However, even in the case of 1000 completely random dropouts, only a MSE of 175.2621 is received for the income estimate. Then, the estimates for the MCAR case become worse and the impacts of the underlying nonresponse error more remarkable, but the consequences are nevertheless not comparable to those of MAR and MNAR.

Measurement Error

After application of the three measurement error models to the simulated data set, a large number of rounded income values is received in the case of RAR and RNAR. While the 100 data sets with rounding behavior at random have between 3113 and 3208 values that are divisible through 10, this amount varies between 3106 and 3202 in the RNAR case. However, the RCAR model takes the percentages of rounded values of the Allbus data set, as has been described before. Thus, in each of the iterations it displays a total number of 2513 income values that are divisible through 10. Obviously, this number widely differs from those of the other models which has to be kept in mind, when comparing the results in the following.

	RCAR	RAR	RNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1493.299	1493.406	1493.407
$Bias(\hat{\mu}_{\text{income}})$	0.34795	0.45484	0.45603
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	0.00023	0.00030	0.00031
$Var(\hat{\mu}_{\text{income}})$	0.00056	0.00023	0.00022
$MSE(\hat{\mu}_{\text{income}})$	0.12163	0.20711	0.20818

Table 2.8: Expected value, variance, bias and MSE of mean estimation of income in order to quantify measurement error

Estimating the mean of the income gives good results for all three error models as the following graphical illustration confirms.

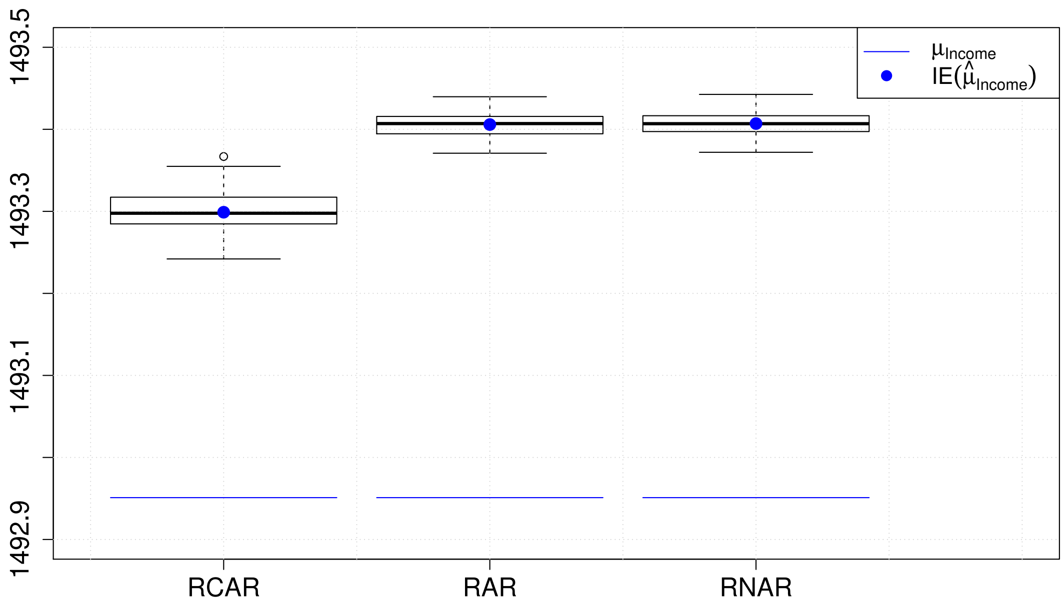

Figure 2.9: Boxplots of income mean estimates for different measurement mechanisms

The average income value of the 100 data sets, respectively, only slightly overestimates the true population mean. However, those differences are statistically not significant and are only extended to differences in the decimal place which is reflected through the small range of values on the y-axis of the boxplots. Here also very small variances are noticed, which means that the estimation of the population mean does not vary much in the different data sets containing the simulated measurement errors. This fact can be explained by the large number of rounded values, which is why the income values of the different data sets and consequently also their corresponding means do not differ much. Even though the percentage of roundings is decreased, the variances would stay small, since it is only rounded to the next threshold of 10, which either means an additional small positive or a small negative change of the income value, which finally does not influence the mean of the income variable noticeably. Due to those very small variances of the estimates, also remarkable small MSEs result for all three models. The RCAR model results in the smallest bias and, despite the largest variance, in the smallest MSE. However, if the number of income values that are rounded to the next threshold of 10 is increased close to the total number that the other two models provide (e.g. additional 2800 rounded values instead of 2200), then statistics approximate. The MSE of the RCAR model then is close to those of RAR and RNAR, which are nearly the same. The findings of the regression coefficient estimation coincide with those of the mean estimation in all aspects. The three resulting MSEs of the different rounding scenarios are very similar, whereas the one of RCAR in most cases is the smallest. However, those differences are not considerable and vanish, if the number of rounded values for RCAR is increased. Summarizing, the estimated regression coefficients in all three scenarios are unbiased which confirms the following, theoretical findings of Schneeweiss et al. (2010) before: In the underlying regression application, the response variable income has rounded values, whereas all the explanatory variables are not rounded. In this situation the values of the estimated regression parameters are not affected. Again, the intercept boxplots are brought in as an example.

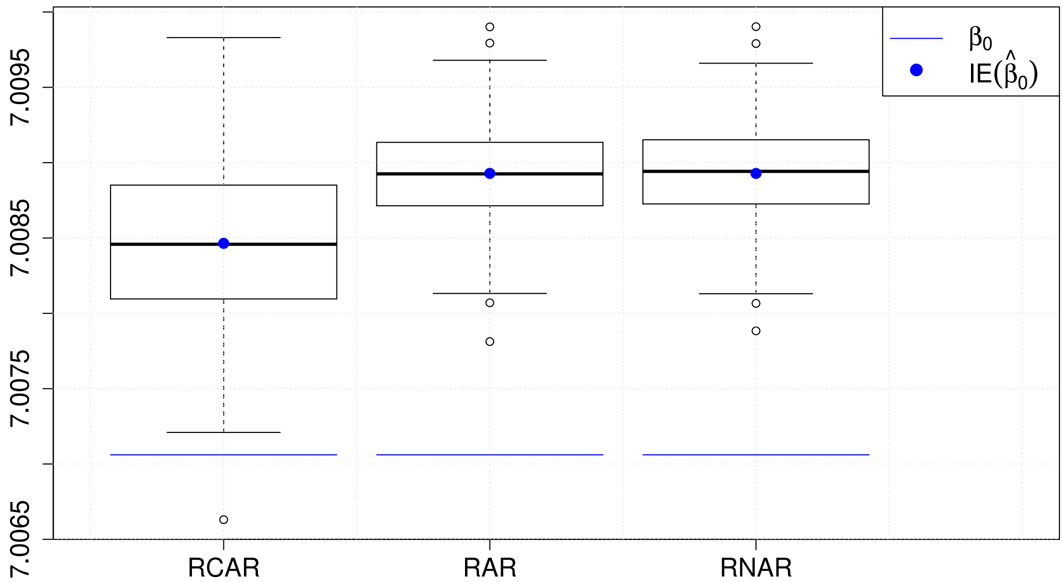

Figure 2.10: Boxplots of intercept estimates for different measurement mechanisms

The boxplots of the remaining regression coefficients in appendix B display even smaller ranges, since the corresponding variances only take values between 10^{-07} and 10^{-12}. As a consequence, the estimated values, represented by the dots, always lie on the line, representing the true values. Those good estimates of the regression coefficients are confirmed by the following table (C.12), containing the corresponding MSEs. Thus, it can be concluded that the range of values of the regression coefficients does not change noticeably due to these kind of errors. Also the significances of the regression effects are not influenced as table (C.13) provides.

	RCAR	RAR	RNAR
$MSE(\hat{\beta}_0)$	0.0000022895	0.0000036333	0.0000036287
$MSE(\hat{\beta}_{\text{age}})$	0.0000000000	0.0000000000	0.0000000000
$MSE(\hat{\beta}_{\text{female}})$	0.0000000058	0.0000000023	0.0000000023
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0000002461	0.0000000917	0.0000000908
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0000002079	0.0000001121	0.0000001104
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0000004746	0.0000005191	0.0000005157
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0000003686	0.0000003532	0.0000003495
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0000004555	0.0000005107	0.0000005064
$MSE(\hat{\beta}_{\text{half-time}})$	0.0000000519	0.0000000703	0.0000000705
$MSE(\hat{\beta}_{\text{part-time}})$	0.0000001699	0.0000002322	0.0000002324
$MSE(\hat{\beta}_{\text{not employed}})$	0.0000000225	0.0000000279	0.0000000281
$MSE(\hat{\beta}_{\text{married living together}})$	0.0000003167	0.0000005033	0.0000005047
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0000005364	0.0000008253	0.0000008284
$MSE(\hat{\beta}_{\text{widowed}})$	0.0000002367	0.0000003829	0.0000003835
$MSE(\hat{\beta}_{\text{single}})$	0.0000003116	0.0000005002	0.0000005027
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0000001028	0.0000001625	0.0000001636
$MSE(\hat{\beta}_{\text{SPD}})$	0.0000000488	0.0000000632	0.0000000634
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0000001028	0.0000001625	0.0000001636
$MSE(\hat{\beta}_{\text{die linke}})$	0.0000000563	0.0000000872	0.0000000877
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0000001024	0.0000001827	0.0000001822
$MSE(\hat{\beta}_{\text{FDP}})$	0.0000000985	0.0000001663	0.0000001665
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0000000679	0.0000001302	0.0000001310

Table 2.9: MSE of regression coefficient estimates in order to quantify measurement error

Drawing a conclusion for the first part of the thesis, it contains that in this application, nonsampling error mostly causes a considerably larger bias compared to sampling error and therefore has stronger impacts on the considered parameter estimation. Thus, when conducting a survey, at least as much attention as focused on reducing sampling error should be spend on the other components of the TSE.

Considering the impacts of the different nonsampling errors, it shows clearly that coverage error has strong effects on parameter estimation, if the sampling probabilities differ noticeably from simple random sampling. In contrast to that, measurement error, which is analyzed here in the specific form of rounding error, yields small deviations between the true parameter values and the corresponding estimates, in fact for all different rounding processes RCAR, RAR and RNAR. This is different in the scenario of nonresponse error. Here, the underlying dropout process is crucial. While MCAR hardly influences parameter estimation, estimates are biased in the case of systematic dropouts including MAR and MNAR.

This finding forms the basis of the second part of the thesis which focuses on nonresponse error as a main error component of the TSE. Nonresponse error, which results in missing data, is an omnipresent problem in survey statistics, since nearly each survey contains missing values, which has to be dealt with in an appropriate way. However, there is no gold standard approach to handle missing data, but a wide range of methods. Often analysts are overcharged with the task of finding the right way to handle the missing data situation, which is why they go back to traditional approaches. Those, however, simply ignore the missing values or are based on assumptions that are not justified. The second part of the thesis shall present the different types of methods including their limitations and drawbacks, shall establish the advantages of modern approaches and apply them to compare the performances. Of course, all those theories and analyses shall be embedded in the context of the TSE approach, in order to associate the two parts of the thesis.

3 Nonresponse as a Main Error Component of the Total Survey Error Approach

3.1 Minimizing Impacts of Nonresponse Error in Survey Data

Focusing on the nonresponse error as a major component of the TSE, two important features arise that aim to minimize the impacts of this source of error. The first family of approaches sets out from a point before or during the data collection process and involves minimizing the impacts of nonresponse error by keeping nonresponse in the survey data at a minimum. This section will deal with those attempts before focusing then on methods whose aim is to correct or improve parameter estimation in the case of present nonresponse in the data (see section 3.2). While concentrating on all those approaches in the field of nonresponse, it may be taken into account that this error is just one major component of the TSE and focusing only on the nonresponse error might have impacts on the other components, since in some cases there is a link between the different errors. Therefore, efforts to minimize nonresponse error can have the consequence of an increased overall TSE, because the impacts of other error sources increased simultaneously. This will become clear at certain points of this chapter.
Minimizing nonresponse error is often put on the same level with increasing response rates. However, response rates per se do not necessarily obtain information about the corresponding bias, which comes along with the dropout. Thus, response rates do not inform about the underlying missing mechanism and consequently about the differences between responder and nonresponder, which is why, literature concerning this field of research suggest alternative measures (cp. Kreuter (2013)), depending on auxiliary information about responders and nonresponders. With additional knowledge about the dropout in terms of paradata or auxiliary variables it is possible, to better asses data and survey quality and moreover, appropriate methods for handling the missing data can be found (see section 3.2).
Also linking surveys to administrative data or combining information from different surveys might help to better understand the reasons, why people do not respond. Often, the cost-benefit equation of taking part in a survey does not seem profitable for the selected people. Then, on the one hand, benefits should be maximized, for example in terms of incentives (cp. Pforr et al. (2015)). This procedure can moti-

vate people to participate in the survey, but often the participation is only based on the promised incentive which might lead to sloppy and non reliable answers. Then, the bias due to measurement error can increase. On the other hand burdens for participants should be minimized. This can for example be achieved by so called multiple-matrix-sampling which is a technique that aims to reduce the number of questions for each participant by randomly dividing the questionnaire in subtests. Although each participant receives only a proportion of the complete set of items, the estimated parameters of interest are equivalent to those obtained by testing all participants on all items. Another approach for reducing respondent burdens are mixed-mode surveys. In these surveys different modes like face-to-face interviews or self-administered designs are provided, from which the participants can choose their favorite procedure. Staying in this context, in general choosing the mode of data collection can have obvious impacts on nonresponse error. Thus, self-administration reduces measurement error, but increases nonresponse (cp. Sakshaug et al. (2010)), compared to designs based on interviewers. But of course, also the interviewer himself can be the source of nonresponse error, which is why, effort should be made in the field of interviewer trainings (cp. West and Olson (2010), O'Brien et al. (2006)). Another source for nonresponse error might be the questionnaire. Bad wording of questions for example can cause dropouts and should therefore be identified, in order to avoid the resulting bias (cp. Kreuter et al. (2008)). For this purpose small pilot studies before the start of the main survey might help to identify unexpected problems (cp. Van Teijlingen and Hundley (2002)). Finally, a last obvious approach for minimizing the nonresponse in a survey is to increase the effort of contacting the selected people (cp. Kreuter et al. (2010), Fricker and Tourangeau (2010)). This way, non-contact bias as a part of nonresponse bias becomes smaller and moreover larger response rates decrease sampling error and desirably improve the coverage of the population of interest. Yet, putting pressure onto selected people might have the negative effect of resulting in worse data quality resembled through larger measurement errors.

Concluding, all those described approaches are just a subset of possibilities which can help to minimize nonresponse error a priori. Obviously, they can also have negative effects on other TSE components and so the pros and cons of each attempt have to be balanced, in order to decide, where effort in terms of money and time is useful to spend. However, in general all those attempts do not completely avoid the dropout of people, which is why, a strong need for methods to handle those missing data situations in the field of surveys arises.

3.2 Correction Methods for Nonresponse Error and their Limitations

When dealing with incomplete data, there are several approaches. According to Little and Rubin (1987) they can generally be divided into different groups of methods:

- Methods based on the available data (ad-hoc methods)
- Methods based on weighting the observed cases (weighting)
- Methods based on replacing the missing values (imputation)
- Methods based on maximum-likelihood (maximum-likelihood procedures)

3.2.1 Ad-Hoc Methods

One of the oldest and without a doubt easiest approaches to deal with missings in survey data are ad-hoc methods which are based on analyzing the available data and in some way ignoring the missings (cp. Weisberg (2005), Spieß (2008) or Little and Rubin (1987)).
The first method within this field is the so called complete-case-analysis, often also referred to as listwise deletion. Here, all observations, having missing data on any of the variables, are discarded and it is then proceeded with the analysis of interest using standard methods. The advantage of this approach is the simplicity. No expertise is necessary and application in all standard software is straightforward. Moreover, this method results in a single data set with given size, which makes univariate statistics comparable. However, those advantages are minor, when considering on the one hand the huge loss of information and on the other hand the resulting estimates which are in general biased, except when the underlying missing mechanism follows MCAR. In this case, the complete cases are representative for the whole sample and the estimates will be unbiased. Then this procedure only becomes a problem, when too many variables have missings so that the sample size becomes too small. Since it is not rational for univariate analyses, to discard values of a particular variable, when they belong to cases that have missings in other variable, another ad-hoc procedure arose.
The available case analysis, often called pairwise deletion, uses all values that have been observed for the relevant variables of a specific research question. As a consequence, the sample base for different analyses changes. This disadvantage affects

comparability and the estimation of standard errors, which requires a specific sample size (cp. Graham (2012)).
Both procedures should be viewed as critical and their application should only be accepted, when dropouts are following MCAR and the number of missings is small. However, even if those conditions are fulfilled, Graham (2012) stated that there are modern methods which bring along at least as good results, in general even results that are better than simple ad-hoc procedures. In order to evaluate this statement and to demonstrate the limitations of ad-hoc methods, they should be briefly applied to the simulated data. The results can be seen in section 3.5. At this point it should be hinted that the different, simulated data sets only have missing values in the income variable. Therefore, the two approaches, complete-case-analysis and available-case-analysis, here are equal and produce same results concerning mean and regression coefficient estimation, which is why the performance of ad-hoc procedures in chapter 3.5 shortens to a single application.

3.2.2 Weighting

The second approach in general is used to compensate for unit nonresponse in surveys, but it is also applied for item nonresponse as in the underlying data situation of the thesis. It involves attaching weights to each subject, included in the analysis, to represent those who were excluded due to missing values. This way, the calculated statistics based on the sample become more representative and bias is reduced. However, as the reasons of nonresponse are complex and individual, bias obviously cannot be eliminated through this approach.
In the case of poststratification weighting, which contains that the different units are assigned their weights after the process of data collection based on external information, the survey sample is adjusted towards the underlying population proportions along a small number of dimensions represented by several variables (cp. Weisberg (2005)). Of course, this approach works best, if those attributes are strongly related with the variable of interest. If those variables are not predictive for the variable of interest, having missing values, bias is not reduced due to weighting, but variance increases. There are different ways to construct the weights, but here it is focused on the adjustment-cell method, where the weighting variables form the so called adjustment cells. Lets assume there are l of them, consisting of responders and nonresponders, respectively, whereas weights based on the inverse of the probability of

selection and response (cp. Little and Rubin (1987)) are attached to the responders in each group, to compensate for the nonresponders in the same group. Thus, let p_i^{sample} be the probability for unit i to be sampled from the population. When viewing a situation based on probability sampling, usually p_i^{sample} is equal for all i and can be determined due to sample size and population size. Let furthermore p_i^{respond} be the probability for individual i to respond, when i has been sampled. This probability in general is not known and has to be estimated. Considering that, constant response probabilities for each unit i within an adjustment cell l are assumed (cp. Raghunathan (2004)), as well as a dropout that is completely at random within the different adjustment cells. This results in weights $w_i = \frac{1}{p_i^{\text{sample}} \cdot \hat{p}_i^{\text{respond}}}$, whereas the weights are equal for all units of an adjustment cell l. Those weights w_i then can be incorporated in the analyses of interest. Thus, for a variable Y, for which not all n units have recorded values, the corresponding weighted mean is calculated by $\frac{1}{n}\sum_{i=1}^{n} y_i \cdot w_i$ and in the context of linear regression the weighted least squares $\sum_{i=1}^{n} w_i(y_i - x_i^T\beta)$ have to be minimized instead of the residual sum of squares.

3.2.3 Imputation

One of the most common approaches for handling missing data are imputation methods. The idea behind is to replace the missing values with plausible values before applying standard statistical analyses. The observed values of the data set form the basis for a predictive distribution from which the imputed values are derived. Obviously, there are different approaches for choosing those values from the predictive distribution which leads to a classification of the imputation methods.

First of all, it has to be differentiated between single and multiple imputation. Whereas single imputation methods are based on the idea that each missing value is replaced by one single value before standard statistical analyses are applied without modification, multiple imputation implies the replacement of each missing value by $T \geq 2$ imputed values. This equals a repeated draw from the predictive distribution. Then statistical analyses are conducted at each resulting data set, which differ only with respect to the imputed values. In a last step, these results are then combined to form one inference. In the literature there can be found different ways to combine those statistics, but the original combination rule that has been adapted over the years, goes back to Rubin Donald (1987). It is based on asymptotic theory, more precisely, it assumes that inferences about θ, the population parameter of interest, can be based on normal distribution (cp. Reiter and Raghunathan (2007)). When the normality assumption appears inappropriate for estimates of the parameters of interest, suitable adaptions should be considered or alternative, robust measures for combinations, such as the median should be applied. Since here, it is focused on the population mean and the regression coefficients for θ, Rubin's rule is appropriate, since those parameters are asymptotically normal, at least for larger samples (cp. Marshall et al. (2009)), which is given in the present scenario. Therefore, later analyses rest upon this combination rule which is also used in the majority of the literature (cp. Marshall et al. (2009)) and therefore presented in detail.

Let $\hat{\theta}^{(t)}$ be the corresponding estimate in data set $t = 1, \cdots, T$ for the population parameter of interest. Then the final point estimate is given by the average of the T estimates, thus by $\hat{\theta} = \frac{1}{T}\sum_{t=1}^{T}\hat{\theta}^{(t)}$. The associated variance of this point estimate results from a combination of within-variance and between-imputation-variance. $\widehat{Var}_{\text{within}} = \frac{1}{T}\sum_{t=1}^{T} Var_{\text{within}}{}^{(t)}$ gives the within-variance, reflecting the variability that would have appeared, even if the data were complete. At this, $Var_{\text{within}}{}^{(t)} = \frac{1}{n} Var(X^{(t)})$ gives the variance of the data within the data set t. Apart from that, $\widehat{Var}_{\text{between}} = \frac{1}{T-1}\sum_{t=1}^{T}(\hat{\theta}^{(t)} - \hat{\theta})^2$ is a measurement of the variance, emerging from the repeated imputation procedure and consequently ascribing to the missing data. Hence, the formula for the variance can be derived and is given by $\widehat{Var}_{\text{within}} + (1 + \frac{1}{T})\,\widehat{Var}_{\text{between}}$. The factor $(1 + \frac{1}{T})$ reflects the fact that only a finite number of estimates provide the basis for estimating the final point estimate by averaging them together.

Thus, the advantage of multiple imputation becomes clear. Due to the repeated imputation, the uncertainty that comes along with the missing values is reflected, whereas single imputation methods treat the imputed values like original values (cp. Dempster and Rubin (1983)). As a consequence standard errors are underestimated and p-values too small. In contrast to that, multiple imputation results only in guesses for the real value which results in additional uncertainty. Since the goal of imputation is not to replace the missing values, but rather to obtain valid inferences, which are within the realm of statistical plausibility of inferences that would have been obtained, had there been no missing data, multiple imputation in general should be preferred in theory. Those theoretical findings shall later in section 3.5 be evaluated.

Yet, imputation methods cannot only be differentiated with respect to the number of imputed values, but also with respect to the way, those values are derived from the predictive distribution. There can be either random draws or the choice can be based on statistical models or algorithms. Thus, orientating on Little and Rubin (1987) the following methods are made out, which can be subdivided even further, as can be seen in Weisberg (2005):

- random imputation
- mean imputation
- regression imputation
- hot or cold deck imputation

Let again X_j be the variable of interest, having some missing values, so that X_j can be split into X_j^{obs} and X_j^{mis}. For purpose of easier notation, let the first m observation be missing, whereas the observations x_{ij} with $i = (m+1), \cdots, n$ are reported. In the case of random imputation, each missing value of X_j is replaced by a random value of the remaining, observable $n-m$ values x_{ij}. In doing so, it can be differentiated between drawing these values with or without replacement. This approach does not impute values that are completely out of range, but it does not use much information from the observed data either.
Hence, mean imputation replaces all missing values of the variable X_j by the unconditional arithmetic mean of the observed values, thus by $\bar{X}_j^{\text{obs}} = \frac{1}{n-m}\sum_{i=m+1}^{n} x_{ij}$. Consequently, the estimated mean of the imputed variable X_j does not differ from the average value of the observed part and the variance of the imputed variable X_j is underestimated by $\frac{n-m-1}{n-1}$ (see appendix A). Using also the information gathered through the other variables, this approach can be extended by using the mean, conditional on the other variable values, recorded in the incomplete cases. Since this approach, called conditional mean imputation, can then be carried forward to the procedure of the regression imputation, it is not focused on here, but referred to Little and Rubin (1987) or Spieß (2008) for details.
When performing regression imputation, an appropriate regression model is fit, where the interested variable X_j with missing values is the dependent variable and the other collected variables without missings, here denoted as $X_1, \cdots, X_k$, take on the role as regressors. Based on the $n-m$ fully observed units of the data set, the regression coefficients β are derived which serve to predict the missing values $x_{1j}, \cdots, x_{mj}$. Then, all missing cases of X_j are replaced by the predicted values derived from the regression equation $\beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k$. This approach theoretically gives good point estimates for the missing values, since information of the individual is used for prediction. However, variances are too low which can be ascribed to the fact that the imputed values always fall right on the regression line, whereas in fact there are always differences between observed values and the predictions, resembled through the regression line. This problem can be solved, if random

normal errors $\epsilon \sim N(0, \sigma_\epsilon^2)$ are added to the predicted values before imputing them. This is what is understood by stochastic regression imputation.
The last category of single imputation approaches is based on algorithms, searching for equal observations with respect to the remaining observable variables $X_1, \cdots, X_k$. Having found one or more observations that are similar, these observations serve as donors and their observed value of the variable of interest is imputed. Again, there are differentiations. First of all, these donors can be searched in the same data set which is called hot deck imputation or in a different data sets which is equivalently called cold deck imputation. It is either possible to search only one donor or several donors whose values are then combined and imputed. Moreover, there are different methods to find those hot or rather cold deck values. One possible approach is to define adjustment cells especially based on different categorical variables of $X_1, \cdots, X_k$ in which observations are close to each other with respect to their variable values. Then, for a missing value of X_j one ore more donors of the corresponding adjustment cell are chosen. On the other hand it is also possible to search the nearest neighbor, which is the fully observed unit that is most similar to the observation whose variable value shall be imputed. In doing so, similarity has to be defined, which means that a metric for measuring the distance between the units has to be determined. There are different options like the maximum deviation, the mahalanobis distance or the euclidean distance.
In the application of this thesis it is focused on the predictive mean matching as a certain form of nearest-neighbor hot-deck imputation. There, the distance is specified through a prediction model for the missing values and for possible donor values. That means a value is filled in for the missing value that is chosen randomly from the (here) 5 observed donor values whose regression-predicted values are closest to the regression-predicted value for the missing value from the regression model. This way, it is ensured that the imputed values are plausible, since predictive mean matching is restricted to observed values. The performances of the different imputation procedures will later in section 3.5 be evaluated.

3.2.4 Maximum Likelihood Procedures

Another approach for handling the missing data problem treats population values as realizations of random variables (cp. Little (1982)). Thus, a distribution assumption, in general multivariate normal distribution, is made for the data. Due to this model distribution, likelihood-based inference (cp. Royall (1970)) can fol-

low. In order to understand the different approaches, classified to the category of likelihood-based procedures, some more theory concerning this field is needed.
Let X be the data, having several missing values and θ the unknown parameter, which governs the distribution of X, abbreviated by $f(X|\theta)$. Thus, interests are mainly focused on θ which determines the assumed model for the given data. Let furthermore M be the previously introduced missing data indicator. M is following a bernoulli distribution, with parameter ξ, so $f(M|X,\xi)$. We are then interested in the joint distribution of X and M which can be specified, according to the definition of conditional densities (cp. Fahrmeir et al. (2007)), as the product of the density of X and the conditional distribution of M given X. That is

$$f(X,M|\theta,\xi) = f(X|\theta) \cdot f(M|X,\xi).$$

However, this relation is only valid, if the joint distribution does not depend on the observation i and furthermore the parameters θ and ξ are distinct (cp. Rubin (1976)). With this joint distribution, a statistical model is specified and the likelihood function can be derived.

In general the likelihood $L(\theta|X)$ is defined as a function of the parameter θ for fixed outcome X, so any function of θ that is proportional to $f(X|\theta)$ is defined as likelihood (cp. Fahrmeir et al. (2007), Schafer (1997)):

$$L(\theta|X) \propto f(X|\theta)$$

Then, a maximum likelihood estimate $\hat{\theta}$ is defined as a value of the unknown parameter θ that maximizes the likelihood $L(\theta|X)$ or rather the log-likelihood $\ell(\theta|X)$ (Spieß (2008)).

According to these definitions the complete-data likelihood corresponding to the joint distribution $f(X,M \mid \theta,\xi)$ is given by

$$L(\theta,\xi \mid X,M) \propto f(X,M \mid \theta,\xi)).$$

Splitting X into the observable part X^{obs} and the missing part X^{mis}, then gives $L(\theta,\xi \mid X,M) \propto f(X^{\text{obs}},X^{\text{mis}},M \mid \theta,\xi)$. In the Maximum Likelihood approach missing values are viewed as random variables that have to be removed from the likelihood function, for example by integrating them out. Integrating out the missing data from the joint distribution results in the joint distribution of observed data and the missing indicator. The observed-data likelihood of θ and ξ then is any function of those two parameters that is proportional to $f(X^{\text{obs}}, M|\theta, \xi)$:

$$
\begin{aligned}
f(X^{\text{obs}}, M \mid \theta, \xi) &= \int f(X^{\text{obs}}, X^{\text{mis}}, M|\theta, \xi) dX^{\text{mis}} \\
L(\theta, \xi \mid X^{\text{obs}}, M) &\propto f(X^{\text{obs}}, M \mid \theta, \xi) \\
&\propto \int f(X^{\text{obs}}, X^{\text{mis}} \mid \theta) \cdot f(M \mid X^{\text{obs}}, X^{\text{mis}}, \xi) dX^{\text{mis}} \qquad (7)
\end{aligned}
$$

This likelihood of equation (7) consists of two separate factors which are determined by different parameters: The distribution of X and the distribution for the underlying missing data mechanism, which has already been formalized before in equations (1), (2), (3) by Rubin (1976). According to those formulas $f(M|X^{\text{obs}}, X^{\text{mis}}, \xi)$ can be rewritten due to which of the missing mechanisms is present. In the case of MCAR and MAR, this factor is independent of X^{mis} and consequently the second term of equation (7) can be drawn out of the integral:

- MCAR:

$$
\begin{aligned}
L(\theta, \xi|X^{\text{obs}}, M) = &\int f(X^{\text{obs}}, X^{\text{miss}}|\theta) \cdot f(M|\xi) dX^{\text{mis}} \\
=&f(M|\xi) \cdot \int f(X^{\text{obs}}, X^{\text{mis}}|\theta) dX^{\text{mis}} \\
=&f(M|\xi) \cdot L(\theta|X^{\text{obs}}) \qquad (8)
\end{aligned}
$$

- MAR:

$$
\begin{aligned}
L(\theta, \xi|X^{\text{obs}}, M) = &\int f(X^{\text{obs}}, X^{\text{mis}}|\theta) \cdot f(M|X^{\text{obs}}, \xi) dX^{\text{mis}} \\
=&f(M|X^{\text{obs}}, \xi) \cdot \int f(X^{\text{obs}}, X^{\text{mis}}|\theta) dX^{\text{mis}} \\
=&f(M|X^{\text{obs}}, \xi) \cdot L(\theta|X^{\text{obs}}) \qquad (9)
\end{aligned}
$$

Thus, when the aim is to receive maximum likelihood estimates for the parameters θ, it is sufficient in the case of MCAR and MAR to maximize the simpler likelihood $L(\theta|X^{\text{obs}})$ which ignores the nonresponse mechanism, instead of $L(\theta, \xi|X^{\text{obs}}, M)$, because here, likelihood-based inferences for θ will be the same due to proportionality of the likelihoods (cp. Little and Rubin (1987)). Therefore, no missing data model has to be set up, but maximum likelihood estimates are only based on the specified data generation model that is assumed for the variables with missing values.

Based on this background, the first maximum-likelihood approach is a direct method in the sense that model parameters and standard errors are estimated directly from the available data in a Structural Equation Model (SEM). For closer information to SEMs see for example Westland (2015), which provides a history of this models or Fox (2006), which familiarizes with the application of SEMs in the software R. To raise the subject shortly, the advantages of SEMs encompass the consideration of latent variables and the possibility of modeling more complex associations compared to a regression model. Therefore, SEMs consist of two model equations, the measurement submodel, defining latent variables with the help of one or more observed variables and the structural submodel, measuring the relationship between the different elements including latent constructs and observable variables. Then the first step is to formulate the SEM, generally based on prior information, literature or careful consideration, only then this model is tested empirically. This order of steps already reveals that SEMS are often used to determine, whether a certain model is valid, rather than finding a suitable model. In the application of the thesis, based on simulation, the underlying model is already known, consequently, the application of a SEM on the data including missing values is standing to reason. Here, we are not dealing with latent constructs, which is why the measurement submodel is not specified, but disregarded. The structural submodel equals the regression model of interest, already introduced in equation (4). Regression models are one of the major applications of structural equation modeling or viewed differently, the linear regression model is just a special case of the SEM.

If the data in a SEM has missing values, which in our underlying situation is the major subject, parameter estimation can be handled with direct Maximum Likelihood or also called Full Information Maximum Likelihood (FIML) which is a Maximum Likelihood approach on all available data. Originally, it was outlined by Finkbeiner (1979) for use with factor analysis. Instead of deleting observations with missing values and just calculating maximum-likelihood estimates based on the fully observed units, which would lead to a strong reduction of the estimation's efficiency, the FIML method uses all available information in all observations by integrating the likelihood function over the variables with missing data. Assuming multivariate normality, the overall likelihood function value is obtained by summing the n casewise likelihood functions of the observed data (cp. Enders (2001b)), which leads to a more complicated function. Thus, for each observation a separate likelihood function is set up and then maximized together to find the estimates of interest.

Let here in general X be the entire data set of $i = 1, \cdots, n$ observations, containing missing values in different variables. Then the likelihood of a certain row i of the data is given (cp. Arbuckle et al. (1996)) by

$$L_i = k_i \ln(2\pi) + \ln(|\Sigma_i|) + (x_i - \mu_i)\Sigma_i^{-1}(x_i - \mu_i)^T \tag{10}$$

which is equal to (-2) times the logarithm of the probability density function of the multivariate normal distribution (see appendix A), which is the probability of the data given the model. Here, k_i denotes the number of non-missing observed variables of observation i, x_i is the filtered row i of the data set, μ_i the filtered model-implied mean row vector and Σ_i the filtered model-implied covariance matrix. Filtering in the context of FIML means the removal of the appropriate missing entries. If for example, the value of the third variable is missing in the observed row, the corresponding μ vector consists of the mean of all the variables except that one of the third variable. Adequately, the covariance matrix diminishes by leaving out the third row and third column of the model-implied covariance. So the determinant of Σ_i is calculated based on all variables that are observable for unit i. Consequently, these separate likelihoods L_i measure the discrepancy between the observed data of unit i and the parameter estimates. This way, for each observation i all available information are used and then combined in the likelihood of the entire data which is obtained by summing over the casewise likelihoods L_i:

$$L = \sum_{i=1}^{n} L_i \tag{11}$$

By maximizing this likelihood, those values are chosen as estimates for the population parameters of interest that are most likely to have resulted in the observed data. Finally, it is important to note that for the process of FIML estimation two assumptions are made: First, in order to use maximum likelihood, distributional assumptions are required for all variables with missing values. Normality is not far to seek, but deviations from this assumption can have noticeable impacts on parameter estimation (cp. Enders (2001a)). In the application of the FIML estimation on the simulated data set to come, the only variable having missing values is the income variable, which has been constructed by a linear regression model (see equation (4)). Consequently, the logarithmized income here is normally distributed conditional on the other covariates. The second assumption is MAR, as we have seen in the introduction, which again is a strong assumption and limits the range of application of the FIML estimation.

Another approach to solve the likelihood equations in (8) or rather (9), to receive parameter estimates, are iterative methods. They are applied, since the log-likelihood equation tends to be complex for incomplete data, amongst others having no obvious maximum, and consequently calculating maximum likelihood estimates can be a major task (cp. Little (1982)). The most common iterative procedure for finding maximum likelihood estimates for parametric models, when data are not fully observed, is the Expectation-Maximization (EM) algorithm (cp. Dempster et al. (1977)) which is named after the two steps it combines by reiteration: The Expectation-step and Maximization-step.

The basic idea of the EM algorithm is that with the help of start values for the (set of) parameter θ it yields plausible values for the missing data which are the basis for re-estimating θ so that a sequence of estimations $\theta^{(t)}, t = 1, \cdots, T$ are received, converging against the maximum $\hat{\theta}$, which is the maximum-likelihood estimate. Thus, the EM algorithm benefits by the interdependence between missing data and parameters θ: The missing data contains relevant information for the estimation of θ, while θ in turn helps to find likely values of the missing data (cp. Schafer (1997)).

More detailed, in the starting situation we are observing data with missing values and given start values for the parameter θ. Here, the conditional expectation of the missing data is calculated, given the observed data and the current parameter estimates. These expectations then replace the missing values, so that the parameters θ can be re-estimated. This approach is finally iterated until the estimated parameter converge. Little and Rubin (1987) provide the underlying theory by presenting a formalization of the algorithm which is roughly presented in the following. Let the starting point again be the simpler Likelihood $L(\theta \mid X^{\text{obs}}) = \int f(X^{\text{obs}}, X^{\text{mis}} \mid \theta) dX^{\text{mis}}$ in equations (8) or rather (9) which should be maximized with respect to θ. For this purpose, focus is set on the complete data density $f(X^{\text{obs}}, X^{\text{mis}} \mid \theta)$ which again can be factored into the density of the observed data and the density of the missing data conditional on the observed data. The adequate decomposition of the log-likelihood then is given by $\ell(\theta|X) = \ell(\theta|X^{\text{obs}}) + \ln f(X^{\text{mis}}|X^{\text{obs}}, \theta)$, whereas $\ell(\theta|X)$ is the complete-data-log-likelihood, $\ell(\theta|X^{\text{obs}})$ is the observed or rather incomplete-data-log-likelihood and finally $\ln f(X^{\text{mis}}|X^{\text{obs}}, \theta)$ is the missing part of the complete-data-log-likelihood. After reposition, the equation is given by

$$\ell(\theta|X^{\text{obs}}) = \ell(\theta|X) - \ln f(X^{\text{mis}}|X^{\text{obs}}, \theta).$$

Let now $\theta^{(t)}$ be the current estimate of parameter θ. After some steps of calculation (see appendix A) this results in

$$\begin{aligned}\ell(\theta|X^{\text{obs}}) = &\underbrace{\int \ell(\theta|X^{\text{obs}}, X^{\text{mis}}) f(X^{\text{mis}}|X^{\text{obs}}, \theta^{(t)}) dX^{\text{mis}}}_{Q(\theta,\theta^{(t)})} \\ &- \underbrace{\int \ln f(X^{\text{mis}}|X^{\text{obs}}, \theta) f(X^{\text{mis}}|X^{\text{obs}}, \theta^{(t)}) dX^{\text{mis}}}_{H(\theta,\theta^{(t)})}.\end{aligned}$$

The algorithm then builds upon part $Q(\theta, \theta^{(t)})$. Under the assumption that the actual parameter value $\theta^{(t)}$ is the true parameter value θ, the E-step finds the expected value of this log-likelihood given the observed data:

$$Q(\theta, \theta^{(t)}) = E(\ell(\theta|X)|X^{\text{obs}}, \theta^{(t)})$$

Then the M-step finds $\theta^{(t+1)}$, such that

$$Q(\theta^{(t+1)}, \theta^{(t)}) \geq Q(\theta, \theta^{(t)}), \quad \forall\, \theta.$$

Thus, the M-step maximizes the expected value with regard to θ, to receive the next estimate $\theta^{(t+1)}$. Consequently, the log-likelihood increases in each iteration which is equivalent to the statement that the plausibility of the current estimate increases given the observed data (cp Little and Rubin (1987)). Obviously, the EM algorithm estimates the parameters of the normal distribution, namely mean, variance and covariances. Those estimates can be delivered to a regression procedure in order to receive consistent regression coefficients. However, the corresponding standard errors are not produced as a by-product, since the derivatives of the log likelihood function are not computed (cp. Dong and Peng (2013)). Consequently, computationally complex extensions are necessary, to receive corresponding standard errors (cp. Baker (1992)) which then gives that the EM algorithm results in quasi the same point estimates as FIML, but due to the complicated standard error calculation, FIML should be preferred, when focus is set on regression coefficient estimation. Consequently, the analysis of regression coefficient estimation with the EM algorithm will be omitted later in the performance section 3.5, also because of not available implementation in **R**.

Since the presented procedures FIML and EM ignore the dropout mechanism, they only provide good estimates for the ignorable mechanisms MCAR and MAR. In the case of MNAR, however, the likelihood of equation (7) does not simplify like in the cases (8) and (9), since the missing data mechanism here depends on X^{mis} and therefore this term cannot be drawn out of the integral, as can be seen in the following:

- MNAR:

$$L(\theta, \xi | X^{\text{obs}}, M) = \int f(X^{\text{obs}}, X^{\text{mis}} | \theta) \cdot f(M | X^{\text{obs}}, X^{\text{mis}}, \xi) dX^{\text{mis}}$$

Consequently, this scenario requires special attention, since here, the dropout process, represented by M, has to be modeled explicitly and inference has to be done regarding θ and ξ. However, finding a model that correctly represents the underlying response mechanism is difficult, so these approaches are highly sensitive for misspecification. Wrong models then can produce strongly biased results (cp. Demirtas and Schafer (2003)). Yet, when finding a model that tends to represent the nonresponse mechanism in a good way, unbiased estimates even for the scenario of nonignorable nonresponse mechanism can be received.
Summarizing, Maximum Likelihood estimation is possible even in the scenario of MNAR, but in general it is complicated, since the joint distribution of data and the missing mechanism has to be considered. Thus, the joint distribution is factorized in an appropriate way, which leads to two model-based approaches: The Selection Models (cp. Heckman (1976)) and Pattern Mixture Models (cp. Little (1993)). Both are based on the basic approach of formulating a statistical model for the joint distribution of data and the missing process and estimating then the corresponding model parameters.

Starting with the Selection Model (SM), the joint distribution of X and M is factorized into the marginal distribution of X and the conditional distribution of M given X:

$$f(X, M | \theta, \xi) = f(X | \theta) \cdot f(M | X, \xi)$$

Behind this factorization is the idea of Heckman's two-step-statistical approach for correcting for non-randomly selected samples (cp. Toutenburg et al. (2004)). The basic idea of this model is the presence of two latent variables: The first one is the dependent variable of interest itself, in later application the income variable, but

here in general it is denoted with Y^*. This variable is only observable for a certain part of the sample, depending on a selection process, so instead of Y^* we only have information about Y. On the other hand there is a non-observable variable M^* which governs the selection process dependent on different explanatory variables (cp. Windzio (2013)). Instead of this unobservable variable, only information about the dropout of an individual are present which are gathered in the previously introduced missing indicator M, taking value 1, if $M^* \leq 0$ and value 0 otherwise. The true latent model of interest is specified by $\mathbb{E}(Y^*|X)$, but fitting it only to the participating part of the sample might lead to biased estimates, since the missing part of the sample might differ with respect to the dependent variable which then leads to wrong inferences. The Heckman selection model identifies this sample-selection bias and corrects it.
Considering that, two models for the two latent variables, namely the so called selection model and the outcome model, have to be fit and combined in an adequate way. In order to model the dependence structure of those models, a correlation is assumed between the error terms of the two model equations, leading to the assumption of a bivariate normal model (cp. Toutenburg et al. (2004)). Then, the selection bias, if present, will be traced back exclusively to this correlation. So if there is no correlation between the error terms, the effects of the outcome model, fitted for the participating subsample, are valid for the whole population, otherwise a correction is necessary.
Let

$$y_i^* = x_i^T \beta + \epsilon_i \tag{12}$$

be the outcome model of the Heckman selection model, with the dependent variable Y^*, covariates X and error term ϵ.

$$\mathbb{P}(m_i = 0|z_i) = \Phi(z_i^T b) \quad \text{or rather} \quad m_i^* = z_i^T b + e_i \tag{13}$$

represents the selection model, a probit model, where the explanatory variables Z determine the probability of being selected and where the dependent variable gives information about, whether Y^* is observable for a certain individual i. In principle, the two equations (12), (13) could have the same set of regressors, but in order to avoid collinearity and to get a good estimate of the selection model, it is desirable to have at least one variable in the selection equation that is expected to affect the selection process, but not directly the variable of interest, except through selection. This is why the variable “willingness” has been included in the simulated data set.

Estimating the selection model (13) for all units gives a prediction of the participation-probability for each individual (cp. Guo and Fraser (2014)). If ϵ and e are not correlated and furthermore ϵ and X are not correlated, then the regression model (12) does not give biased estimates. However, since Y^* is only partly observable, those assumptions are in general not fulfilled. Then the expected value of y_i^* does not only depend on x_i, but also on the selection process which decides over $m_i = 0$ or $m_i = 1$ and consequently over the observability of y_i^* (cp. Windzio (2013)). Thus, instead of the latent model of interest $\mathbb{E}(Y^*|X)$, a possible model based on the available information has to be specified correctly and estimated afterwards. The true, latent model only for the participated part is given by

$$\begin{aligned} \mathbb{E}(Y|X, M = 0) = \mathbb{E}(Y^*|X, M^* > 0) = \mathbb{E}(X^T\beta + \epsilon|X, e > -Z^Tb) = \\ = X^T\beta + \mathbb{E}(\epsilon|e > -Z^Tb) \end{aligned} \tag{14}$$

Equation (14) then shows that the error of the model of interest is not random, but depends on the selection process which itself is conditioned by Z^Tb. Fitting the model without considering the underlying situation, would result in biased estimates of the regression model, as indicated by the additional term $\mathbb{E}(\epsilon|e > -Z^Tb)$ that can be transformed, according to truncated normal distribution (cp. derivation in appendix A), to

$$\mathbb{E}(\epsilon|e > -Z^Tb) = \sigma_s \underbrace{\frac{\phi(Z^Tb)}{\Phi(Z^Tb)}}_{\lambda} . \tag{15}$$

σ_s is the selection bias and the following ratio is called inverse Mills Ratio and can be abbreviated as λ. The selection bias takes value 0, if there is no correction, otherwise it can be calculated by $\sigma_\epsilon \cdot \rho_{\epsilon,e}$, whereas σ_ϵ is the standard deviation of ϵ and $\rho_{\epsilon,e}$ the correlation coefficient between the two error terms ϵ and e.

The inverse Mills Ratio λ is the ratio of the probability density function over the inverse cumulative distribution function of the standard normal distribution (cp. Toutenburg et al. (2004)) and equals a transformation of the predicted individual probabilities from the selection model of (13). Thus, the inverse Mills Ratio measures the individual observability of the latent dependent variable Y^* and can take values between $[0, \infty[$. Small values serve as indicator for a small probability of participation and high values represent a high probability of participation. By including λ as an additional explanatory covariate in the outcome model (12), which is fitted only on the participating part of the sample, the sample-selection bias is corrected. Solving the regression equation is then straightforward and can be done by standard least squares method (cp. Guo and Fraser (2014)). By including the inverse Mills Ratio as additional covariate, the Heckman Selection Model uses information from participants who dropped out and therefore did not respond. Thus, a great advantage of the Heckman Selection Model contains that the same model that would have been chosen for the naive approach, can be modified by a model for the missing mechanism and then used the same way. Finding an appropriate model for the selection process, however, is definitely the pitfall of this approach. Also the assumption of an underlying normal distribution restricts the application of this approach to special cases.

In contrast to the Heckman SM, it is spoken of a Pattern Mixture Model (PMM), if the joint distribution of data X and missing process M is specified through a model for the marginal distribution of M and the conditional distribution of X given M:

$$f(X, M|\theta, \xi) = f(M|\xi) \cdot f(X|M, \theta)$$

It is important to note that in general the parameters θ and ξ of the PMM are not identical with the parameters of the SM. Both models may only be equivalent for the case of MCAR (cp. Little (1994)). Nevertheless, the same parameter notation is chosen here for reasons of clarity.
This PMM approach quasi represents the analyst's or the survey sampler's point of view, having to deal with the given missing data, represented by the marginal distribution $f(M|\xi)$, but actually being interested in the underlying true values, represented by the conditional distribution $f(X|M, \theta)$. Thus, on the one hand, PMMs specify the conditional distribution of the data, given that the variable of interest is observed or missing respectively and on the other hand the marginal distribution of the binary missing indicator. Consequently, two pattern are taken as a starting point, the pattern of responder ($M = 0$) and nonresponder ($M = 1$). This leads to $f(X|\theta)$ being a mixture of two distributions, which then results in two models and the name 'pattern mixture'. Only for the case of MCAR those are equivalent, otherwise it is dealt with two separate models and consequently two different distributions with different parameters, namely $f(X|M = 0, \theta^0)$ and $f(X|M = 1, \theta^1)$. Nevertheless, substantive interest concerns the distribution and the corresponding parameters averaged over patterns. Hence, maximum-likelihood estimates are received as a mixture of the estimates resulting from both models. Concentrating for example on estimating the mean of the variable of interest with

missing values, denoted as before with Y, it is given by

$$\hat{\mu_Y} = \mathbb{P}(M = 0) \cdot \hat{\mu}_Y^0 + \mathbb{P}(M = 1) \cdot \hat{\mu}_Y^1, \tag{16}$$

whereas $\hat{\mu}_Y^0$ or rather $\hat{\mu}_Y^1$ are the averages of Y for the responders and nonresponders, respectively (cp. Little (1994)). At this point it conveys that some of the parameters from the PMM are not identified from the data. In the case of the marginal mean estimation, the likelihood does not give an estimate for the component $\hat{\mu}_Y^1$, resembling the mean of the incomplete cases. PMMs are, in general and especially in the case of univariate nonresponse, which we are focused on here, underidentified (cp. Little (1993)), since in the stratum of nonresponders there is no knowledge about the distribution of the variable of interest conditional on the other covariates. In order to then being able to identify those parameters, additional information have to be given or restrictions regarding the missing mechanism M have to be imposed. Those restrictions, explained in Little (1994) are based on two continuous variables, with one having missing values and both following normal distributions. The dropout then is assumed to be either completely at random or depending on one or both of those variables. This dependence structure may obviously be in many different ways and is not known in general therefore it is assumed to be an arbitrary unspecified function of a linear combination of those continuous variables. The simulated data basis of this thesis, however, provides a dropout process that depends only on categorical variables in the case of MAR and additionally on the income variable itself in the MNAR scenario. The only considered continuous variable besides income is age. Assuming a dependence structure between this variable and the missing mechanism gives biased estimates, since it does not agree with the underlying data situation.

Therefore, the application of the PMM later primarily concentrates on assumed priori information, instead of those restrictions. This means, the inestimable parameters are set equal to the parameters, or rather to a function of the parameters, resulting from the distribution of the responders, whereas this connection is based on priori information. Obviously, those different restrictions are often viewed as the drawback of the PMMs, but the need of additional assumptions also forces the analysts to think about their justification and usefulness. Thus, this can form the starting point of sensitivity analysis (cp. Thijs et al. (2002)) or the approach of partial identification (cp. Manski (2003)).

3.3 Partial Identification

The preceding chapters of this thesis have shown that development regarding handling missing data has been fostered and by now there are many different approaches providing good precise point estimates. However, many of those methods are based on strong assumptions especially with regard to the underlying nonresponse mechanism. Without those assumptions many of the models would be non-identifiable, as we have seen in the previous section about PMMs. Non-identification means that different values of the parameter θ, the model should estimate, generate the same probability distribution $\mathbb{P}$ of the observable variables, so that even in the case of an infinite number of observations from the model, the true parameter value is not known. Lehmann and Casella (2006) provide the corresponding definition: If Y is distributed according to $\mathbb{P}(Y \mid \theta)$, then θ is said to be unidentifiable on the basis of Y, if there exists $\theta_1 \neq \theta_2$ for which $\mathbb{P}(Y \mid \theta_1) = \mathbb{P}(Y \mid \theta_2)$. The usual approach in such a scenario, when the number of known parameters in a model is smaller than the number of parameters that have to be estimated, is to restrict the model by setting parameters equal or constant, in order to receive identifiable statistical models (cp. Casella and Berger (1990)). Often, these assumptions are untenable or cannot be checked, but made nevertheless, in order to be able to apply a certain method. Thus, in many cases assumptions are made due to convenience and not due to plausibility. Stronger assumptions in general lead to more powerful inferences, but less credibility and reliability (cp. Manski (2003)). Obviously, the results have to be considered as critical and consequently a new approach comes to the fore.
Manski (2016) described the estimation of parameters without making any assumptions about nonresponse. This way, all values that the missing data might take, are regarded which results in an upper and a lower bound for the population parameter of interest. The embedded set of possible values is then called identification region or identified set, abbreviated in the following as H. Moreover, it is either spoken of a point-identified parameter, if the identification region consists only of one point, or otherwise of partially identified parameters. Consequently, in the field of partial identification, it is not only distinguished between parameters that are identifiable and those that are non-identifiable, but it is also possible to identify the parameters of interest in parts and thus to receive interval estimates (cp. Manski (2003)). Even though uncertainty remains with this approach, results can contain important information.

Let again θ be the statistic of interest, which can be viewed as a parameter of the outcome probability distribution $\mathbb{P}(Y)$ or rather $\mathbb{P}(Y|X)$, whereas for reasons of clarity it is concentrated on the first notation. With the previously introduced definition of the missing data indicator variable M and the application of the Law of Total Probability (cp. Fahrmeir et al. (2007)), the outcome distribution of interest, which serves as the basis for deducing θ, can be displayed in

$$\mathbb{P}(Y) = \mathbb{P}(Y|M = 0) \cdot \mathbb{P}(M = 0) + \mathbb{P}(Y|M = 1) \cdot \mathbb{P}(M = 1). \tag{17}$$

In general only a subsample of the population is observed which brings along that empirical distributions are used to estimate their population counterparts. The Strong Law of Large Numbers (cp. Fahrmeir et al. (2007)) then implies that the resulting estimates are consistent and converge almost surely to the real set. Since in the context of the second part of the thesis we are dealing with a simulated data set that serves as population gold standard, we are not viewing samples and consequently we continue with the notation of (17). Here, the distribution of the observed outcome, as well as the distribution of the missing process, are known, but not the distribution of missing outcomes $\mathbb{P}(Y|M = 1) = \gamma$, since no information about Y is given for the missing data. Therefore, γ is not identified, but all this knowledge gives rise to conclusions about the identification region of the outcome distribution and consequently of the identification region of the related parameter. Thus, $\mathbb{P}(Y)$ lies in the identification region

$$H[\mathbb{P}(Y)] = [\mathbb{P}(Y|M = 0) \cdot \mathbb{P}(M = 0) + \gamma \cdot \mathbb{P}(M = 1)], \quad \gamma \in \Gamma_Y, \tag{18}$$

whereas Γ_Y denotes the set of all probability distributions on Y and γ is limited to $[0, 1]$. Hence, the identification region of θ comprises all the values that can arise, when $\mathbb{P}(Y)$ ranges over all of its possible values η encompassed in (18):

$$H[\theta] = \{\theta(\eta), \eta \in H[\mathbb{P}(Y)]\}$$

Manski also refers this general definition to some specific parameters like the mean of a function g of Y, having the bounds $g_0 = \inf_Y g(Y)$ and $g_0 = \sup_Y g(Y)$. The Law of Iterated Expectations, saying in general that for a continuous variable Y and a binary variable Z, $\mathbb{E}(Y) = \mathbb{E}_Z(\mathbb{E}_Y(Y|Z))$ is valid, gives

$$\mathbb{E}[g(Y)] = \mathbb{E}[g(Y)|M = 0] \cdot \mathbb{P}(M = 0) + \mathbb{E}[g(Y)|M = 1] \cdot \mathbb{P}(M = 1),$$

which then, due to the knowledge of all elements except $\mathbb{E}[g(Y)|M = 1]$, which can take values in the interval $[g_0, g_1]$, results in the following identification region for $\mathbb{E}[g(Y)]$:

$$\begin{aligned} H[\mathbb{E}[g(Y)]] =& [\mathbb{E}[g(Y)|M = 0] \cdot \mathbb{P}(M = 0) + g_0 \cdot \mathbb{P}(M = 1); \\ & \mathbb{E}[g(Y)|M = 0] \cdot \mathbb{P}(M = 0) + g_1 \cdot \mathbb{P}(M = 1)] \end{aligned}$$

Obviously, the width of the resulting interval varies with the probability of missing data $\mathbb{P}(M = 1)$. It is informative, even if it has infinite length due to either $g_0 = -\infty$ or $g_1 = \infty$ (cp. Manski (2003)).

This basic step, presented in general and for the specific population parameter of the mean, only uses information offered by the data and is therefore referred to as empirical evidence. Obviously, the empirical evidence does not reveal anything about the distribution of missing data, which is why many possible $\mathbb{P}(Y)$ are imaginable. Therefore, partial identification may also consist of a second step. The aim is to constrain the set of possible distributions for the missing outcome $\Gamma_{Y,M=1} \in \Gamma_Y$ by including justified assumptions, which are based on the idea that missing data and observed data do not differ too much. Different assumptions may shrink the possible range of distributions from point estimation to estimates with less identification power. Thus, additional information about $\mathbb{P}(Y)$ can be gathered and corresponding identification regions that are narrower and a middle ground between precision, which can be reached by strong assumptions like MCAR or MAR (cp. Plass et al. (2015)) and credibility, can be received. In contrast to sensitivity analyses, which regard the collection of all precise results from successively relaxed assumptions, the starting point of partial identification is total uncertainty, where then assumptions are added gradually (cp. Plass et al. (2015)).
In this context Manski (2003) suggested a few possible assumptions for the distribution of interest $\mathbb{P}(Y)$, which should be presented and discussed in the following. The first possibility is to assume that the distribution of responder and nonresponder does not differ, which can be expressed through

$$\mathbb{P}(Y) = \mathbb{P}(Y|M = 0) = \mathbb{P}(Y|M = 1)$$

and reminds on the definition of MCAR. With this strong assumption, $\mathbb{P}(Y)$ is point identified, which can be seen by the fact that now each of the elements in equation (17) can be estimated. However, the assumption can neither be proved wrong, nor be checked. In general it is often not valid, so Manski proposed also weaker assumptions that rely on instrumental variables X. In Manski (2003) two attempts are explained. The first one contains distributional assumptions that use instrumental variables to identify the distribution of outcomes. The second one is based on the statistical independence between outcomes and instrumental variables. Both assumptions can be relaxed, so that their acceptance becomes more plausible. Nevertheless, at the same time they then lose identifying power. The first assumption based on instrumental variables is an equality of the distributions between responder and nonresponder for units having same values in those variables X, so

$$\mathbb{P}(Y|X) = \mathbb{P}(Y|X, M = 0) = \mathbb{P}(Y|X, M = 1).$$

This assumption is non-refundable, since the empirical evidence does not reveal anything about $\mathbb{P}(Y|X, M = 1)$. However, it is possible to point identify $\mathbb{P}(Y)$, namely by $\mathbb{P}(Y) = \sum_X \mathbb{P}(Y|X, M = 0) \cdot \mathbb{P}(X)$, because of the Law of Total Probability and the previously formulated assumption. Weakening of this assumption can be achieved by assuming that instead of the distribution, only the mean of observed and missing outcomes is the same, conditional on X:

$$\mathbb{E}[g(Y)|X] = \mathbb{E}[g(Y)|X, M = 0] = \mathbb{E}[g(Y)|X, M = 1]$$

Or finally mean missing monotonicity can be assumed,

$$\mathbb{E}[g(Y)|X, M = 0] \geq \mathbb{E}[g(Y)|X] \geq \mathbb{E}[g(Y)|X, M = 1],$$

which means that for each realization of X, the mean value of $g(Y)$, when Y is observed, is greater than or equal to the mean value of $g(Y)$, when Y is missing.
The other approach is based on assuming statistical independence, whereas statistical independence of outcome and instrumental variables is the strongest assumption, denoted with

$$\mathbb{P}(Y|X) = \mathbb{P}(Y).$$

In this case, the assumption is refutable and the identification power can range from a point identification of $\mathbb{P}(Y)$ to no identification power, if there is statistical independence between M and X. Also in this case, there are weaker versions of this assumption. On the one hand, statistical mean independence of outcomes and instruments can be assumed:

$$\mathbb{E}[g(Y)|X] = \mathbb{E}[g(Y)]$$

Or on the other hand mean monotonicity of outcomes and instruments:

$$\mathbb{E}[g(Y)|X_1] \geq \mathbb{E}[g(Y)|X_2], \forall X_1 \geq X_2$$

All those assumptions may lead to smaller identification regions compared to using empirical evidence alone. However, besides the gain of credibility, this shrinkage also obviously brings along a loss in precision and consequently less stronger conclusions. There is no universal guideline about which middle ground and consequently which assumption to choose, instead it has to be considered in each situation, where importance should be attached and consequently which assumptions are reasonable. This approach to deal with missing data obviously differs from the other methods presented before. Thus, this chapter of partial identification rather proposes a new perspective on the underlying problem, which is recommended to follow up in other scientific works. However, in this thesis, which will concentrate on the application of the presented methods in the chapters to come, the application of this partial identification approach would go beyond the constraints and therefore it is focused on the methods providing point estimates for the statistics of interest. Nevertheless, when applying the different nonresponse methods, the corresponding assumptions are emphasized and discussed critically, in order to show the usefulness of the partial identification approach. Moreover, in many cases it is possible, to change the assumptions of the approaches and to review then the performances which falls in the category of sensitivity analysis.

3.4 Simulated Data Basis

In order to evaluate the different introduced methods, whose objective is to yield good estimates for population parameters, again a suitable data basis is needed. For this purpose, it is approached in the same way, as it has been done in chapter 2.1, where 100 data sets with the different nonresponse mechanism MCAR, MAR and MNAR have been created, respectively, to show the impacts on parameter estimation. However, there are little differences to the data sets that are created now: Besides the comparison of the different correction methods' performances, the focus of the next sections is also to analyze whether some methods lead to good results for some missing data mechanisms, whereas they are not useful for others. Since these comparisons can only be drawn, if the number of dropouts in the different data sets are equal, importance is attached to this fact, now. Therefore, the individual probabilities for a dropout in the case of MAR are increased deliberately by adding 0.014 to the probabilities that have resulted from the model orientated on the Allbus data set before. This randomly appearing number guarantees that the number of dropouts in the 100 iterations of the MAR model is on average and according to its range very close to that of MNAR, which has a mean and median of 1091 dropouts. The number of dropouts in the MCAR case is then consequently also set to 1091. Even though those changes entail a stronger deviation from the underlying Allbus data set, it is necessary for this part of the thesis. Thus, the simulated data sets for this part of the thesis again consists of 100 data sets following MCAR, MAR and MNAR, respectively.

3.5 Performance of different Correction Methods

3.5.1 Ad-Hoc Methods

Evaluating the performance of the complete- or rather available-case analysis, indicates, what estimates can be expected, when the missing data is simply ignored. Analogous to the structure of the first part of the thesis, it is always first concentrated on the income mean estimation and afterwards on that of regression coefficients.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1491.440	1546.533	1529.321
$Bias(\hat{\mu}_{\text{income}})$	-1.51143	53.58223	36.36994
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00101	0.03589	0.02436
$Var(\hat{\mu}_{\text{income}})$	142.92347	201.75756	198.20299
$MSE(\hat{\mu}_{\text{income}})$	145.20789	3072.81341	1520.97589

Table 3.1: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of ad-hoc methods

Obviously only in the case of MCAR, unbiased estimates are received. The resulting income value of 1491.440, calculated as an average of all 100 MCAR iterations, is very close to the true value of 1492.951, which results in a very small bias and a comparatively small MSE. One can conclude that ad-hoc methods in the case of MCAR dropouts yield good mean estimates, even in a situation of over 30% missing values and consequently a large loss of information. In such a scenario it is justified

that analysts resort to simple ad-hoc methods. However, it is important to note that the MCAR situation is on rare occasions and in general cannot be checked. Therefore, it is critical to use these methods for mean estimation, since in the other two scenarios of MAR and MNAR, biased estimates are received. The average mean estimates then are 1546.533 and 1529.321, whereat the MAR dropout results in an even larger bias, a littler bit more variance and consequently a larger MSE.
In order to keep the performance subsection clear, the tables (C.14) and (C.15), containing the evaluation of the average regression coefficient estimation and the corresponding p-values, are located in the appendix, whereas the associated interpretation can be found here. Since the regression coefficients have very small values, it is often easier to have a look at the relative biases, instead of the biases, in order to evaluate the estimation in the case of the ad-hoc approach for all three scenarios MCAR, MAR and MNAR. Obviously most of the 21 regression coefficients showed the smallest absolute value of the relative bias, when the dropout followed MCAR. In these cases the differences compared to the absolute values of MAR or rather MNAR in general are apparent. Those findings show a tendency towards the fact that when ad hoc methods are applied, the estimation of regression coefficients are in average biased less, when data is MCAR. But since this is not the case anywhere near all regression coefficients, lets have next a look on the corresponding MSEs. Here no pattern can be noticed, given information about which dropout mechanism leads to the smallest or largest MSE of regression parameter estimation, when missing data is simply ignored. As a consequence that would mean that when estimating regression coefficients with a simple ad-hoc approach, the goodness does not vary noticeably for different missing data mechanisms. In all three scenarios the average estimates $\mathbb{E}(\hat{\beta})$ are close to the true values β, but due to the small values, it is difficult to assess the dimension of bias.
Therefore, next focus is set on the corresponding p-values, given information about how the significances of the different regression effects change, when ignoring the missing data following MCAR, MAR and MNAR, respectively. Table (C.15) shows that the p-values of the three scenarios are in the same range in the application of complete-case-analysis. It follows that significances do not differ between MCAR, MAR and MNAR. $\hat{\beta}_{\text{would not vote}}$ constitutes the only exception, since the corresponding p-value of MAR and MNAR comes below the threshold of 0.05, whereas in the case of MCAR this effect is only statistically significant at the level of 0.1. More important, however, is the result that has already been implied in section 2.3, where the effect of nonresponse error was analyzed: Except the highly significant effects like that of “age”, “female”, “half-time”, “part-time”, “not-employed” and “married

living together", the regression effects attenuate, when missing data is present and ignored. The other imaginable change that nonsignificant effects become significant, when data contains missing values, does not exist here. So finally it can be summarized that the dimension of bias for the estimation of regression coefficients is almost equal for MCAR, MAR and MNAR, when ad-hoc methods are applied, even though in the case of MCAR and MAR the dropout only depends on the predictors, but not on the variable of interest, which is the income. As a consequence, the decision of using ad-hoc methods for estimate regression coefficients, can be made quasi independent of the underlying nonresponse mechanism. Due to the weakening of the statistical effects, the application should, however, be viewed as critical.

3.5.2 Weighting

In order to be able to compare the performances of different correction methods, the weighting approach is applied to the same data sets as all the other methods, even though weighting in general is rather used for unit nonresponse. However, application for the item nonresponse problem is also feasible. As it has already been described in the theoretical subsection before, the choice of suitable weighting variables is decisive for a good performance of this approach. In order to decrease the estimation bias that resulted before in the ad-hoc approach, first the categorical variables gender, education and willingness are chosen as weighting variables, since due to our simulation background, it is already known that two of those variables, namely gender and education, have an influence on the income values and all three variables determine the MAR dropout.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1491.906	1492.272	1477.161
$Bias(\hat{\mu}_{\text{income}})$	-1.04534	-0.67873	-15.79006
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00070	-0.00046	-0.01058
$Var(\hat{\mu}_{\text{income}})$	108.82952	119.01859	136.46785
$MSE(\hat{\mu}_{\text{income}})$	109.92226	119.47926	385.79388

Table 3.2: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of weighting (weighting variables: gender, education, willingness)

The corresponding average mean estimation in table (3.2) shows that in the case of MCAR again unbiased estimates result. Compared to the ad-hoc estimation before, the value of 1491.906 is even closer to the true value and the variance decreased as well which then gives an improved MSE. Also in the MAR scenario unbiased estimates are received, which may be traced back to the fact that the choice of weighting variables agrees with the MAR dropout variables. Of course, this is not possible in a real data situation, because then the variables, which are responsible for the dropout, are not known, or at least not completely known. However, having this prior knowledge given, shows that then weighting offers a quite good possibility for correcting the biased estimates. Finally, in the MNAR scenario we may not expect unbiased estimates as in the MAR case, since here the dropout additionally depends on the income values itself, which cannot be compensated with the help of weighting variables. Thus, the absolute bias of 15.79006 obviously exceeds the values of MAR and MCAR. However, compared to the performance of the previously ad-hoc approach, even in the MNAR scenario bias and variance is decreased, which results in better mean estimations.

For comparing next the performance of regression parameter estimation between simple ad-hoc and weighting approaches, focus has to be set on tables (C.16) and (C.17) in the appendix, indicating the estimation results of the weighting approach, as well as the corresponding p-values. 10 of those 21 estimated regression coefficients in the MCAR scenario are closer to the true values compared to the corresponding estimates of the complete-case-analysis. This shows that weighting in this application has not resulted in a smaller bias, when MCAR is present. Comparing besides the biases also the MSEs for MCAR, gives an identical picture. In about half of the cases the ad-hoc approach yields smaller MSEs, in the other half it is the weighting approach. This result is standing to reasons, since dropout cannot be compensated by weighting here, because it is completely random. For both other scenarios the relative bias of regression coefficient estimates is smaller compared to the ad-hoc approach in two third of the cases. This can be seen as a first tendency towards a good correction of the weighting approach. Yet, the MSEs for all of those 21 coefficient estimates in the MAR and MNAR scenario are larger compared to the estimates of the complete-case-analysis, which results in the conclusion that this weighted regression did not improve the outcome.

Viewing at last the corresponding p-values in table (C.17) shows that they are very close to those of the ad-hoc method. Going into detail, in the majority of the statistically significant effects, p-values especially in the scenarios of MAR and MNAR, here are slightly larger than in the ad-hoc estimation, but those differences in general amount to differences at the second decimal place or even less, so consequently significances and interpretations do not change. Then, based on this single weighting approach, the conclusion has to be drawn that weighted regression does not

improve this kind of parameter estimation, when nonresponse is present. However, it has to be mentioned that this result might change, when other weighting variables are used.
All in all, we have seen that weighting can improve the parameter estimation and therefore is a first step away from simply ignoring the missing data. However, we have also seen that this is not always the case like in the estimation process of the regression coefficients. This brings up the question, whether the improved results of mean estimation are consistent or depend on the choice of the weighting variables. Until now, we have used weighting variables which knowingly had a relationship with the variable of interest and the dropout. Since this prior knowledge is in reality not given or rather limited, the performance of this weighting approach shall be considered also for different weighting variables. In order to remain in a certain extent, these analyses are limited to the parameter estimation of the mean income. It is differentiated between the case, where all categorical covariates of the regression model (4) are used as weighting variables and those cases, where only one variable is used as weighting variable, respectively. There, at first willingness is chosen as an example for a variable that influences the MAR and MNAR dropout, but not the values of income. The second choice is the variable education, having an influence on both components and finally professional activity, which amongst other variables predicts the income, but not the dropout. The results can be seen in the following tables.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1468.194	1469.151	1458.997
$Bias(\hat{\mu}_{\text{income}})$	-24.75687	-23.80027	-33.95418
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.01658	-0.01594	-0.02274
$Var(\hat{\mu}_{\text{income}})$	125.35169	76.84832	109.03060
$MSE(\hat{\mu}_{\text{income}})$	738.25415	643.30095	1261.91701

Table 3.3: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of weighting (weighting variables: gender, education, professional activity, family status, election intention)

First of all, it is interesting, how parameter estimation for MAR changes, when instead of all the MAR dropout variables, all independent covariates, playing a role for the creation of the income values, are now chosen as weighting variables. As expected, the average mean becomes worse, resulting in a bias of -23.80027. Since those weighting variables are all predictable for the income, the variance decreases, however, the corresponding MSE is far larger as in the weighting approach before.

Also in the case of MNAR the mean estimation becomes worse. The resulting bias of -33.95418 and the MSE of 1261.91701 are not far from the results of the ad-hoc approach. Viewing finally the results of MCAR shows that also in this scenario far worse estimates compared to the ad-hoc and the previous weighting approach result. Consequently, this choice for weighting variables is not suggestive. In order to find out, whether these bad results can be traced back to the chosen combination or rather the amount of weighting variables, we are now viewing mean estimation with only one weighting variable, respectively.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1491.457	1547.458	1530.003
$Bias(\hat{\mu}_{\text{income}})$	-1.49439	54.50708	37.05162
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00100	0.03651	0.02482
$Var(\hat{\mu}_{\text{income}})$	143.03381	214.44520	208.37248
$MSE(\hat{\mu}_{\text{income}})$	145.26701	3185.46716	1581.19489

Table 3.4: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of weighting (weighting variable: willingness)

Willingness is a categorical variable that is simulated without having a deliberate relationship to the income values. Thus, it is not surprising that parameter estimation in the MCAR case is not influenced, when choosing it as a weighting variable. However, amongst other variables, willingness has an influence on the simulated nonignorable dropout. Such a kind of weighting variable leads in average to bad mean estimations in the MAR and MNAR scenario as column 2 and 3 of table (3.4) shows.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1492.071	1523.667	1512.389
$Bias(\hat{\mu}_{\text{income}})$	-0.88042	30.71588	19.43825
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00059	0.02057	0.01302
$Var(\hat{\mu}_{\text{income}})$	121.22653	99.72755	122.418547
$MSE(\hat{\mu}_{\text{income}})$	122.00168	1043.19281	500.26422

Table 3.5: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of weighting (weighting variable: professional activity)

In contrast to the choice of willingness, now professional activity is chosen as weighting variable, which does not have an influence on the dropout, but is predictive for the values of income. Summarizing shortly, this choice gives an unbiased estimation for the MCAR case, but does not perform in a satisfying way otherwise.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1491.644	1509.600	1492.656
$Bias(\hat{\mu}_{\text{income}})$	-1.30712	16.64919	-0.30049
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00088	0.01115	-0.00020
$Var(\hat{\mu}_{\text{income}})$	125.22030	299.49475	245.53625
$MSE(\hat{\mu}_{\text{income}})$	126.92886	576.69030	245.62654

Table 3.6: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of weighting (weighting variable: education)

Finally, as a result of further deliberations, education is chosen as single weighting variable, since on the one hand it has an influence on income values, but also on the MAR and MNAR dropout. In all three cases this weighting approach results in good estimates with no or rather small bias, which are better than the estimates of ad-hoc, respectively.

Thus, weighting can lead to improved estimations compared to simply ignoring the missing data. However, prior knowledge concerning the weighting variables is necessary, since otherwise inappropriate weighting can even result in larger bias compared to simple ad-hoc methods. Moreover, this application has shown that the choice of weighting variables obviously depends on the underlying dropout mechanism. Correction has worked best for MCAR, when the weighting variable is a reliable predictor for the variable of interest. The estimation has not differed much from that one of ad-hoc, if the weighting variable is independent of the income variable. But bias has increased remarkably in the case of many predictive weighting variables. In contrast to that, the weighting correction is a success for MAR, if the weighting variables have an influence on the dropout and the variable values of interest. If you have the choice between a weighting variable that predicts the variable of interest, but not the dropout or a weighting variable that only predicts the dropout, but not the vaiable of interest, according to these analyses the first one should be chosen in the case of MAR. The conclusion about MNAR contains a surprisingly good outcome for the scenario of the weighting variable education. All the other performances yield biased estimates, in parts hardly better than in the ad-hoc approach.

3.5.3 Imputation

In the following section of single imputation, we are comparing the performances of the different single imputation strategies that have been presented in the theoretical part before. As a result, it should be worked out which of these approaches give better estimations compared to the simple ad-hoc method and under which conditions they are recommended to be a feasible option for the nonresponse problem.

Random Imputation
First, it is started out with simple random imputation which is conducted with replacement due to the large number of missing values.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1538.434	1575.896	1564.237
$Bias(\hat{\mu}_{\text{income}})$	45.48266	82.94523	71.28634
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	0.03047	0.05556	0.04775
$Var(\hat{\mu}_{\text{income}})$	173.877218	249.06419	243.98141
$MSE(\hat{\mu}_{\text{income}})$	2242.54928	7128.97514	5325.72409

Table 3.7: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of single imputation (random imputation)

Viewing the results for the average income mean estimation shows that in all three cases of MCAR, MAR and MNAR the deviations between the average estimates and the true values increase distinctly compared to the results of the ad-hoc methods. Also the MSEs of the estimates become large and take the values 2242.54928, 7128.97514 and 5325.72409, compared to the MCAR, MAR and MNAR reference values 145.20789 3072.81341 and 1520.97589 of the complete-case-analysis. Without going in greater detail, these results clearly indicate that in the case of mean estimation, random imputation should not be an option for dealing with missing data, independent of the underlying nonresponse mechanism.
Concentrating then on the evaluation of table (C.18) in the appendix, this result can be transferred to the regression coefficient estimation, too. Unless the intercept, in all of the other 21 cases of regression coefficient estimation, additional bias arises which leads to larger MSEs compared to the results of the simple ad-hoc procedure. Thus, random imputation does not improve parameter estimation, but instead in addition even obscures effects.
Table (C.19) contains the corresponding p-values which for all three nonresponse

scenarios obviously deviate from the true values. Thus, only highly significant effects can still be noticed in the imputed data sets, while all the other effects vanished, expressed through larger p-values. Only p-values of non significant regression coefficients decreased, which does not have an influence on significance interpretations. Thus, the performance of p-value estimation is worse compared to the ad-hoc performance. Of course, it cannot be expected that the p-values in all 21 regression coefficient cases are close to the true values, since about one third of the income values are missing respectively, so that a lot of the association structure in the remaining observed data cannot be recognized. However, a comparison with the performance of ad-hoc is feasible.

Mean Imputation

As next single imputation procedure, focus lies on mean imputation, which is one of the most common imputation procedures. Replacing missing values with the mean of the variable for responders naturally causes that mean estimation yields the same results as complete-case analysis. Thus, the following table (3.8) is identical to table (3.1) and will not be interpreted further.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1491.440	1546.533	1529.321
$Bias(\hat{\mu}_{\text{income}})$	-1.51143	53.58223	36.36994
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00101	0.03589	0.02436
$Var(\hat{\mu}_{\text{income}})$	142.92347	201.75756	198.20299
$MSE(\hat{\mu}_{\text{income}})$	145.20789	3072.81341	1520.97589

Table 3.8: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of single imputation (mean imputation)

It is still to be said at this point that mean estimation assumes that responder and nonresponder are on average alike which is not true for the cases of MAR and MNAR. Thus, in these cases single mean imputation does not make sense. In the remaining scenario of MCAR it is not worth expending the effort of conducting mean imputation, when focus is on mean estimation, since estimates are identical to simply ignoring the missing values. Thus, besides having now a full data set without missing values, there are no advantages compared to ad-hoc approaches, since the problem of a severely distorted distribution for the variable of interest, traced back to the appearance of missing values, remains and leads to complications with summary measures like the underestimation of standard errors.

Moreover, mean imputation does not lead to proper estimates of regression coefficients, as tables (C.20) and (C.21) show. As it has been the case before, when applying random regression, almost without exception regression coefficient estimates are more biased, have a larger MSE compared to ad-hoc and associations tend to be diluted even more.

Deterministic Regression Imputation

Staying in the field of single imputation, but going one step further regarding the usage of more information about the other observed variables, leads us to regression imputation. As has been described before, missing values are replaced by predictions based on a regression equation, established by the observed part of the data. Of course this approach yields better results, when good predictors are chosen. Since analyses here are based on the simulated data set, where the creation of the income variable knowingly depends on the covariates age, gender, education, professional activity, family status and election intention (see equation (4)), this best prediction model is known and chosen here, only deviating by the fact that instead of the logarithmic income as in the simulation process, the income is chosen as dependent variable. However, it is important to note that this level of knowledge is not given in reality, so that a real application of regression imputation would probably lead to worse predictions and consequently worse estimates. Nevertheless, in choosing those covariates as predictors, it can be shown which results can be reached by regression imputation in an ideal world of given information. Regression imputation with different regression models are not run and compared at this point, since it would exceed the extend of the thesis.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1492.481	1491.596	1482.350
$Bias(\hat{\mu}_{\text{income}})$	-0.47000	-1.35474	-10.60144
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00032	-0.00091	-0.00710
$Var(\hat{\mu}_{\text{income}})$	96.91094	55.67593	65.41347
$MSE(\hat{\mu}_{\text{income}})$	97.13183	57.51124	177.80390

Table 3.9: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of single imputation (deterministic regression imputation)

Interpreting the outcomes of the mean estimation based on regression imputation gives an obvious improvement of the estimates in all three nonresponse scenarios. For MCAR and MAR the mean estimates are in average unbiased and for MNAR

an absolute bias of 10.60144 is received. This is also a good performance which can be explained due to the fact that income does not have such a large influence in the MNAR model and all the other predictors of this missing model are taken into account in the regression model of this imputation approach. Then the corresponding MSEs of 97.13183, 57.51124 and 177.80390 are only a fraction of the corresponding ad-hoc values, which brings us to the conclusion that regression imputation can be a good option for mean estimation with missing values, if information about suggestive predictors are given.
Whether those implications are also true for the regression coefficient estimation, shows table (C.22). Here, a similar picture as in the other single imputation approaches arises: In general the biases and MSEs of the resulting regression estimates are larger than in the ad-hoc approach. This conclusion contains that if the aim is to fit a linear regression model with missing values in the dependent variable, it is better to ignore the missing values, independent whether they are following MCAR, MAR of MNAR, instead of predicting the values via a very good prediction model. Viewing the p-values in table (C.23), reveals a change compared to the impacts of the previous single imputation performances. The p-values of the different regression coefficients change in magnitude, but there is no systematic trend as before. In some cases they increase in contrast to the true values (e.g. $\beta_{\text{volks-, hauptschule}}, \beta_{\text{FDP}}$), in others they decrease (e.g. $\beta_{\text{married living together}}, \beta_{\text{die linke}}$). As a consequence, some effects seem to be significant at a certain significance level, even though this level actually is lower. But even more worth reporting than the attenuation of some effects is the fact that due to the smaller p-values in the regression imputed data sets, some effects become significant at a lower significance level than they actually are, which even peaks in the nonsignificant effect of $\beta_{\text{die linke}}$ that becomes significant at the level 0.1. The obtaining of statistical significance, when it should not be, is a huge disadvantage and can be ascribed to the underestimation of standard errors and the corresponding overestimation of statistical tests. The p-values resulting from the different nonresponse data sets are close to each other and only differ distinctly, in the case of $\beta_{\text{would not vote}}$, where the p-value of the MCAR case increases, the p-value of the MAR case stays quasi the same and the p-value of the MNAR case decreases compared to the true value, respectively. Thus, summarizing, all the significant effects are noticed, but the associations are assumed to be even stronger than they really are. These results show that regression imputation should be used with caution, when regression coefficient estimation is the main subject of the analysis.

Stochastic Regression Imputation

In order to handle the previously described disadvantage of the deterministic regression imputation, stochastic regression imputation is used which is based on additional random terms that increase the variance and leads to better standard error estimates.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1492.302	1492.470	1483.334
$Bias(\hat{\mu}_{\text{income}})$	-0.64878	-0.48114	-9.61653
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00044	-0.00032	-0.00644
$Var(\hat{\mu}_{\text{income}})$	169.77077	107.89755	140.95240
$MSE(\hat{\mu}_{\text{income}})$	170.19168	108.12904	233.43007

Table 3.10: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of single imputation (stochastic regression imputation)

This modification does not influence mean estimation remarkably, as table (3.10) shows. Deviations of the final point estimates from the true values are quasi unchanged in all three nonresponse scenarios. Only variances increased as hinted before, which leads to larger MSEs. However, this loss can be taken, if there comes along an improvement in the estimation of regression coefficients or rather concerning the resulting p-values.

In order to find out, next tables (C.24) and (C.25) are viewed. Compared with the results of deterministic regression imputation the regression coefficient estimates here in general are less biased, but the corresponding MSEs are smaller in only about half of the cases. Even higher interest is on the change of the p-values. Indeed, compared to deterministic regression imputation the p-values of stochastic regression imputation increased, which results in the consequence that statistical significances, where they should not be, do not appear anymore. However, due to the increased p-values the problem of dilution of effects comes to the fore again. In the cases of $\beta_{\text{mittlere reife}}$, $\beta_{\text{volks-, hauptschule}}$ and β_{FDP} statistical effects are not even apparent for the cases of MCAR, MAR and MNAR, whereas ad-hoc methods here show weakly significant effects. In the cases of $\beta_{\text{no graduation}}$, β_{single}, $\beta_{\text{would not vote}}$ this tendency of dilution can be viewed, too, but effects remained statistically significant at a higher significance level compared to the ad-hoc and the true values. Especially for the MNAR scenario, those effects often remain visible. Thus, stochastic regression imputation defrauds of even more statistical significant effects than deterministic regression imputation, but it does not create statistical significant effects, where there are none.

Predictive Mean Matching Imputation

As a last possible single imputation approach, methods remain that are based on donors. Here we are viewing the performance of the predictive mean matching imputation, which is, as described in the theoretical part before, a form of nearest-neighbor hot-deck imputation with a specific distance function.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1489.049	1536.064	1519.272
$Bias(\hat{\mu}_{\text{income}})$	-3.90224	43.11290	26.32063
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00261	0.02887	0.01763
$Var(\hat{\mu}_{\text{income}})$	244.57129	310.15347	326.29167
$MSE(\hat{\mu}_{\text{income}})$	259.79879	2168.87587	1019.06746

Table 3.11: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of single imputation (predictive mean matching imputation)

The resulting average mean income estimation in table (3.11) does not entail an improvement compared to the (stochastic) regression imputation before, since biases and MSEs in all three scenarios are far larger. Yet, these results are better than the average estimates of the ad-hoc approach, of the random imputation and mean imputation. Consequently, this approach can only convince in this application, if its performance in regression coefficient estimation and especially the resulting p-values stands out positively.

Viewing table (C.27) in the appendix gives that the p-values increased much, so that hardly any of the statistically significant effects remain statistically significant in the cases of imputed values. Since this result is even worse than in the ad-hoc method, table (C.26), containing the regression coefficient estimates, will not be discussed.

Thus, it can be summarized conclusively that random imputation and mean imputation did not at all provide good results and should therefore not be chosen in order to handle missing data, independent of the underlying dropout mechanism. Deterministic and stochastic imputation yield very good mean estimates for MCAR, MAR and MNAR which might, however, be ascribed to the well-known and chosen prediction model. In the case of predictive mean matching imputation, where the prediction model is not chosen by hand, the mean estimation obviously has a larger bias and MSE. Furthermore, all three single imputation methods, based on prediction models, did not perform satisfactorily with respect to p-value estimation. Deterministic regression imputation stands out, since here all effects are remained in the imputed data set, but due to the underestimation of some p-values these results have to be interpreted with caution as well. In contrast to that, stochastic regression imputation and predictive mean imputation diluted statistical effects too much, compared to the simple ad-hoc approach.

As it has been mentioned before, instead of single imputation, trends go towards multiple imputation, since this approach reflects the uncertainty of the estimation by replacing each missing value with several imputed values. Orientating on literature, the number of imputed values here is set to 5. Due to the bad performances of random and mean imputation before, those approaches are omitted for the multiple imputation and focus is set on regression imputation and predictive mean matching imputation which are both performed with the **R** package **'mice'** (version 2.22, see Van Buuren and Groothuis-Oudshoorn (2011)).

Multiple Regression Imputation

Starting out with multiple regression imputation, the resulted income mean estimate of the MCAR scenario still has the smallest bias, but with a value of -2.05960 it is not completely unbiased as it has been in the single imputation version. In the case of MAR and MNAR, however, the estimates have a far larger bias of 40.46051 and 25.12561. Due to the additional variation, variances of the estimates increased, which leads to larger MSEs in all three cases.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1490.891	1533.412	1518.077
$Bias(\hat{\mu}_{\text{income}})$	-2.05960	40.46051	25.12561
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00138	0.02710	0.01683
$Var(\hat{\mu}_{\text{income}})$	812.23570	855.858657	809.56319
$MSE(\hat{\mu}_{\text{income}})$	816.47764	2492.91117	1440.85935

Table 3.12: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of multiple imputation (multiple regression imputation)

Regarding the coefficient estimates in table (C.28) shows that estimates of the multiple regression approach are indeed differing from those of single imputation, which results in the fact that for some regression coefficient estimates corresponding MSEs are smaller for the multiple imputation approach, for others they are larger compared to the single regression imputation approach. Therefrom no conclusions can be drawn, which is why focus is set on the resulting p-values. Due to the increased variability of the multiple imputation, which comes along with the larger number of imputed values, standard errors are less underestimated and consequently p-values become larger. Therefore, especially the resulted p-values of the multiple regression imputation are of interest, since in the single regression imputation case, tendencies could be noticed that some of the estimated p-values were too large, as in nearly all other procedures, but some others were indeed underestimated. Here, the p-values became very large, which means that hardly any of the statistical effects can still be seen in the imputed data sets. All those findings do not change remarkably when the number of imputed values is increased from $T = 5$ to $T = 10$. Then, mean estimation improves a little bit for the MCAR scenario, while it changes little to the worse for MAR and MNAR. Influences on regression coefficient estimation are not considerable, so that resulting p-values remain large and statistical effects are hardly visible. Thus, imputing missing values multiple times by a regression model obviously obtains the statistical association structure less than simply ignoring the missingness and testing significances in the observed part of the data.

Multiple Predictive Mean Matching Imputation

Even better comparability with the single imputation approach enables the predictive mean matching method, since here also the single imputation approach has been applied with the help of the **'mice'** package. Comparing first the income mean estimates gives that except of the increasing variance, which results in larger MSEs, the estimates in all three scenarios do not differ remarkably from those of the single approach.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1496.957	1534.194	1518.697
$Bias(\hat{\mu}_{\text{income}})$	4.00636	41.24305	25.74577
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	0.00268	0.02763	0.01725
$Var(\hat{\mu}_{\text{income}})$	806.57083	851.27628	805.54525
$MSE(\hat{\mu}_{\text{income}})$	822.62178	2552.26566	1468.39009

Table 3.13: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of multiple imputation (multiple predictive mean matching imputation)

Comparing the regression coefficient estimates with those of the single imputation version gives no clear tendency towards the degree of unbiasedness. In some of the regression coefficient cases both or just one of the statistical measurements of bias and MSE are larger, when missing values are imputed multiple times, in other cases they are larger, when a single value is imputed via predictive mean matching. Thus, special attention is payed to the p-values, in order to find an answer to the question, whether multiple imputation here improves the estimation compared to single imputation. It shows that the resulting, increased p-values again tend to dilute associations and heavily prevent from noticing statistical effects in the imputed data sets.

In conclusion, against expectations those two multiple imputation approaches did not improve the results in this application. There is no clear tendency towards smaller MSEs of the estimates for the mean income and the regression coefficients and furthermore p-values are too large, so that significant statistical effects are hidden. Those results will not change in the case of increased number of imputed values. Choosing $T = 10$ instead of $T = 5$ will not improve the estimation in general. While mean estimation in all three scenarios has nearly the same average bias and only small reduction concerning the MSE, the MSEs of estimated regression coefficients tend to become only a little bit smaller, but in a range that does not influence p-values and consequently significance of statistical effects.

3.5.4 Maximum Likelihood Procedures

FIML

First of all, it is started out with parameter estimation based on maximum likelihood procedures ignoring the dropout process. At this, the first estimation approach is FIML which is conducted in **R** with a structural equation model established with the package **'lavaan'** (version 0.5.17, see Rosseel (2012)). Before application, the data has to be prepared, since **'lavaan'** is not able to handle nominal variables. Thus, the variables education, family status and election intention are recoded into dummy variables which are then incorporated in the appropriate model. The income variable serves as dependent variable and no latent constructs are included, so that the SEM then equals a linear regression model. In order to receive an estimation for the income mean, first, the appropriate regression model is fitted and estimated via full information maximum likelihood, meaning that all available information are used to set up the likelihood for a certain observation before summarizing over all individual likelihoods to receive the likelihood for the entire data (see equations (10) and (11)). As a result, regression coefficient estimates are received which are used to predict the missing income values. Finally, the average of these predictions, is used as estimate of the interested mean income. The results are presented in the following table.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1492.481	1491.596	1482.350
$Bias(\hat{\mu}_{\text{income}})$	-0.46977	-1.35493	-10.60141
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00032	-0.00091	-0.00710
$Var(\hat{\mu}_{\text{income}})$	96.91332	55.66915	65.41175
$MSE(\hat{\mu}_{\text{income}})$	97.13400	57.66915	177.80175

Table 3.14: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of Maximum Likelihood Procedures (FIML estimation in Structural Equation Model)

Obviously, in all three scenarios, the average mean estimates are closer to the true values as in the ad-hoc approach. For MCAR and MAR the parameters are not biased and for MNAR, where the application of the modelbased approaches ignoring the missing data mechanism is not justified, a bias of size -10.60141 arises. However, this still is smaller than in most other ad-hoc, weighting and imputation applications. Moreover, variances of the estimates are small, which results in comparable small MSEs and the conclusion of a good performance. These good results in the case of MNAR seem unexpected at first sight, since FIML assumes a MAR dropout. However, it should not be left behind that also in the MNAR model, dropout depends, besides the income itself, on other variables, which are considered in the FIML estimation, so that valid information are used. Moreover, the simulated MAR and MNAR model do not deviate too much, since in the MNAR model income does not have such a large influence on the dropout of the income values. Consequently, the performance of FIML would probably worsen, if another MNAR model would have been constructed and evaluated that is based on a stronger influence of the income variable itself.

In order to receive regression coefficient estimates based on FIML the approach is similar to that of mean estimation. However, the fitted regression models differ with respect to the dependent variable, since now the logarithmic income is used, which is consistent to the usual regression model of interest before. Moreover, the prediction step is not necessary, but the received regression coefficient estimates, listed in table

(C.32) can be interpreted directly. They are close to those of the ad-hoc procedure and in not only a few cases the resulting MSEs are even a little bit smaller. Those smaller MSEs in the case of FIML estimation occur in all three scenarios. Due to the previously described theoretical background, which stated at least the assumption of a MAR dropout, no improvement in the case of MNAR is expected. However, this result should not be attached too much importance, since those differences account only for small decimal places and possible reasons have already been adduced in the section of mean estimation before. Furthermore, table (C.33) with the corresponding p-values shows that for all three dropout mechanism highly significant effects remain highly significant, moderate significant effects weaken and non significant effects remain non significant. Actually, this result is nearly identical to that of the ad-hoc procedure, where all statistically significant effects at least are noticeable.

EM Algorithm

Using next the EM algorithm for estimating the income mean in the presence of incomplete data, a parametric model specification is required, in order to establish and maximize the likelihood function. Here, the basic assumption is normal distribution for the continuous variable of interest, having several missing values. Applying then the EM algorithm, in order to receive an estimate for the mean income gives the average value of the observed income values, which is equal to the ad-hoc result. A step further, making sense with regard to the underlying simulation situation, is the expansion of the normal assumption to more than just the income variable, equaling the multivariate normal model described in Dempster et al. (1977). The **R** package **'norm'** (version 1.0.9.5, see Novo and Schafer (2013)) provides the facility to perform maximum-likelihood estimation on a matrix of incomplete data using the EM algorithm. It will consequently be used and the resulting mean estimates are presented in the following table.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1492.482	1491.651	1482.400
$Bias(\hat{\mu}_{\text{income}})$	-0.46926	-1.30011	-10.55061
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.00031	-0.00087	-0.00707
$Var(\hat{\mu}_{\text{income}})$	96.88839	55.68690	65.45040
$MSE(\hat{\mu}_{\text{income}})$	97.10859	57.37717	176.76574

Table 3.15: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of Maximum Likelihood Procedures (EM algorithm)

Since those results are quasi equivalent to those of FIML estimation, they will not be interpreted further. However it is noteworthy that this good performance might not pertain to situations, where the assumption of normality is unreasonable. Therefore, it would be interesting for other data situation to evaluate this approach and also to expand it to applications assuming other parametric models, thus different to the normal model. Until now, however, no general implementation of the EM algorithm for handling missing data is provided in **R**, which is a useful outlook for future research in this field of application. Moreover, the **'norm'** package is limited to the estimation of means, variances, covariances and correlation coefficients. This is why regression coefficient estimation with the EM algorithm is not applied in this thesis. However, as it has been described in the theoretical part before, regression coefficient estimates will be analogous to those of FIML, but standard errors calculation would be more complex, which is why FIML should here be preferred nevertheless.

Heckman Selection Model

After application of the Maximum Likelihood estimation via FIML and EM algorithm, which are based on the ignorance of the missing process and therefore are theoretically only suggestive for the scenarios of MCAR and MAR, focus next is set on the application of maximum likelihood approaches, which are based on the explicit modeling of the dropout process. The previously introduced Heckman SM is the first of those approaches. In **R** it can be applied with the **'selection'** function of the package **'sampleSelection'** (version 1.0.2, see Toomet and Henningsen (2008)). Since this function is only able to deal with numeric and binary variables, the respective variables have to be recoded, as it has already been necessary in the application of FIML estimation before. Furthermore, a dichotomous variable "observed" has to be constructed that takes value 1, if the interested variable income is observable and value 0 otherwise. Thus, it is the complement of the previously used missing indicator variable M. As described in the theoretical part concerning the SM, besides the analysis model of interest a selection model has to be established which are both defined within the **'selection'** function. At that, the "observed" variable acquires the position of the dependent variable of the selection model. Gender, education and willingness are chosen as corresponding covariates, since those are the variables used to simulate the MAR dropout process. So again, more information is used than is usually available in a real world application. However, this way, the best possible performance of the Heckman SM can be seen, since the dropout mechanism is modeled as good as possible. The second defined model within the **'selection'** function again is consistent with the regression model of all other applications (see equation (4)). For estimating the mean income, however, this regression model as part of the Heckman SM is fit with the income as dependent variable, instead of the logarithmic income. The results of the analysis model then serve to predict the missing income values by considering the correction term consisting of the inverse Mills Ratio and the selection bias. Herefrom, the following estimates arise which are received by applying the estimation method **'2step'**, which is equivalent to the original Heckman 2-step-procedure:

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1532.048	1499.985	1495.219
$Bias(\hat{\mu}_{\text{income}})$	39.09717	7.03424	2.26804
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	0.02619	0.00471	0.00152
$Var(\hat{\mu}_{\text{income}})$	88375.310019	1048.99845	972.67315
$MSE(\hat{\mu}_{\text{income}})$	89903.989898	1098.47905	977.81714

Table 3.16: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of Maximum Likelihood Procedures (Heckman Selection Model)

Obviously, this approach may result in good performances, if the dropout mechanism is modeled close to the true underlying dropout process which, however, actually is not known. Here, the average estimate of the MNAR scenario is hardly biased. The relative bias of 0.00152 is very small and so is the corresponding MSE of 977.81714, compared for example with those of the ad-hoc approach, having value 1520.976. Also in the MAR scenario, the resulting estimates are far better than in the ad-hoc approach. The average mean estimate of 1499.985 deviates only by 7.03424 from the true value and the MSE of 1098.47905 is quite small, too. Thus, in these two cases the application of this likelihood based approach is reasonable. This is not the case for the MCAR scenario. Here, an assumption for the dropout mechanism is made in the modelbased approach that is not accurate. Consequently, it is not surprising that the outcome is bad compared to the ad-hoc performance or other simpler weighting and imputation approaches. However, if the aim is to estimate the mean of a variable having missing values following MAR or even MNAR, the Heckman SM yields good results. In order to evaluate the performance for different selection models, the Heckman SM for the mean income estimation is also applied with changes in the selection equation. On the one hand only the variable willingness is used to explain the dropout and on the other hand the variables willingness and family status are chosen. At this point it is important to note repeatedly that the dropout is simulated independent of the latter one. The results, which are presented in table (3.17), but which are limited to the average estimates, show a change for the worse. Nearly all resulting estimates, independent of the underlying dropout mechanism, deviate more from the true values than in the previous case, when the selection model has been fit with all the dropout variables. The only exception is constituted in the MCAR scenario, where family status is incorporated in the selection model. The estimates become better, since a wrong dropout variable is assumed in a scenario, where the dropout actually is arbitrary. However, in the case of MAR and MNAR, the estimates for the new selection models are still far better than in the ad-hoc approach which here pleads for the application of the Heckman SM, even if not too many detailed information are known about the dropout.

selection variables		MCAR	MAR	MNAR
	μ_{income}	1492.951	1492.951	1492.951
willingness	$\mathbb{E}(\hat{\mu}_{\text{income}})$	1525.629	1505.043	1499.039
willingness, family status	$\mathbb{E}(\hat{\mu}_{\text{income}})$	1529.352	1508.392	1502.106

Table 3.17: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of Maximum Likelihood Procedures (Heckman Selection Model with different selection variables)

If the aim is to receive regression coefficient estimates in such situations, the Heckman SM can also be applied. However, the result slightly differs from the outcomes of other approaches, since here, an additional covariate is included in the analysis regression model of interest, namely the inverse Mills Ratio, whose value is also presented in the outcome. Nevertheless, the other estimated regression coefficients are comparable with those of other applications, since they are already corrected with respect to the bias that possibly arose due to the censoring of certain observations. The results, presented in table (C.34) in the appendix, show regression coefficient estimates that are drifting away from the true values slightly more than the ad-hoc estimates. Also the p-values are larger than in the ad-hoc approach and all the more larger than the true values. While the significances of the statistical effects remain the same as in the ad-hoc approach for MAR and MNAR, in the MCAR scenario some significant effects like the one of "mittlere reife", "no graduation" and "volks-, hauptschule" are not apparent anymore. Thus, the Heckman SM could convince in the context of mean estimation, but not for regression coefficient estimation, even though also here full priori knowledge was used. In the case of wrong specification, the estimates are expected to become worse, so that the application of the Heckman SM is only justified for cases of known information about the dropout mechanism.

Pattern Mixture Model

The application of the PMM, which is the second approach, based on explicit modeling of the dropout process can either be based on further restrictions of the missing mechanism or on priori information, as has been explained before in the theoretical chapter. First, in the case of mean income estimation, the approach based on restrictions is briefly mentioned, but due to reasonableness not carried out much further and focus will then be set on evaluation of the approach based on priori information.

Restricting the missing mechanism according to Little (1994), sets focus on the two variables income and age, since those restrictions are based on normal data and all the other variables, which in fact are simulated as having an influence on the dropout, are categorical. Age, however, actually does not have an influence on the dropout, so assuming MAR in the PMM does not make sense, since it means that the dropout only depends on age. Consequently, then the mean income is biased for all three simulated dropout scenarios. In the case of MCAR a mean income of 1491.498 is received, for MAR 1552.947 and for MNAR 1535.67, which shows that this assumption is not suitable. Hence, formulating a statistical model that assumes MNAR instead of MAR improves the results, since then, besides the influence of age, also the influence of the income values itself on the dropout is assumed, which is at least closer to the simulated data background. However, assuming this scenario means that it has to be predefined in which way the dropout depends on those two variables. This is done via an assumed linear combination of the form $X_{\text{age}} + \lambda \cdot X_{\text{income}}$, whereas λ is assumed to be not zero and furthermore known (cp. Little (1994)). Determining λ in a suitable way can give good results. For $\lambda = 0.42$ for example this PMM gives a mean estimate of 1497.077 for the underlying MAR data situation and 1493.370, if the data is MNAR. As expected this estimate is bad for MCAR, resulting in an mean income estimate of 1239.602. However, it becomes clear that those results strongly depend on the choice of λ, requiring priori information, not only about which variables are responsible for the dropout, but also about there magnitude of influence compared to the other selection variables. This knowledge, however, is unrealistic and even in this situation, based on simulated data, λ has been chosen by trial and error.

In real life situations, however, it is more often the case that priori information are given concerning the connection of responder and nonresponder. Then the PMM is suitable, as the following approach and the corresponding evaluation shows. Here, the missing data indicator M serves as dependent variable of a logit model, where the variables gender, education and willingness are chosen as influential covariates according to the simulation process presented in equation (5). This model then results in predictions for the individual dropout probability. In order to first estimate the average income mean herefrom, equation (16) in the theoretical chapter of the PMM is the basis. The average income mean of the responder is calculated and an assumption for the connection between the mean of the responder and those of the nonresponder, which is unknown, is made. In this case it is assumed that the mean of the nonresponder is a function of the mean of the responder, which here is determined by $\mu_{\text{income}|\text{M}=1} = 0.9 \cdot \mu_{\text{income}|\text{M}=0}$, meaning that here the mean of the nonresponder is in average 90% of the income mean of the group of the responder. This relationship is not chosen randomly, but it orientates on the true underlying connection of the income means of the two groups. Since the true income values are known before the dropout process was applied, the income mean of the stratum of the nonresponder is known here in contrast to real data situations. It appeared that in the MAR scenario, in average the income mean is 88.844% of the responder mean and in the MNAR scenario the corresponding value is 92.354%. Thus, the relationship in the model is determined by 90%, which lies in the middle. Obviously, the performance of the PMM strongly depends on this assumption, which is why, the application will then be also conducted for different assumptions. But first, this approach will be applied further. With the help of the predicted individual dropout probabilities and their corresponding converse probabilities, for each observation it will be sampled between the mean of the responder and the assumed mean for the nonresponder. The average over all those sampled values then resembles the corresponding income mean averaged over the pattern of responder and nonresponder. The following table (3.18) gives an overview of the results.

	MCAR	MAR	MNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1444.873	1498.356	1481.817
$Bias(\hat{\mu}_{\text{income}})$	-48.07786	5.40515	-11.13352
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.03220	0.00362	-0.00746
$Var(\hat{\mu}_{\text{income}})$	142.13432	189.41969	242.76899
$MSE(\hat{\mu}_{\text{income}})$	2453.61493	218.63529	366.72420

Table 3.18: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of Maximum Likelihood Procedures (Pattern Mixture Model)

The MCAR column will not bet interpreted further, since the assumption of this functional dependence structure between responder mean and nonresponder mean is not true in this scenario. Therefore, the bad results are according to expectations.

Obviously, the assumption of identical income means between responder and nonresponder would be correct here which would lead to the good ad-hoc estimates. In the cases of MAR and MNAR, however, the PMM performs good. The average income estimates are comparable with those of the Heckman SM, whereas in the MAR scenario the estimate is even less biased and in the MNAR scenario the average income estimate deviates more from the true value. However, both average estimates do not have a large bias and have, due to the small variances, small MSEs. Thus, this approach is a good alternative, if information about the dropout are known. Here, the knowledge about the influential dropout variables is used and furthermore the knowledge about the functional dependence structure between responder and nonresponder income mean.
In order to further evaluate the performance of the PMM, different assumptions are made. First, the underlying logit model is changed with respect to the independent variables, while the dependence structure between income means remain the same. As in the Heckman SM, here only the variable willingness is chosen and the variables willingness and family status. Second, the assumption that the nonresponder mean accounts for 90% of the responder mean is changed to 85%, 95% or rather 50%, while the dropout mechanism is modeled by the original logit model. The following table summarizes the results, but it is again limited to the average estimates.

selection variables		MCAR	MAR	MNAR
	μ_{income}	1492.951	1492.951	1492.951
willingness	$\mathbb{E}(\hat{\mu}_{\text{income}})$	1444.897	1498.325	1481.612
willingness, family status	$\mathbb{E}(\hat{\mu}_{\text{income}})$	1444.911	1481.544	1481.544
$\mu_{\text{income}\mid M=1} = g \cdot \mu_{\text{income}\mid M=0}$		MCAR	MAR	MNAR
$g = 0.85$	$\mathbb{E}(\hat{\mu}_{\text{income}})$	1421.590	1474.268	1458.066
$g = 0.95$	$\mathbb{E}(\hat{\mu}_{\text{income}})$	1468.156	1522.445	1505.569
$g = 0.50$	$\mathbb{E}(\hat{\mu}_{\text{income}})$	1258.607	1305.648	1291.804

Table 3.19: Expected Value of income in order to evaluate the performance of Maximum Likelihood Procedures (Pattern Mixture Model with different selection variables and different assumptions for the relationship between patterns)

It can be seen that changes in the independent selection variables of the underlying nonresponse model and consequently a small change in the individual dropout probabilities do not have strong influences on mean estimation. In fact, then the average mean estimation only changes with respect to decimal places. In contrast to that the assumption concerning the relationship of the income mean of the different groups, has a larger influence. As before, attention is not focused on the MCAR scenario, since average estimates are bad for all different cases. In the scenarios of

MAR and MNAR, however, it can be seen that if g and consequently the average mean of nonresponder is assumed larger than it actually is, bearing relation to the mean of the responder, it follows that the resulting income mean for all observations is overestimated. Smaller g consequently result in an underestimation. The bias is remarkable, even if g deviates only little from the true underlying relationship. For the completely inappropriate value of $g = 0.50$ the bias is huge. This results in the conclusion that PMMs for mean estimation only yield good estimates, in the case of exact prior information.

Focusing next on regression coefficient estimation, one encounters difficulties with this approach, since assumptions about the relationship between the beta coefficients of responder and nonresponder have to be made. While it is standing to reason to make assumptions for the differences in mean, there are no reasons to determine how regression coefficients should change for nonresponder compared to responder. Moreover, analyzing the true underlying data structure here shows that regression coefficients do not change in the same directions, which means that some beta estimates are in average larger for nonresponder compared to responder, others are equal or smaller. Consequently, this way of regression coefficients estimation is doubtful, which is why, it is only applied for the optimal case, meaning that the powerful selection variables are considered and furthermore the relationship of the responder regression coefficients and nonresponder regression coefficients is modeled due to the underlying true link. The results, presented in table (C.36) are close to the true values, nevertheless, those results will never be reached in a real life applications, since too much prior knowledge is used. Consequently, sensitivity analysis are necessary here and the idea of partial identification establishes. Due to the scope of this thesis, however, the approach of partial identification is confined to theoretical introduction, but not applied here. Thus, in future works this idea should be followed up further, in order to expand the applications PMMs towards parameter estimation other than mean estimation. For this purpose also the implementation of the PMM in **R** would be desirable.

4 Limits of Simulation and extended Applications due to Points of Criticism

After evaluation of the impacts of the different error sources in chapter 2 and the performances of different approaches for handling missing data in chapter 3.5, a final summary of the results (see chapter 5) is remaining. However, many of the outcomes of this thesis may not simply be generalized, since they strongly depend on the underlying simulated data basis and on decisions that had been made in this context. Therefore, this chapter shall give an overview over the limits and points of criticism concerning the simulation and application of different missing data procedures which might have led to overoptimistic results compared to real life situations. Some of these aspects have already been mentioned in the course of the thesis, nevertheless, they shall be pointed out here again and carried forward.

First of all, the similarity between the data basis and the Allbus 2014 has been ensured as effectively as possible, nevertheless not all the relationships between the different variables could have been considered and adopted. Consequently, a possible point of criticism might arise, since some variables like age and professional activity are not simulated with a certain dependency, which in reality would clearly be the case. Another difference between the simulated data basis and the Allbus 2014 concerns the variable of interest income. The simulated income values are truncated at a maximum of 10904, whereas the income variable of the Allbus data set contains very few upward outliers, exceeding this threshold. Those large income values obviously have a strong influence on parameter estimation. Drawing for example samples from a data set containing large outliers, then results are expected, to differ in the cases of inclusion or exclusion of these powerful observations. In contrast to that, in a data situation without remarkable outliers, which is the case in the underlying simulated data situation, error sources like sampling error or coverage error, will not have such a remarkable impact on parameter estimation, expressed through smaller biases and MSEs. Thus, those non-existing outliers might be a reason for the good parameter estimations in the scenarios of present error models.

Another reason for the small biases and MSEs in the presence of constructed errors might be the small statistical effects which are estimated in the Allbus data set and are then also assumed for the simulated data set. Comparable small β-coefficient for the different categories of education, for example, mean that this variable does not have a strong influence on the logarithmized income. If then for example, in the context of coverage error, the sample probability for high educated people is

doubled and for others remained constant or even is reduced, the estimated outcome parameters in this situation are not biased too much, when the logarithmized income does not differ remarkably in the different education categories. The same phenomenon can also be viewed in the presence of other error models. If, for example, the dropout of the income depends amongst other variables on education, but the income itself does not differ a lot for the different categories of education, this dropout model will not result in parameter estimates that are strongly biased.
Furthermore, in order to match with the range of the other β-coefficients, also the influence of the income itself on the dropout in the MNAR model (cp. equation (6)) is chosen comparatively small with a value of $\beta_{\text{income}} = 0.00005$. This choice has been justified in section 2.2, nevertheless, the determination of this fixed value probably has a large influence on further analyses. Due to this value, the MAR and MNAR model, all analyses of the second part of the thesis are based on, do not differ too much from each other, since other predictors are the same in those models and only the additional dropout-influence of the income, which, however, is quite small, makes the difference. This similarity of the dropout-models probably is the reason, why FIML estimation and the EM algorithm, both assuming MAR, also performed surprisingly well in the MNAR scenario. In order to check this supposition, another MNAR model is set up that differs from the evaluated MNAR model with respect to the chosen β_{income} value. While the effect has been determined by $\beta_{\text{income}} = 0.00005$ in the logit model before, the influence of this variable on the dropout is now remarkably increased by changing it to $\beta_{\text{income}} = 0.0005$. Since this new determination leads to higher predicted, individual dropout-probabilities, a larger number of missing income values ensues, which, however, prevents from comparing the impacts of the different nonresponse models. Consequently, the estimated dropout-probabilities are shrunken by subtracting 0.147, respectively, as it has already been proceeded in chapter 3.4, so that in average we receive approximately the same number of missing income values. Nevertheless, in general now different income values are missing in the presence of the different nonresponse models. Thus, the new underlying data situation differs remarkably from the MNAR scenario before and consequently also from the MAR scenario and the performances of certain correction methods can again be evaluated. Focus hereby is set on the performance of the complete-case-analysis for reasons of comparison and the different likelihood-based approaches. The performance of the EM algorithm is spared out, since it gives point estimates nearly identical to those of FIML estimation. For the application of the PMM the assumptions regarding the relationship between responder and nonresponder is updated and adapted due to the new underlying data

situation, where now nonresponder are assumed to have a 1.35 times larger average income than nonresponder. So again full information knowledge is used and the application of the PMM confines to the ideal case of prior knowledge and no sensitivity analysis is performed at this point. Of course, the results of the PMM then have to be regarded critically, since in reality results will be nowhere near these good performances. The following table refers to the mean income estimation, while table (C.37) in the appendix evaluates the corresponding regression coefficient estimation.

	ad-hoc	FIML	Heckman SM	PMM
μ_{income}	1492.951	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1343.804	1372.319	1403.999	1490.412
$Bias(\hat{\mu}_{\text{income}})$	-149.14719	-120.63248	-88.95165	-2.53903
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.09990	-0.08080	-0.05958	-0.00170
$Var(\hat{\mu}_{\text{income}})$	196.14158	119.18878	1121.90554	264.12542
$MSE(\hat{\mu}_{\text{income}})$	22441.02713	14671.38374	9034.30205	270.57211

Table 4.1: Expected value, variance, bias and MSE of mean estimation of income in order to evaluate the performance of different correction methods on alternative MNAR model

Both tables show that this modified MNAR model obviously results in larger biases after application of different nonresponse methods. Especially in the case of FIML strong deviations from the true values can be recognized. This confirms the previously made supposition that the likelihoodbased methods, ignoring the nonresponse mechanism, only performed that good in the case of MNAR, because the influence of the income variable itself on the dropout was chosen small and consequently the MNAR model was similar to the MAR model. In the now presented modified MNAR model, however, we can see a clear domination of the likelihoodbased methods that explicitly model the nonresponse mechanism.

Another explanation for the good performance of FIML and EM algorithm in the MNAR scenario, described and evaluated in the thesis before, might be the fact that the variables gender and education are predictors of both models, the dropout-model (cp. equation (6)) and the regression model that is used to construct the income variable itself (cp. equation (4)). Thus, when applying the methods FIML and EM algorithm, which are based on the MAR assumption, we are indeed not including the income as predictor, but other variables which partly explain the income values. If the predictors of the dropout-model completely differ from those of the regression model that is used to construct the income variable itself, a larger bias of the estimated parameters is expected. However, this aspect should be checked

in another setting, since here, the modification would not be compatible with the underlying data situation, where those two variables are indeed simulated as having an influence on both, income and dropout.
Another point of criticism contains that in general the considered error models, explained in chapter 2.2, are very specific. Thus, conclusions like “sampling error does not influence mean estimation” or “coverage error has a larger impact than other error sources” are not valid. Instead, the simulation background has to be involved in the interpretation and results have to be questioned critically. In the case of sampling error, the impact obviously depends on the sample size which is determined to certain values in this thesis, namely $n = 500, 1000, 2000, 3000$. In order to make general statements, even smaller critical sample sizes should be regarded. Choosing for example $n = 100$ does not give unbiased estimates anymore, but an average mean estimate of 1498.893. Of course, results would also change in the field of coverage error, if sample sizes were chosen differently. For the case of $n = 100$ for example, the coverage error model would result in an average income estimate of 1366.91. Concerning this error source, also the choice of sample probabilities, listed in table (2.1) are decisive. They are determined on literature basis here, but are nevertheless chosen quite arbitrary. Therefore, at this point two further coverage models are viewed which differ with respect to the assigned drawing probabilities. The first one favors low educated people, which means they now have a doubled probability of being sampled, whereas middle educated people still have a probability of $\frac{1}{N}$ and the probability of high educated people is halved. So here, we have a reversed probability distribution compared to the analyzed scenario in the thesis. The differentiation between male and female participants, however, remains as before, so that women are still overrepresented. The second additional coverage error model does not consider the variable gender, but differentiates only with respect to education, whereas this marginal drawing probability remains like in table (2.1). Again those additional models are only evaluated for mean income estimation and results are listed in tables (4.2) and (4.3). As expected, the impacts of the coverage error strongly depend on the chosen sampling probabilities. In the case of an overrepresentation of low educated people, smaller average income values compared to table (2.4) result for all sample sizes. If the sample probability does not depend on gender, but the marginal sampling probabilities of the different education categories remain as chosen before, better estimates are received, since the income difference between men and women does not play a role anymore. Thus, not only the variables causing the coverage error are decisive for the impacts on parameter estimation, but also the proportion of sampling probabilities between different groups of people.

	$n = 500$	$n = 1000$	$n = 2000$	$n = 3000$
μ_{income}	1492.951	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1271.826	1289.486	1342.178	1432.835
$Bias(\hat{\mu}_{\text{income}})$	-221.12528	-203.46508	-150.77285	-60.11562
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	-0.14811	-0.13628	-0.10099	-0.04027
$Var(\hat{\mu}_{\text{income}})$	1561.66306	553.85316	155.13169	51.95490
$MSE(\hat{\mu}_{\text{income}})$	50458.05251	41951.89194	22887.58248	3665.84267

Table 4.2: Expected value, variance, bias and MSE of mean estimation of income in alternative coverage model with overrepresentation of low educated and female units in order to quantify coverage error

	$n = 500$	$n = 1000$	$n = 2000$	$n = 3000$
μ_{income}	1492.951	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1505.806	1500.013	1499.617	1495.699
$Bias(\hat{\mu}_{\text{income}})$	12.85476	7.06166	6.66633	2.74786
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	0.00861	0.00473	0.00447	0.00184
$Var(\hat{\mu}_{\text{income}})$	1638.64462	734.88158	227.72206	50.16994
$MSE(\hat{\mu}_{\text{income}})$	1803.88947	784.74862	272.16202	57.72069

Table 4.3: Expected value, variance, bias and MSE of mean estimation of income in alternative coverage model with overrepresentation of high educated units in order to quantify coverage error

Also in the case of nonresponse and measurement error, there are factors determining magnitude and direction of impacts on parameter estimation. While in the error scenarios of MCAR and RCAR the number of missing or rather rounded values is decisive, in the case of MAR, MNAR, RAR and RNAR specific models are considered, whose meaningfulness and validity may not be tested. In the context of measurement error for example, it is concentrated on rounding errors, whereby the corresponding model assumed rounding exclusively to the next threshold of 10. This kind of error model hardly biased mean estimation which suggests the assumption that rounding to the next threshold of 10 is a too optimistic model that is out of proportion with reality. Therefore, additional measurement error models with the same construction background, as described in section 2.2, are viewed that in contrast now consider rounding to the next threshold of 100 and 1000, instead of 10. The number of rounded values again orientates on the percentage of rounded values in the Allbus, so that in the case of rounding to the next threshold of 100 obviously less income values are rounded compared to the previous threshold of 10. Then in the third rounding model only a small number of income values are rounded to the next threshold of 1000. More precisely, for example in the underlying MNAR scenario we have in average 233.48 rounded up income values, 203.78 observations that are rounded off and 3062.74 income values that remain as they originally have been. Those circumstances, meaning the comparatively small number of rounded values and the compensation due to similar numbers of up and off rounded income values, probably cause the small deviations from the true values, as the resulting evaluation of the mean income estimation in the following tables (4.4) and (4.5) shows. For better comparison, finally a last rounding model is viewed that contains rounding to the next threshold of 1000, whereas the probability for rounding is identical to the constructed and evaluated rounding model, described in the measurement error subsection of chapter 2.2. Consequently, the same income values are rounded, but to the next threshold of 1000, not 10. This model deviates from the Allbus role model, but shows impacts of larger rounding intervals, when other conditions are hold constant. The estimation then obviously worsens, but is still tolerable, as the outcome of mean estimation, listed in table (4.6), makes clear.

	RCAR	RAR	RNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1493.370	1493.286	1493.292
$Bias(\hat{\mu}_{\text{income}})$	0.41914	0.33491	0.34061
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	0.00028	0.00022	0.00023
$Var(\hat{\mu}_{\text{income}})$	0.05134	0.04228	0.04174
$MSE(\hat{\mu}_{\text{income}})$	0.22702	0.15444	0.15776

Table 4.4: Expected value, variance, bias and MSE of mean estimation of income in alternative rounding model with threshold 100 in order to quantify measurement error

	RCAR	RAR	RNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1501.143	1495.046	1495.084
$Bias(\hat{\mu}_{\text{income}})$	8.192074	2.09501	2.13290
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	0.00549	0.00140	0.00143
$Var(\hat{\mu}_{\text{income}})$	7.24193	2.29290	2.43778
$MSE(\hat{\mu}_{\text{income}})$	74.35201	6.68197	6.98705

Table 4.5: Expected value, variance, bias and MSE of mean estimation of income in alternative rounding model with threshold 1000 in order to quantify measurement error

	RCAR	RAR	RNAR
μ_{income}	1492.951	1492.951	1492.951
$\mathbb{E}(\hat{\mu}_{\text{income}})$	1502.467	1505.378	1505.392
$Bias(\hat{\mu}_{\text{income}})$	9.51600	12.42701	12.44089
$Bias_{\text{rel}}(\hat{\mu}_{\text{income}})$	0.00637	0.00832	0.00833
$Var(\hat{\mu}_{\text{income}})$	6.10699	1.66200	1.72996
$MSE(\hat{\mu}_{\text{income}})$	96.66130	156.09254	156.50560

Table 4.6: Expected value, variance, bias and MSE of mean estimation of income in rounding model with threshold 1000 and increased rounding probability in order to quantify measurement error

As before, the results of RCAR are not directly comparable with those of RAR and RNAR, since the number of rounded values here clearly deviates. However, small deviations from the true income value still remain in all three scenarios, even though now more remarkable measurement errors, in form of wider rounding intervals, are constructed and applied compared to the rounding model evaluated in the context of the thesis. On the one hand this indicates that the impact of rounding error does not only depend on the choice of the rounding interval, but on other factors. Schneeweiss et al. (2010), for example, point out the influence of the underlying distribution of the unrounded data. Moreover, the outcomes make clear that even those modified measurement error models with larger rounding thresholds seem to be too overoptimistic, since in general not all responder do have the same rounding behavior. So choosing just one rounding threshold for all units that should have a measurement error is not realistic. Moreover, the people might also round to other thresholds like 50, 500 or certainly will not pay attention to correct rounding rules. Then the influence of rounding errors might probably be more noticeable, resulting in larger biases and MSEs. The same aspect is valid for the nonresponse error source. Here, certain variables are chosen to have an influence on the dropout, but in reality the situation is more complex. There is a mass of reasons to refuse an answer in a survey. Those causes in general are not the same for all nonresponder and moreover they are often not captured through variables which permits from modeling the dropout process.

Consequently, there are several facts concerning this simulation approach that might lead to overoptimistic results in situations of present survey errors, shown up through small biases and small MSEs. In addition, it shouldn't be forgotten that the different error sources here are regarded separately, which clearly is not the case in real life situations. There, all error sources of the TSE appear together which results in even worse parameter estimation.

But not only the results of the first part of the thesis shall be considered as overoptimistic, but also the performances of different missing data approaches will probably not be this good in real life situations due to the large amount of additional information about the dropout process that is used here in most applications. Since this fact has already been discussed in chapter 3.3 and will also be brought up in the final conclusions, it will not be gone into detail at this point.

5 Conclusion and Further Research

Finally summarizing the findings of the thesis leads to several conclusions. First of all, the concept of the TSE showed that survey methodology has to deal with several different error sources affecting statistics in different ways. Obviously, it is not possible to prevent from all those biasing effects, but the decomposition into different error components showed that focus should not only be set on sampling error, which is discussed the most in survey literature. Sampling error in this application did not bias mean estimation and hardly had an influence on regression coefficient estimation. In contrast to that, coverage error had a remarkable impact. So did nonresponse error in the cases of MAR and MNAR. Finally rounding error, which has been analyzed as a special case of measurement error, yields small deviations between the true parameter values and the corresponding estimates. With those findings concerning the TSE on the one hand gaps in research literature can be identified and on the other hand guidelines for users can be derived by emphasizing the important factors that can lead to biased results. However, there is a lack in routine measurements that could help the users to detect and deal with different error sources especially due to the fact that different error sources are linked and approaches for handling one error might influence the bias of other error components. As a result of these remaining tasks in the field of TSE, until now the TSE approach more or less remains a theoretical construct which clearly is a weakness of the approach.

As introduced previously, one major component of the TSE approach is the nonresponse error. This source of error is omnipresent, since on the one hand it arises in nearly every survey and on the other hand unlike the other components of the TSE it is directly visible namely through missing data. Thus, finding a way to handle this problem is a lasting challenge for analysts, leading to the second part of the thesis which deals exclusively with this error component and focuses on approaches to treat missing values. The theoretical introduction and especially the application of different missing data strategies reveals weaknesses, limitations and problems, respectively. Most of all it showed that despite the large amount of literature concerning this topic, much more research is necessary in this field, since until now there is no gold standard method which became prevalent. Instead it showed that for different nonresponse scenarios, other approaches performed best in terms of smallest bias and smallest MSE.

In the case of MCAR ad-hoc approaches, based on simply ignoring the missing data, are theoretically standing to reason and resulted in the best mean estimates com-

pared to all other applications. For MAR and MNAR ad-hoc methods lead to biased results, which is why other approaches should be chosen. Choosing right weights can give better estimates and also imputing values following special schemes can result in improved parameter estimation. Thus, for example regression imputation yields good mean estimates especially for the MAR scenario, where no bias is received.
In contrast to that, the estimation of regression coefficients with imputed data sets leads to dilution of statistical effects. In this regard, even multiple imputation could not improve the outcome. All in all, the performances of imputation satisfied the expectations only in certain cases. Maximum Likelihood Procedures ignoring the underlying dropout process, lead to unbiased mean estimation in the case of MCAR and MAR, but also in the MNAR scenario deviation from the true value is comparative small. This confirms findings of a previous study (cp. Schafer and Graham (2002)), stating that MAR methods might indeed also perform well in MNAR scenarios. A possible explanation might be the fact that dropout depends on the value of the dependent variable itself, but other variables which partly explain the dependent variable are observable. Drawing next conclusion about the resulting regression coefficient estimates of these methods gives that they differ only little from ad-hoc estimates. This result is equivalent to all other previously evaluated procedures. It shows that due to missing data, the dependence structure of the variables is slacken and consequently estimates of regression coefficients are not becoming much better than in ad-hoc approaches, where only observed values are used. While extremely significant or nonsignificant statistical effects remain visible, other effects are obscured. So, summarizing the results on regression coefficient estimation gives the following: With respect to the corresponding bias, ad-hoc methods for all different regression coefficients performed in the upper midfield compared to all the other approaches. Most often weighting, FIML, Heckman SM and PMM resulted in a smaller bias, while stochastic regression imputation has been the only imputation approach whose biases have been comparable. All the other imputation approaches resulted in a larger bias for the regression coefficients, independent on which nonresponse mechanism was present. However, for a certain regression coefficient not always the same approach performed best for the different missing mechanism scenarios. Regarding the corresponding MSEs, this is also the case, meaning that while for a certain regression coefficient a particular approach can result in the smallest MSE in the case of MCAR, another approach can perform better for MAR and MNAR. Here, no system is noticeable. However, regarding the MSE, in general ad-hoc methods, PMM and FIML result in the smallest outcomes and therefore performed best. Yet, also the MSEs of the weighting approach and the Heckman SM usually are very

close. In the field of imputation, stochastic regression imputation and the multiple imputation approaches resulted in noticeably smaller MSEs than random and mean imputation. These results showed that in the field of regression coefficient estimation, all presented methods did not bring much advantages compared to the ad-hoc approach. Therefore, future work in the field of missing data should concentrate on parameter estimation that describe dependencies.

Last, focus of the summary is further set on modern, modelbased nonresponse methods, which are based on explicit modeling of the nonresponse process, namely Heckman SM and PMM. While their results of regression coefficient estimation have already been embedded in the comparison before, focus is now set on their performance in the field of mean estimation. Both, the Heckman SM, as well as the PMM yield good mean estimates and their use is consequently recommended for this purpose. However, their performances strongly depend on prior knowledge regarding the dropout process which is an obvious disadvantage, since these information in general are not given. Moreover, for an easier application, more work has to be done with respect to the implementation of those models. Whereas there is already a package in **R** that is able to fit Heckman SMs, this is not the case for PMMs. For this reason, the application of the PMM in this thesis is limited, but further attention should be directed to this topic in later works. Nevertheless, those two modern approaches already have shown that there are ways, to receive good parameter estimates, even when data are not simply MAR. Yet, in the case of MCAR, the application of those modelbased approaches are obviously meaningless.

However, there is no gold standard way to diagnose the underlying missing data mechanism which is without a doubt one of the biggest problems in the field of nonresponse error. Research has already found out several application examples, where ignorability is or is not known to hold (cp. Schafer (1997)), but it should definitely be spend more effort on this topic in following works. This thesis illustrates which group of method is suitable for a certain nonresponse mechanism, but it is unprofitable, when the underlying dropout process is not known. Thus, in general assumptions are made, in order to be able to use some of these methods, even though those fundamental assumptions can neither be agreed with, nor be refused. This is the point, where sensitivity analysis comes in and performances are evaluated, when assumptions are not valid or made differently. In this thesis partial differentiation as a new approach to deal with missing data is explained, but due to reasons of the thesis' scope, it is not applied. Consequently, in this field future work is desirable, since it is a quite new approach and not much literature about the application is published.

So all in all, this thesis pointed out several crucial points of survey statistics by simulating a data set that is comparable to survey data, but it also presented approaches to solve these problems, while embedding all this in a field of several open questions and further need for research.

References

Arbuckle, J. L., G. A. Marcoulides, and R. E. Schumacker (1996). Full information estimation in the presence of incomplete data. *Advanced structural equation modeling: Issues and techniques 243*, 277.

Baker, S. G. (1992). A simple method for computing the observed information matrix when using the em algorithm with categorical data. *Journal of Computational and Graphical Statistics 1*(1), 63–76.

Biemer, P. P. (2010a). Overview of design issues: Total survey error. *Handbook of survey research 2*, 27–57.

Biemer, P. P. (2010b). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly 74*(5), 817–848.

Biemer, P. P. and L. E. Lyberg (2003). *Introduction to survey quality*, Volume 335. John Wiley & Sons.

Buzas, J., L. Stefanski, and T. Tosteson (2005). *Measurement Error*, pp. 1241–1282. Springer VS.

Casella, G. and R. L. Berger (1990). Statistical Inference. Duxbury Press. *Belmont, Calif*.

Chesher, A. (1991). The effect of measurement error. *Biometrika 78*(3), 451–462.

Collins, L. M., J. L. Schafer, and C.-M. Kam (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological methods 6*(4), 330.

Deming, W. E. (1944). On errors in surveys. *American Sociological Review 9*(4), 359–369.

Demirtas, H. and J. Schafer (2003). On the performance of random-coefficient pattern-mixture models for non-ignorable drop-out. *Statistics in Medicine 22*(16), 475–492.

Dempster, A., N. Laird, and D. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society 39*(1), 1–38.

Dempster, A. and D. Rubin (1983). Overview. In W. Madow, I. Olkin, and D. Rubin (Eds.), *Incomplete Data in Sample Surveys (Vol. 2): Theory and Bibliography*, pp. 3–10. New York: Academic Press.

Dong, Y. and C.-Y. J. Peng (2013). Principled missing data methods for researchers. *SpringerPlus 2*(1), 1.

Enders, C. K. (2001a). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological methods 6*(4), 352.

Enders, C. K. (2001b). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling 8*(1), 128–141.

European Health Examination Survey (2016). Characteristics of participants and non-participants. http://www.ehes.info/nopahes/literature.html#characteristics. [9 November 2016].

Fahrmeir, L., R. Künstler, I. Pigeot, and G. Tutz (2007). *Statistik: Der Weg zur Datenanalyse*. Springer-Verlag.

Faulbaum, F. (2014). Total survey error. In N. Baur and B. Jörg (Eds.), *Handbuch Methoden der empirischen Sozialforschung*, pp. 439 – 453. Springer VS.

Finkbeiner, C. (1979). Estimation for the multiple factor model when data are missing. *Psychometrika 44*(4), 409–420.

Fox, J. (2006). Teacher's corner: Structural equation modeling with the sem package in r. *Structural equation modeling 13*(3), 465–486.

Fricker, S. and R. Tourangeau (2010). Examining the relationship between nonresponse propensity and data quality in two national household surveys. *Public Opinion Quarterly 74*(5), 934–955.

GESIS (2016). Allbus. http://www.gesis.org/en/allbus/allbus-home/. [18 May 2016].

Graham, J. W. (2012). *Missing data: Analysis and design*. Springer Science & Business Media.

Groves, R. M. (1998). *Survey Errors and Survey Costs*. Wiley Series in Probability and Statistics. Wiley.

Groves, R. M. and L. Lyberg (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly 74*(5), 849–879.

Guo, S. and M. W. Fraser (2014). *Propensity Score Analysis: Statistical Methods and Applications*. Sage Publications.

Hanisch, J. (2005). Rounded responses to income questions. *Allgemeines Statistisches Archiv 89*(1), 39–48.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economics and Social Measurement 5*(4), 475–492.

Heitjan, D. F. and D. B. Rubin (1991). Ignorability and coarse data. *The Annals of Statistics 19*(4), 2244–2253.

Kauermann, G. and H. Küchenhoff (2010). *Stichproben: Methoden und praktische Umsetzung mit R.* Springer-Lehrbuch. Springer Berlin Heidelberg.

Kreuter, F. (2013). Facing the nonresponse challenge. *The Annals of the American Academy of Political and Social Science 645*(1), 23–35.

Kreuter, F., G. Müller, and M. Trappmann (2010). Nonresponse and measurement error in employment research: Making use of administrative data. *Public Opinion Quarterly 74*(5), 880–906.

Kreuter, F., T. Yan, and R. Tourangeau (2008). Good item or bad- can latent class analysis tell?: The utility of latent class analysis for the evaluation of survey questions. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 171*(3), 723–738.

Lehmann, E. L. and G. Casella (2006). *Theory of point estimation.* Springer Science & Business Media.

Little, R. J. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association 88*(421), 125–134.

Little, R. J. (1994). A class of pattern-mixture models for normal incomplete data. *Biometrika 81*(3), 471–483.

Little, R. J. A. (1982). Models for nonresponse in sample surveys. *American Statistical Association 77*(378), 237–250.

Little, R. J. A. and D. B. Rubin (1987). *Statistical analysis with missing data.* John Wiley & Sons.

Little, R. J. A. and D. B. Rubin (2002). *Statistical analysis with missing data* (2 ed.). Wiley-Interscience.

Manski, C. F. (2003). *Partial identification of probability distributions*. Springer Science & Business Media.

Manski, C. F. (2016). Credible interval estimates for official statistics with survey nonresponse. *Journal of Econometrics 191*(2), 293–301.

Marshall, A., D. G. Altman, R. L. Holder, and P. Royston (2009). Combining estimates of interest in prognostic modelling studies after multiple imputation: Current practice and guidelines. *BMC Medical Research Methodology 9*(57).

Neyman, J. (1934). On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society 97*(4), 558–625.

Novo, A. A. and J. L. Schafer (2013). *norm: Analysis of multivariate normal datasets with missing values*. R package version 1.0-9.5.

O'Brien, E. M., M. C. Black, L. R. Carley-Baxter, and T. R. Simon (2006). Sensitive topics, survey nonresponse, and considerations for interviewer training. *American journal of preventive medicine 31*(5), 419–426.

Pforr, K., M. Blohm, A. G. Blom, B. Erdel, B. Felderer, M. Fräßdorf, K. Hajek, S. Helmschrott, C. Kleinert, A. Koch, et al. (2015). Are incentive effects on response rates and nonresponse bias in large-scale, face-to-face surveys generalizable to Germany? Evidence from ten experiments. *Public Opinion Quarterly 79*(3), 740–768.

Plass, J., T. Augustin, M. E. Cattaneo, and G. Schollmeyer (2015). Statistical modelling under epistemic data imprecision: Some results on estimating multinomial distributions and logistic regression for coarse categorical data. ISIPTA.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raghunathan, T. E. (2004). What do we do with missing data? Some options for analysis of incomplete data. *Annu. Rev. Public Health 25*, 99–117.

Reiter, J. P. and T. E. Raghunathan (2007). The multiple adaptations of multiple imputation. *Journal of the American Statistical Association 102*(480), 1462–1471.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software 48*(2), 1–36.

Royall, R. (1970). On finite population sampling theory under certain linear regression models. *Biometrika 57*(2), 377–387.

Rubin, D. B. (1976). Inference and missing data. *Biometrika 63*(3), 581–592.

Rubin Donald, B. (1987). Multiple imputation for nonresponse in surveys.

Sakshaug, J. W., T. Yan, and R. Tourangeau (2010). Nonresponse error, measurement error, and mode of data collection: Tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly 74*(5), 907–933.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data.* CRC press.

Schafer, J. L. and J. W. Graham (2002). Missing data: Our view of the state of the art. *Psychological methods 7*(2), 147.

Schneeweiss, H. and T. Augustin (2006). Some recent advances in measurement error models and methods. *Allgemeines Statistisches Archiv 90*(1), 183–197.

Schneeweiss, H., J. Komlos, and A. Ahmad (2010). Symmetric and asymmetric rounding: A review and some new results. *AStA Advances in Statistical Analysis 94*(3), 247–271.

Schneeweiß, H. and H.-J. Mittag (1986). *Lineare Modelle mit fehlerbehafteten Daten.* Physica-Verlag Heidelberg-Wien.

Smith, G. (2008). Does gender influence online survey participation?: A record-linkage analysis of university faculty online survey response behavior. *ERIC Document Reproduction Service No. ED 501717*.

Spieß, M. (2008). *Missing-Data Techniken : Analyse von Daten mit fehlenden Werten*, Volume 1. LIT.

Stocker, H. (2016). Lecture Script: Zensierte (censored) und gestutzte (truncated) abhängige Variablen, Sample Selection. Available from `https://www.uibk.ac.at/econometrics/einf/22.pdf` [16 October 2016].

Thijs, H., G. Molenberghs, B. Michiels, G. Verbeke, and D. Curran (2002). Strategies to fit pattern-mixture models. *Biostatistics 3*(2), 245–265.

Toomet, O. and A. Henningsen (2008). Sample selection models in R: Package sampleSelection. *Journal of Statistical Software 27*(7).

Toutenburg, H., C. Heumann, and T. Nittner (2004). Statistische Methoden bei unvollständigen Daten. Sfb 380 discussion paper, Sonderforschungsbereich 386, Ludwig-Maximilians-Universität München, Germany. Available at http://epub.ub.uni-muenchen.de.

Van Buuren, S. and K. Groothuis-Oudshoorn (2011). mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software 45*(3), 1–67.

Van Teijlingen, E. and V. Hundley (2002). The importance of pilot studies. *Nursing Standard 16*(40), 33–36.

Weisberg, H. F. (2005). *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. The University of Chicago Press.

West, B. T. and K. Olson (2010). How much of interviewer variance is really nonresponse error variance? *Public Opinion Quarterly 74*(5), 1004–1026.

Westland, J. C. (2015). A brief history of structural equation models. In *Structural Equation Models*, pp. 9–22. Springer.

Windzio, M. (2013). *Abgeschnittene, zensierte und selektive Daten: Tobit-Regression und Heckman-Modell*, Chapter 10, pp. 255 –275. Springer VS.

Winker, P. (2010). *Empirische Wirtschaftsforschung und Ökonometrie*. Springer-Lehrbuch. Springer.

A Mathematical Derivations

Reformulating definitions of the MSE:

$$
\begin{aligned}
MSE(\hat{\theta}) &= \mathbb{E}(\hat{\theta}-\theta)^2 \\
&= \mathbb{E}[(\hat{\theta}-\mathbb{E}(\hat{\theta})+\mathbb{E}(\hat{\theta})-\theta)^2] \\
&= \mathbb{E}[(\hat{\theta}-\mathbb{E}(\hat{\theta}))^2+2((\hat{\theta}-\mathbb{E}(\hat{\theta}))(\mathbb{E}(\hat{\theta})-\theta))+(E(\hat{\theta})-\theta)^2] \\
&= \mathbb{E}[(\hat{\theta}-\mathbb{E}(\hat{\theta}))^2]+2\mathbb{E}[(\hat{\theta}-\mathbb{E}(\hat{\theta}))(\mathbb{E}(\hat{\theta})-\theta)]+\mathbb{E}[(\mathbb{E}(\hat{\theta})-\theta)^2] \\
&= \mathbb{E}[(\hat{\theta}-\mathbb{E}(\hat{\theta}))^2]+2(\mathbb{E}(\hat{\theta})-\theta)(\mathbb{E}(\hat{\theta})-\mathbb{E}[\hat{\theta}])+\mathbb{E}[(\mathbb{E}(\hat{\theta})-\theta)^2] \\
&= \mathbb{E}[(\hat{\theta}-\mathbb{E}(\hat{\theta}))^2]+2(\mathbb{E}(\hat{\theta})-\theta)\mathbb{E}(\hat{\theta}-\mathbb{E}(\hat{\theta}))+\mathbb{E}[(\mathbb{E}(\hat{\theta})-\theta)^2] \\
&= \mathbb{E}[(\hat{\theta}-\mathbb{E}(\hat{\theta}))^2]+\mathbb{E}[(\mathbb{E}(\hat{\theta})-\theta)^2] \\
&= Bias(\hat{\theta})^2+Var(\hat{\theta})
\end{aligned}
$$

Variance underestimation in the presence of mean imputation:
Let X_j be the variable of interest with observations x_{ij}. x_{ij} is missing for $i = 1, ..., m$ and observed for $i = (m+1), ..., n$. The corresponding sampling variance is calculated by

$$\sigma^2 = \frac{1}{n-m-1}\sum_{i=m+1}^{n}(x_{ij}-\bar{x}_j)^2 \quad \text{with} \quad \bar{x}_j = \frac{1}{n-m}\sum_{i=m+1}^{n}x_{ij}$$

Applying unconditional mean imputation for all m missing values of X_j gives $x_{ij} = \bar{x}_j$ for $i = 1, ..., m$. Then the corresponding sampling variance is calculated by

$$
\begin{aligned}
\sigma^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)^2 \\
&= \frac{1}{n-1}\left(\sum_{i=1}^{m}(x_{ij}-\bar{x}_j)^2+\sum_{i=m+1}^{n}(x_{ij}-\bar{x}_j)^2\right) \\
&= \frac{1}{n-1}\left(\sum_{i=1}^{m}(\bar{x}_j-\bar{x}_j)^2+\sum_{i=m+1}^{n}(x_{ij}-\bar{x}_j)^2\right) \\
&= \frac{1}{n-1}\sum_{i=m+1}^{n}(x_{ij}-\bar{x}_j)^2
\end{aligned}
$$

Then σ^2 in the case of conditional mean imputation is underestimated by $\dfrac{n-m-1}{n-1}$.

Derivation of casewise likelihood function in context of FIML estimation:
Let X be a k-dimensional random variable, being normally distributed with expected value vector μ and covariance matrix Σ. The corresponding density function then is given by

$$f(x) = \frac{1}{\sqrt{(2\pi)^k|\Sigma|}} \exp -\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^T.$$

Taking (-2) times the logarithm of the probability density function gives

$$\begin{aligned}
(-2)\ln(f(x)) =& -2\ln\left[\frac{1}{\sqrt{(2\pi)^k|\Sigma|}} \exp -\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^T\right] \\
=& -2\ln((2\pi)^{-\frac{k}{2}}) - 2\ln(|\Sigma|)^{-\frac{1}{2}} - 2\ln\left[\exp -\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^T\right] \\
=& k\ln(2\pi) + \ln(|\Sigma|) + (x-\mu)\Sigma^{-1}(x-\mu)^T
\end{aligned}$$

Derivating basic formula of EM algorithm:
Let the starting point be the rearranged decomposition of the complete-data log-likelihood:

$$\ell(\theta \mid X^{\text{obs}}) = \ell(\theta \mid X) - \ln f(X^{\text{mis}} \mid X^{\text{obs}}, \theta)$$

Taking expectation with respect to X^{mis} given X^{obs} and the actual parameter estimate $\theta^{(t)}$ on both sides of the equation results in

$$\begin{aligned}
\int \ell(\theta \mid X^{\text{obs}}) \cdot f(X^{\text{mis}} \mid X^{\text{obs}}, \theta^{(t)}) dX^{\text{mis}} =& \underbrace{\int \ell(\theta \mid X) \cdot f(X^{\text{mis}} \mid X^{\text{obs}}, \theta^{(t)}) dX^{\text{mis}}}_{Q(\theta,\theta^{(t)})} - \\
&- \underbrace{\int \ln f(X^{\text{mis}} \mid X^{\text{obs}}, \theta) \cdot f(X^{\text{mis}} \mid X^{\text{obs}}, \theta^{(t)}) dX^{\text{mis}}}_{H(\theta,\theta^{(t)})}
\end{aligned}$$

$$\underbrace{\int \ell(\theta \mid X^{\text{obs}}) \cdot f(X^{\text{mis}} \mid X^{\text{obs}}, \theta^{(t)}) dX^{\text{mis}}} = Q(\theta,\theta^{(t)}) - H(\theta,\theta^{(t)})$$

$$\begin{aligned}
&= \ell(\theta \mid X^{\text{obs}}) \cdot \textstyle\int f(X^{\text{mis}} \mid X^{\text{obs}}, \theta^{(t)}) dX^{\text{mis}} \\
&= \ell(\theta \mid X^{\text{obs}}) \cdot 1 \\
&= \ell(\theta \mid X^{\text{obs}})
\end{aligned}$$

Derivating the inverse Mills Ratio:

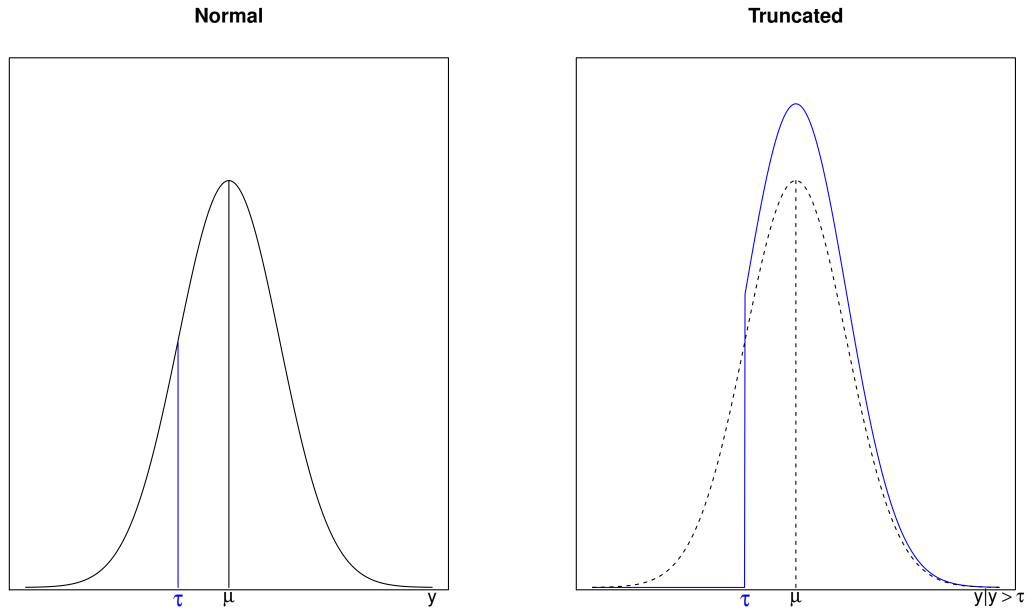

Figure A.1: Normal distributed and truncated variable (graphical representation based on Stocker (2016))

Let $y \sim N(\mu, \sigma^2)$, then

$f(y \mid \mu, \sigma) = \frac{1}{\sigma}\theta\left(\frac{\mu - y}{\sigma}\right)$, $\mathbb{P}(y \leq \tau) = \Theta\left(\frac{y-\mu}{\sigma}\right)$ and $\mathbb{P}(y > \tau) = 1 - \Theta\left(\frac{y-\mu}{\sigma}\right)$

(cp. left plot of figure (A.1))

Let $y \mid y > \tau$, then area under curve still has to be 1 and consequently has to be adapted by $f(y \mid y > \tau, \mu, \sigma) = \dfrac{f(y \mid \mu, \sigma)}{\mathbb{P}(y > \tau)} = \dfrac{\frac{1}{\sigma}\theta\left(\frac{y-\mu}{\sigma}\right)}{1 - \Theta\left(\frac{y-\mu}{\sigma}\right)}$

Taking expectations gives:

$$\mathbb{E}(y \mid y > \tau, \mu, \sigma) = \mu + \sigma \underbrace{\frac{\theta\left(\frac{\mu-\tau}{\sigma}\right)}{\Theta\left(\frac{\mu-\tau}{\sigma}\right)}}_{\text{inverse Mills Ratio } \lambda}$$

B Figures

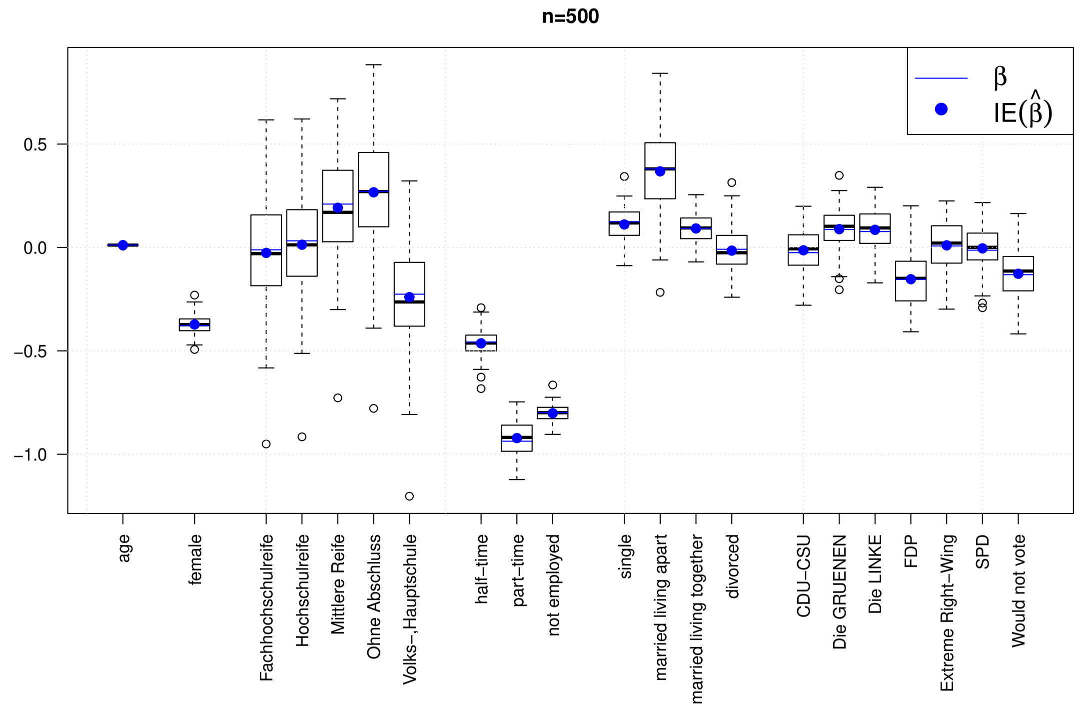

Figure B.1: Boxplots of regression coefficient estimates for sample size 500

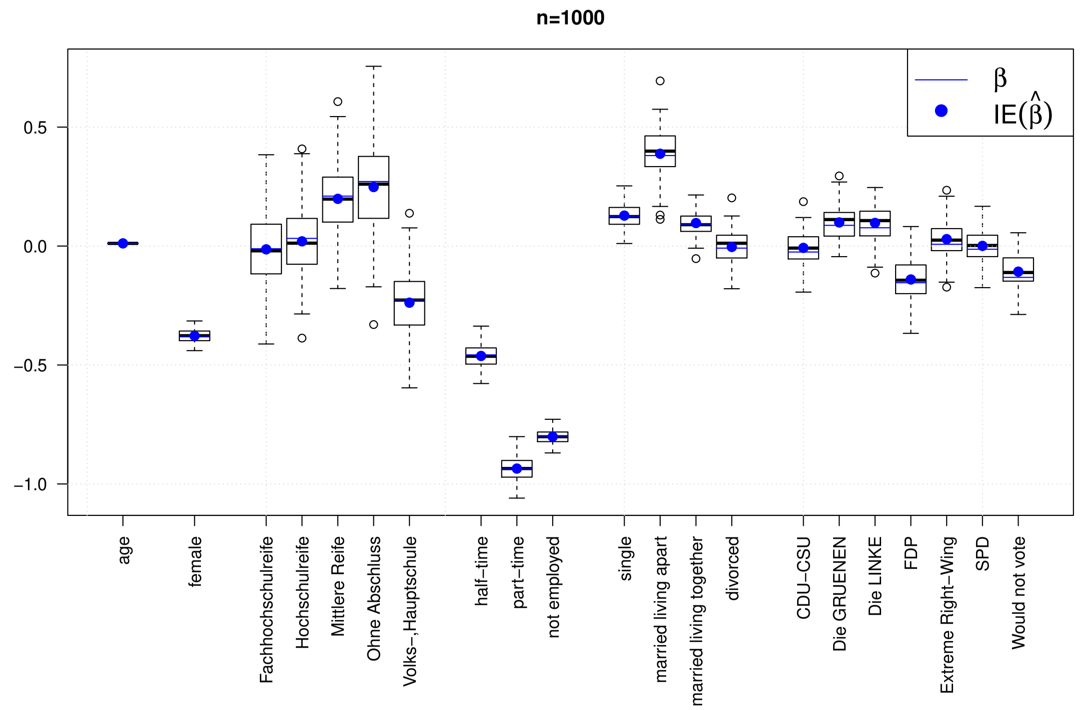

Figure B.2: Boxplots of regression coefficient estimates for sample size 1000

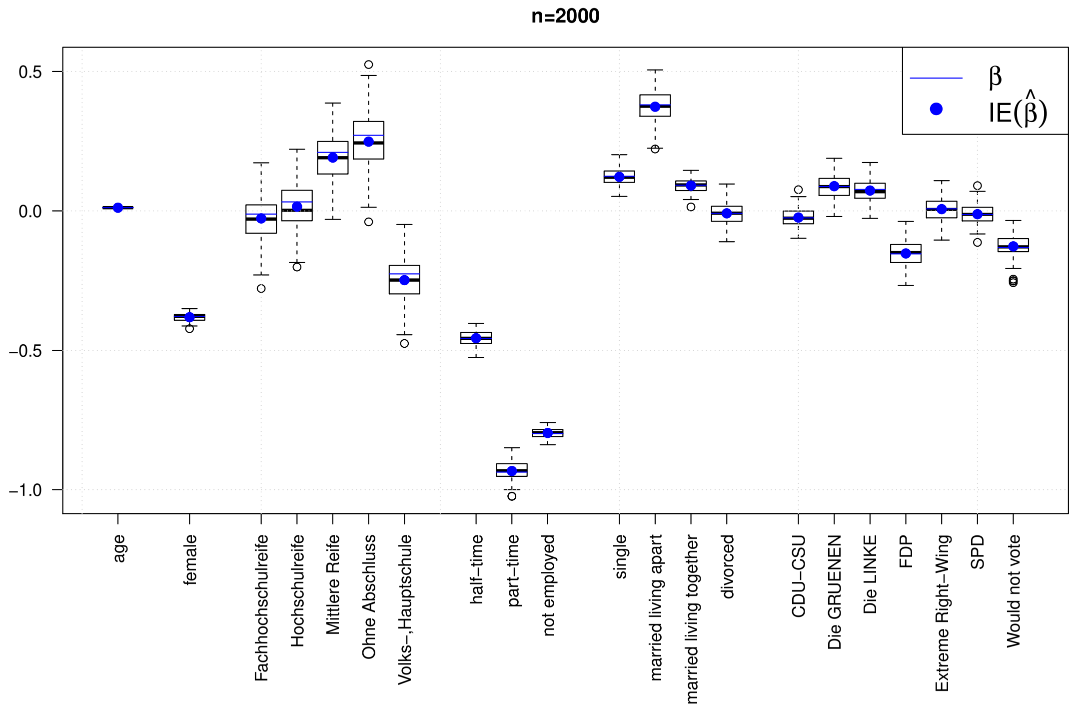

Figure B.3: Boxplots of regression coefficient estimates for sample size 2000

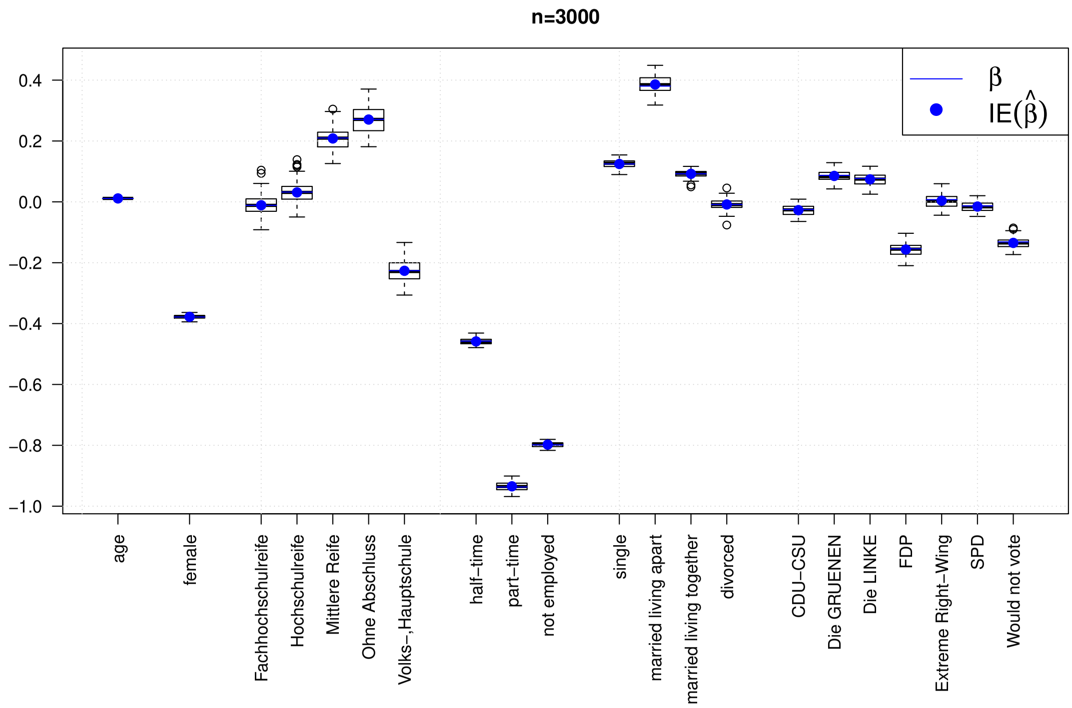

Figure B.4: Boxplots of regression coefficient estimates for sample size 3000

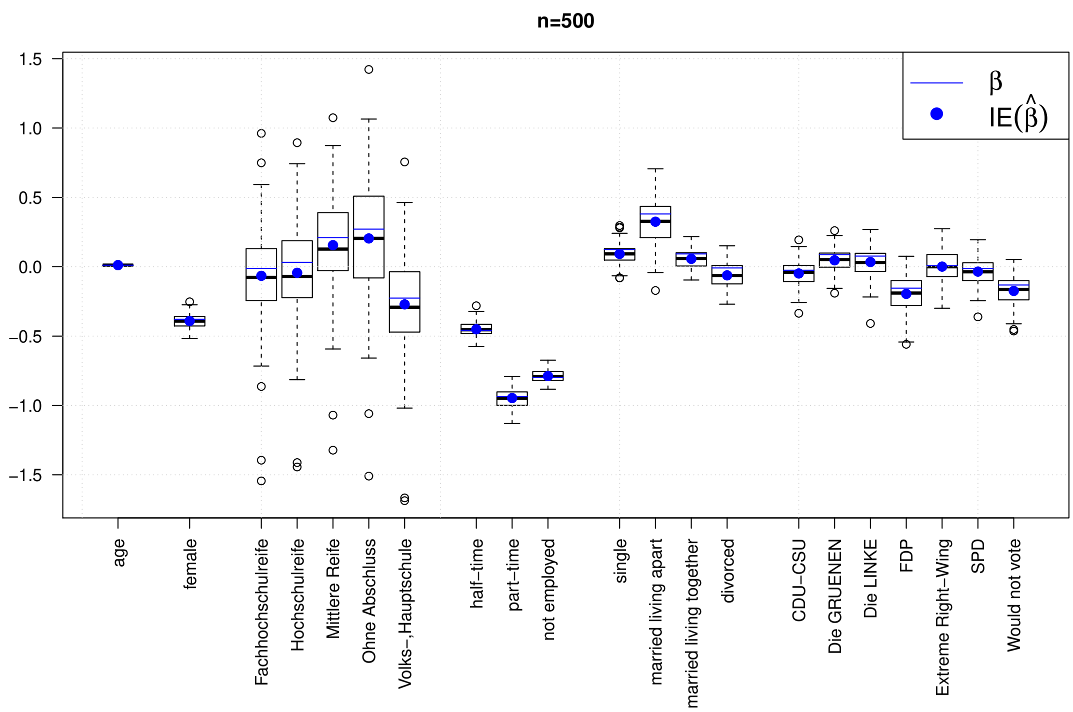

Figure B.5: Boxplots of regression coefficient estimates for sample size 500 containing coverage error

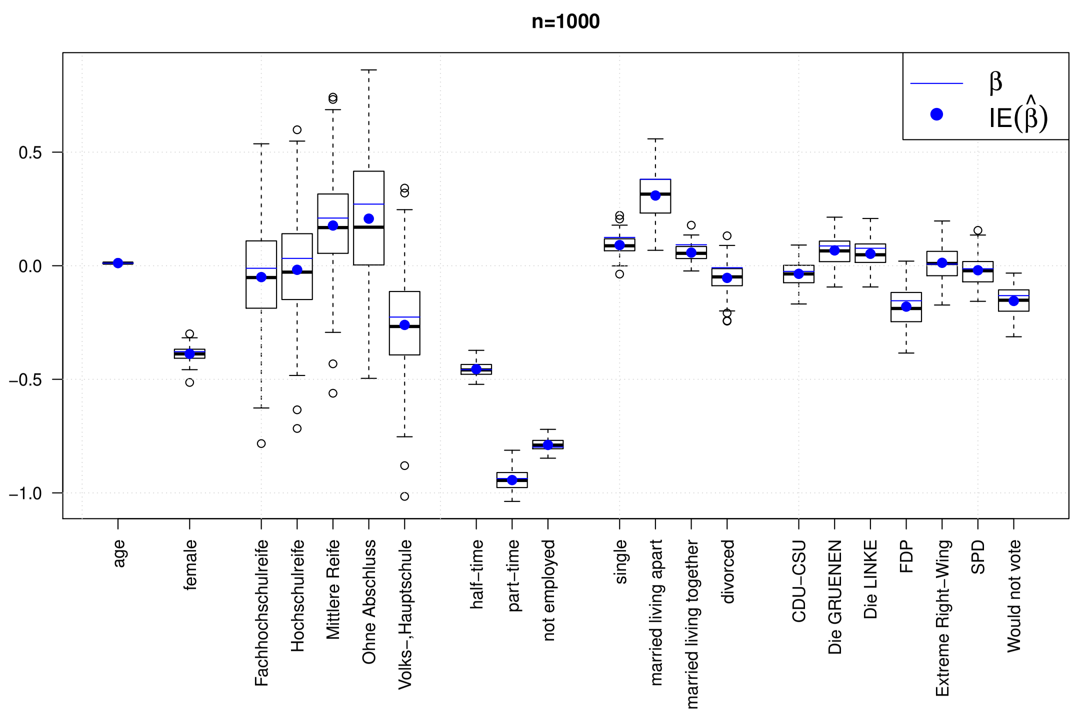

Figure B.6: Boxplots of regression coefficient estimates for sample size 1000 containing coverage error

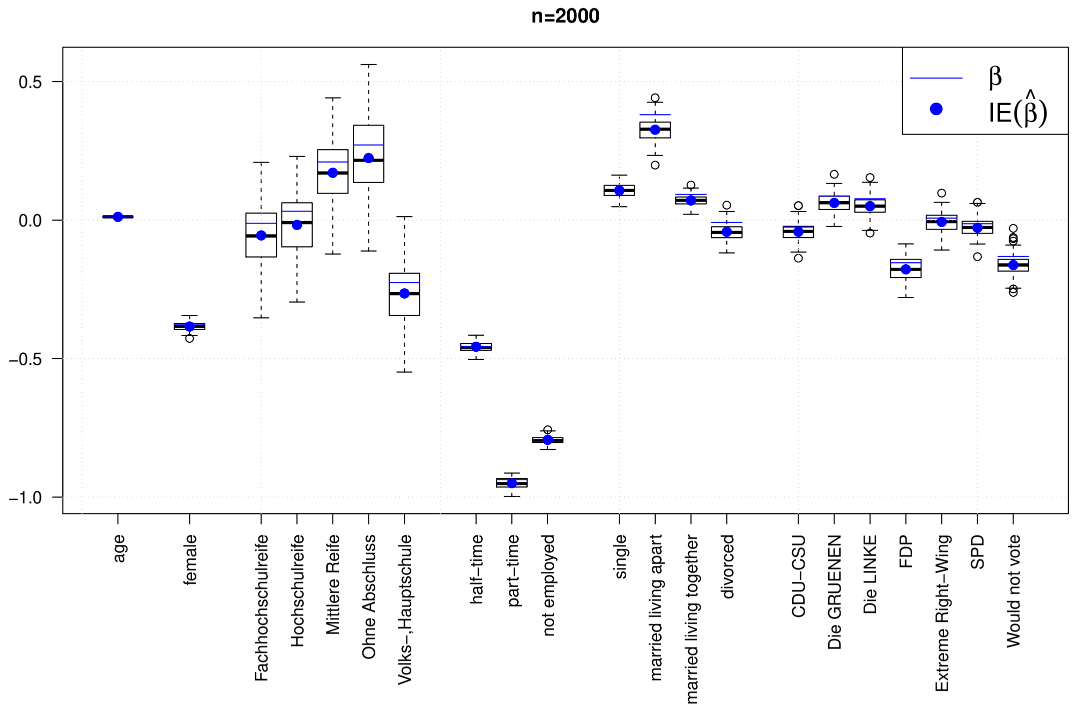

Figure B.7: Boxplots of regression coefficient estimates for sample size 2000 containing coverage error

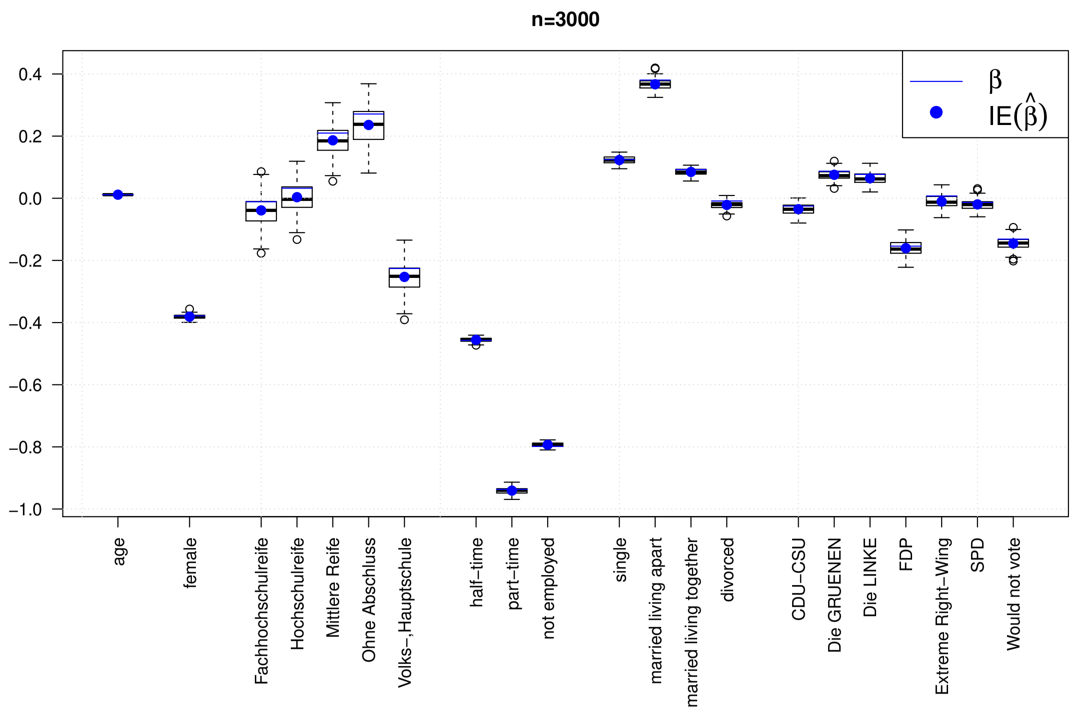

Figure B.8: Boxplots of regression coefficient estimates for sample size 3000 containing coverage error

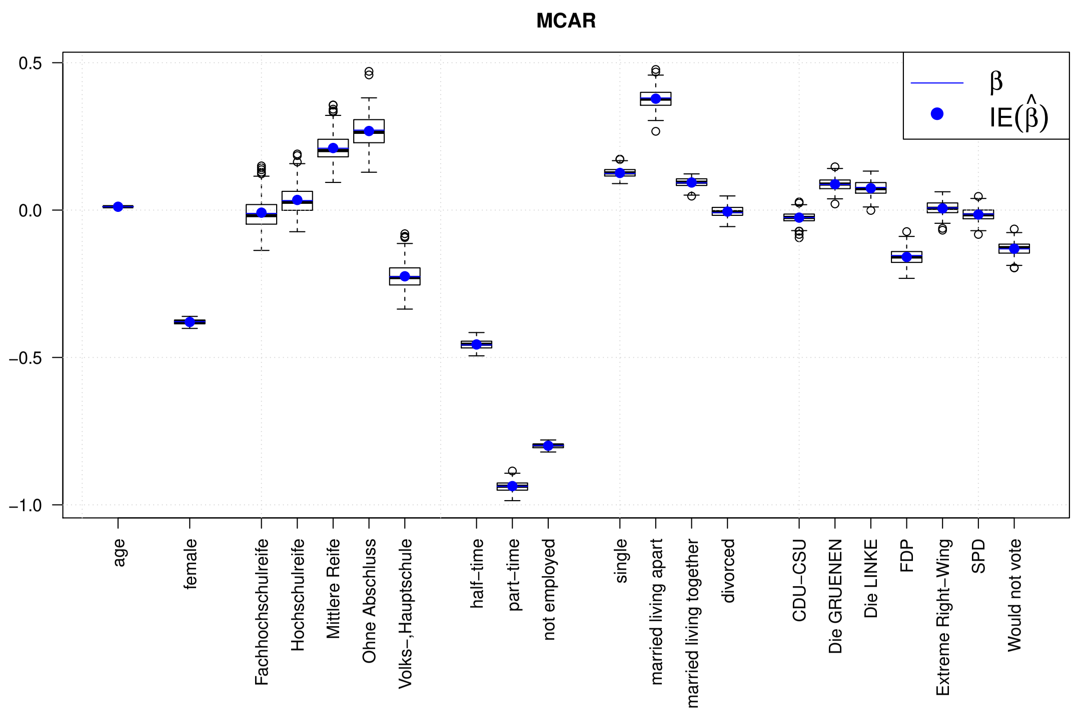

Figure B.9: Boxplots of regression coefficient estimates for data containing nonresponse error following MCAR

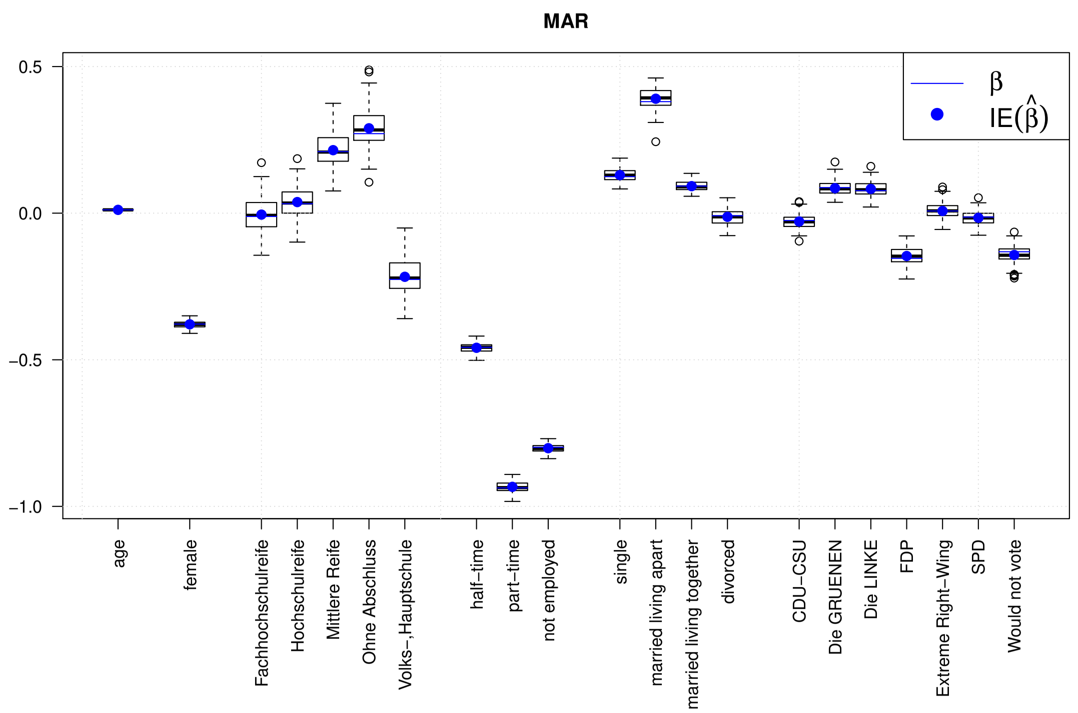

Figure B.10: Boxplots of regression coefficient estimates for data containing nonresponse error following MAR

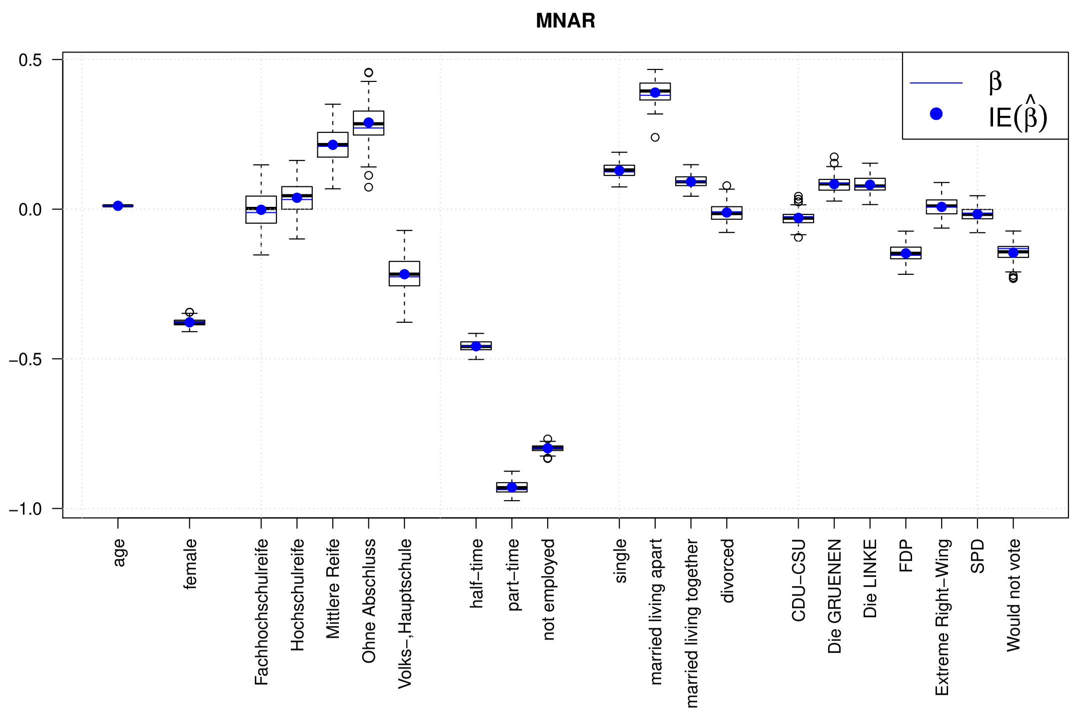

Figure B.11: Boxplots of regression coefficient estimates for data containing nonresponse error following MNAR

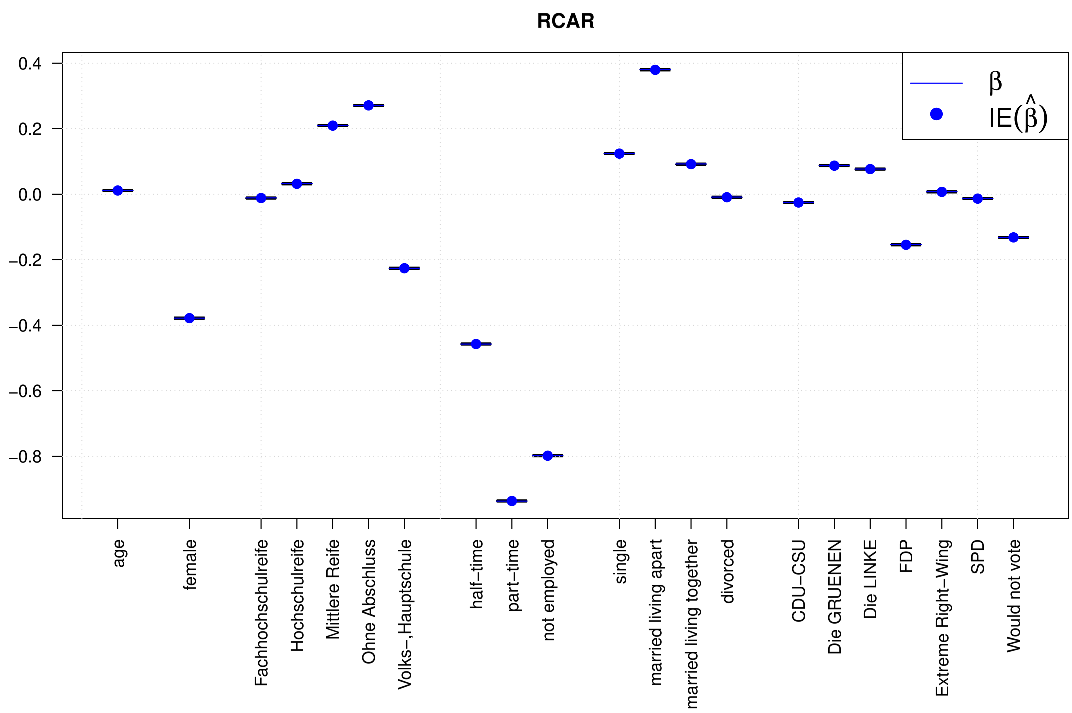

Figure B.12: Boxplots of regression coefficient estimates for data containing measurement error following MCAR

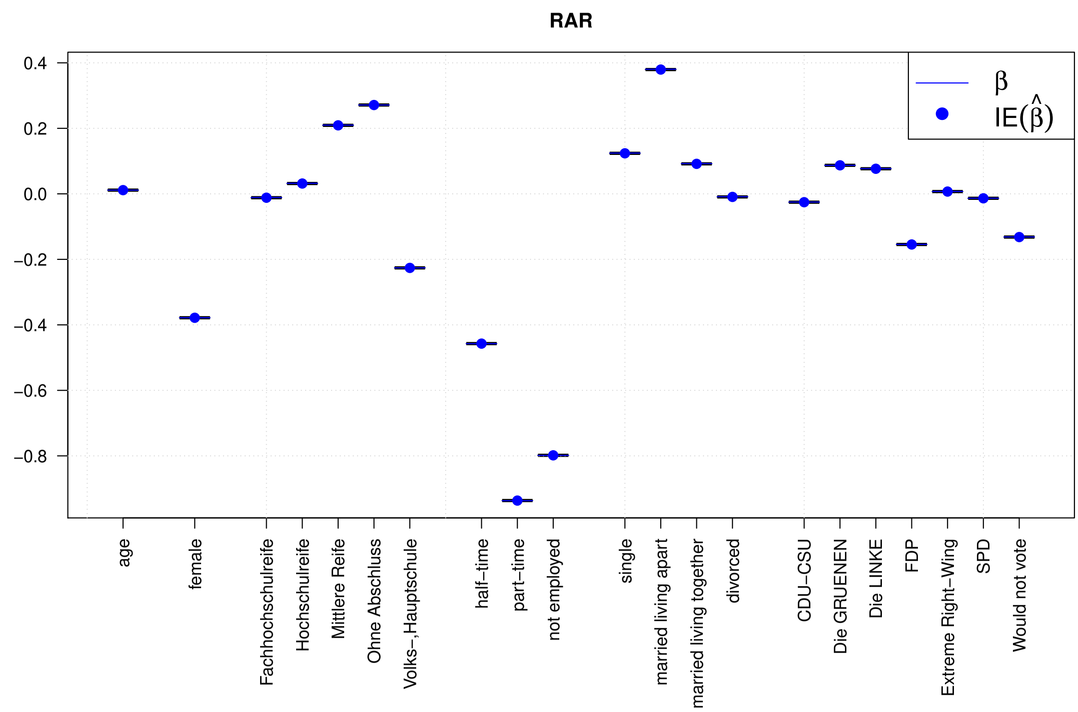

Figure B.13: Boxplots of regression coefficient estimates for data containing measurement error following MAR

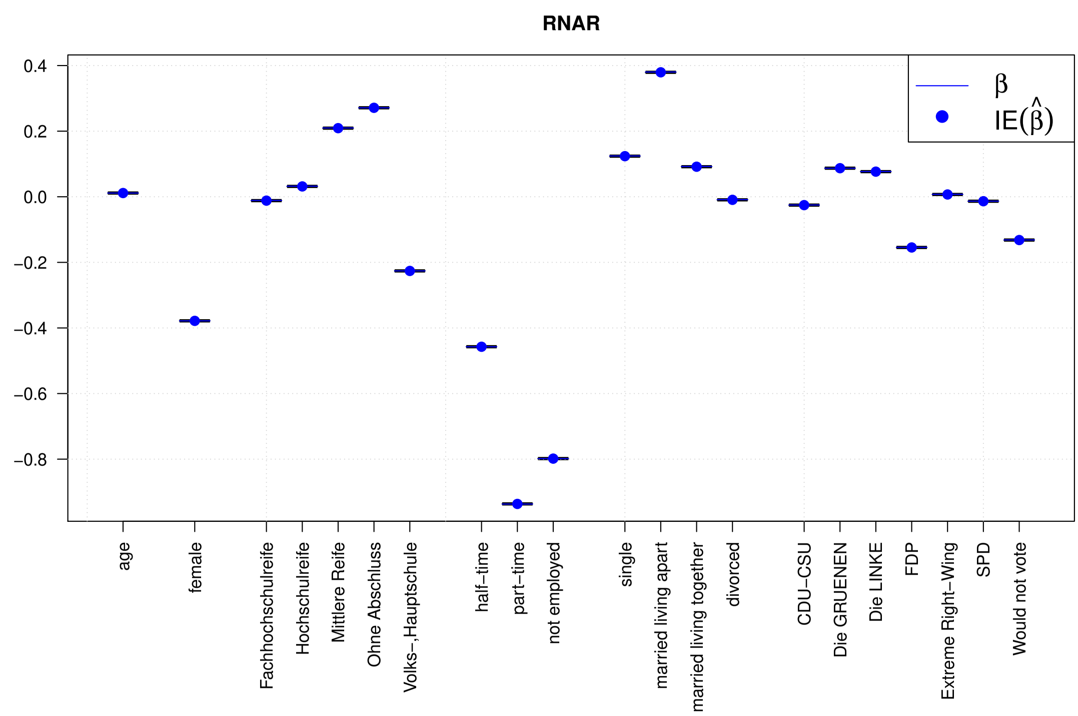

Figure B.14: Boxplots of regression coefficient estimates for data containing measurement error following MNAR

C Tables

	age	income
Min.	18.00	132.00
1st Qu.	38.00	698.80
Median	50.00	1155.00
Mean	50.15	1493.00
3rd Qu.	61.00	1934.00
Max.	89.00	10904.00

Table C.1: Summaries of metric variables in the simulated data set

		n	%
gender	male	1746	0.499
	female	1754	0.501
education [1]	no graduation	81	0.023
	volks-, hauptschule	992	0.283
	mittlere reife	1114	0.318
	fachhochschulreife	257	0.073
	hochschulreife	1019	0.291
	other graduation	37	0.011
professional activity	full-time	1483	0.424
	half-time	410	0.117
	part-time	243	0.069
	not employed	1364	0.390
family status	married living together	1780	0.509
	married living apart	78	0.022
	widowed	257	0.073
	divorced	317	0.091
	single	1068	0.305
election intention	CDU-CSU	1013	0.289
	SPD	773	0.221
	die gruenen	479	0.137
	die linke	359	0.103
	extreme right-wing	238	0.068
	FDP	161	0.046
	Other Party	127	0.036
	would not vote	350	0.100
willingness to take part in the interview	very easy	1116	0.319
	rather easy	1524	0.435
	rather difficult	674	0.193
	very difficult	186	0.053

Table C.2: Absolute and relative frequencies of dichotomous and categorical variables in the simulated data set

[1] Due to the specific school system in Germany the different Graduations will not be translated

Categories[2]	β Coefficients	Standard Error	P-Value
intercept	7.0070608	0.1076221	0.0000000000
age	0.0113721	0.0007065	0.0000000000
female	−0.3783041	0.0180399	0.0000000000
no graduation	0.2712025	0.0984070	0.0058831082
volks-, hauptschule	−0.2259655	0.0830450	0.0065408724
mittlere reife	0.2098522	0.0830075	0.0115117133
fachhochschulreife	−0.0112989	0.0873593	0.8970973124
hochschulreife	0.0322277	0.0832688	0.6987555776
half-time	−0.4575121	0.0291291	0.0000000000
part-time	−0.9369052	0.0345564	0.0000000000
not employed	−0.7984611	0.0194205	0.0000000000
married living together	0.0923461	0.0300279	0.0021189162
married living apart	0.3803864	0.0622169	0.0000000000
widowed	−0.0087343	0.0431919	0.8397563199
single	0.1244900	0.0339140	0.0002454705
CDU-CSU	−0.0250408	0.0462772	0.5884698278
SPD	−0.0133277	0.0470551	0.7770122356
die gruenen	0.0870818	0.0491068	0.0762638735
die linke	0.0768770	0.0507520	0.1299242680
extreme right-wing	0.0074006	0.0540235	0.8910480181
FDP	−0.1541756	0.0583351	0.0082560314
would not vote	−0.1315889	0.0512836	0.0103321856

Table C.3: Regression coefficients of the independent variables in the simulated data set

	age	income
Min.	18.00	37.00
1st Qu.	35.00	800.80
Median	50.00	1300.00
Mean	49.44	1545.00
3rd Qu.	62.00	2000.00
Max.	91.00	60000.00

Table C.4: Summaries of the metric variables in the Allbus 2014 data set

[2]reference categories: male, other graduation, full-time, divorced, Other Party

		n	%
gender	male	1762	0.508
	female	1709	0.492
education [3]	no graduation	64	0.018
	volks-, hauptschule	974	0.281
	mittlere reife	1144	0.330
	fachhochschulreife	270	0.078
	hochschulreife	975	0.281
	other graduation	39	0.011
professional activity	full-time	1571	0.453
	half-time	350	0.101
	part-time	208	0.060
	not employed	1339	0.386
family status	married living together	1934	0.559
	married living apart	58	0.017
	widowed	224	0.065
	divorced	272	0.079
	single	973	0.281
election intention	CDU-CSU	862	0.300
	SPD	663	0.230
	die gruenen	386	0.134
	die linke	281	0.098
	extreme right-wing	190	0.066
	FDP	119	0.041
	Other Party	99	0.034
	would not vote	278	0.097
willingness to take part in the interview	very easy	1094	0.315
	rather easy	1541	0.444
	rather difficult	654	0.188
	very difficult	182	0.052

Table C.5: Absolute and relative frequencies of dichotomous and categorical variables in the Allbus 2014 data set

	$n = 500$	$n = 1000$	$n = 2000$	$n = 3000$
β_0	7.0070608279	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0188242414	7.0020441576	7.0242113040	7.0106541436
$Bias(\hat{\beta}_0)$	0.0117634135	-0.0050166703	0.0171504762	0.0035933157
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0016787943	-0.0007159450	0.0024475992	0.0005128135
$Var(\hat{\beta}_0)$	0.0824542464	0.0290291949	0.0095546657	0.0019973929
$MSE(\hat{\beta}_0)$	0.0825926243	0.0290543619	0.0098488046	0.0020103048
β_{age}	0.0113720951	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0113665651	0.0113141737	0.0114217487	0.0113556442
$Bias(\hat{\beta}_{\text{age}})$	-0.0000055300	-0.0000579215	0.0000496535	-0.0000164509
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.0004862778	-0.0050932967	0.0043662621	-0.0014466016
$Var(\hat{\beta}_{\text{age}})$	0.0000025652	0.0000015655	0.0000004694	0.0000000998
$MSE(\hat{\beta}_{\text{age}})$	0.0000025652	0.0000015689	0.0000004719	0.0000001000
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793	-0.3783040793

[3]Due to the specific school system in Germany the different Graduations will not be translated

$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.3725534172	-0.3774336202	-0.3813361226	-0.3777145687
$Bias(\hat{\beta}_{\text{female}})$	0.0057506620	0.0008704590	-0.0030320434	0.0005895106
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	0.0152011632	0.0023009507	-0.0080148312	0.0015582983
$Var(\hat{\beta}_{\text{female}})$	0.0021166700	0.0007933948	0.0002371454	0.0000416721
$MSE(\hat{\beta}_{\text{female}})$	0.0021497401	0.0007941525	0.0002463387	0.0000420196
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.2667692152	0.2482611756	0.2484705686	0.2705662344
$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0044332383	-0.0229412778	-0.0227318849	-0.0006362191
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.0163466010	-0.0845909672	-0.0838188762	-0.0023459194
$Var(\hat{\beta}_{\text{no graduation}})$	0.0740871331	0.0360337091	0.0100320007	0.0019232689
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0741067867	0.0365600113	0.0105487393	0.0019236736
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2401763251	-0.2379387108	-0.2482590915	-0.2264865895
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.0142107790	-0.0119731647	-0.0222935454	-0.0005210433
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.0628891404	-0.0529866827	-0.0986590466	-0.0023058530
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0599933814	0.0201714650	0.0070767878	0.0012519858
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0601953276	0.0203148217	0.0075737900	0.0012522573
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.1921316063	0.1985470586	0.1912779493	0.2083843601
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0177205615	-0.0113051092	-0.0185742184	-0.0014678077
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.0844430708	-0.0538717770	-0.0885109675	-0.0069944841
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0646062556	0.0218427094	0.0073653740	0.0012599010
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0649202739	0.0219705149	0.0077103756	0.0012620555
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0261523086	-0.0137676807	-0.0269136641	-0.0109995069
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0148533917	-0.0024687638	-0.0156147472	0.0002994100
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	-1.3145854479	-0.2184956145	-1.3819685012	0.0264989974
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0691581168	0.0229086837	0.0076033194	0.0012431210
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0693787401	0.0229147785	0.0078471397	0.0012432106
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0132570382	0.0198921241	0.0155041277	0.0311086429
$Bias(\hat{\beta}_{\text{hochschulreife}})$	-0.0189707071	-0.0123356213	-0.0167236176	-0.0011191024
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	-0.5886451853	-0.3827640175	-0.5189198756	-0.0347248126
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0631801248	0.0223943216	0.0074832986	0.0014152285
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0635400125	0.0225464891	0.0077629780	0.0014164809
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4635059420	-0.4618865270	-0.4564965353	-0.4585660256
$Bias(\hat{\beta}_{\text{half-time}})$	-0.0059938730	-0.0043744579	0.0010155338	-0.0010539566
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	-0.0131010161	-0.0095614044	0.0022196874	-0.0023036694
$Var(\hat{\beta}_{\text{half-time}})$	0.0043138652	0.0021968697	0.0006481069	0.0001413350
$MSE(\hat{\beta}_{\text{half-time}})$	0.0043497917	0.0022160056	0.0006491382	0.0001424458
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9217667994	-0.9351950712	-0.9330820608	-0.9346102285
$Bias(\hat{\beta}_{\text{part-time}})$	0.0151384165	0.0017101446	0.0038231550	0.0022949874

$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0161578954	0.0018253123	0.0040806209	0.0024495406
$Var(\hat{\beta}_{\text{part-time}})$	0.0066663949	0.0026538802	0.0010475571	0.0002074612
$MSE(\hat{\beta}_{\text{part-time}})$	0.0068955666	0.0026568048	0.0010621736	0.0002127282
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.8015241987	-0.8011926588	-0.7965262807	-0.7975925442
$Bias(\hat{\beta}_{\text{not employed}})$	-0.0030630789	-0.0027315390	0.0019348392	0.0008685757
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	-0.0038362279	-0.0034210043	0.0024232103	0.0010878122
$Var(\hat{\beta}_{\text{not employed}})$	0.0017426613	0.0009005808	0.0002659352	0.0000702024
$MSE(\hat{\beta}_{\text{not employed}})$	0.0017520438	0.0009080421	0.0002696788	0.0000709569
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0920676192	0.0967508619	0.0908529327	0.0921221804
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0002784913	0.0044047514	-0.0014931778	-0.0002239301
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.0030157337	0.0476982879	-0.0161693630	-0.0024249001
$Var(\hat{\beta}_{\text{married living together}})$	0.0054620172	0.0024183112	0.0006143188	0.0001582624
$MSE(\hat{\beta}_{\text{married living together}})$	0.0054620948	0.0024377131	0.0006165484	0.0001583126
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3681931929	0.3880023402	0.3736138009	0.3857459803
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0121931989	0.0076159484	-0.0067725908	0.0053595885
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0320547716	0.0200216112	-0.0178045035	0.0140898534
$Var(\hat{\beta}_{\text{married living apart}})$	0.0357831874	0.0114628438	0.0032180375	0.0008689526
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0359318615	0.0115208465	0.0032639055	0.0008976778
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0151447621	-0.0037141123	-0.0087044185	-0.0085167130
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0064104819	0.0050201679	0.0000298617	0.0002175672
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.7339450746	0.5747660700	0.0034189064	0.0249095767
$Var(\hat{\beta}_{\text{widowed}})$	0.0118443623	0.0047536147	0.0015703975	0.0003322459
$MSE(\hat{\beta}_{\text{widowed}})$	0.0118854566	0.0047788168	0.0015703984	0.0003322932
β_{single}	0.1244899786	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1113918121	0.1281221820	0.1221197386	0.1244998525
$Bias(\hat{\beta}_{\text{single}})$	-0.0130981665	0.0036322033	-0.0023702400	0.0000098739
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.1052146258	0.0291766726	-0.0190396050	0.0000793148
$Var(\hat{\beta}_{\text{single}})$	0.0061412553	0.0026393525	0.0008664834	0.0002047678
$MSE(\hat{\beta}_{\text{single}})$	0.0063128172	0.0026525454	0.0008721015	0.0002047679
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0136398494	-0.0077270092	-0.0237686035	-0.0271556921
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0114009971	0.0173138373	0.0012722430	-0.0021148456
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.4552959930	0.6914238036	0.0508067091	-0.0844558367
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0094649706	0.0045630179	0.0012392035	0.0002950376
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0095949533	0.0048627868	0.0012408221	0.0002995102
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0042791273	0.0008088493	-0.0116774568	-0.0153175330
$Bias(\hat{\beta}_{\text{SPD}})$	0.0090485924	0.0141365690	0.0016502628	-0.0019898133
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.6789302758	1.0606892509	0.1238218442	-0.1492988600
$Var(\hat{\beta}_{\text{SPD}})$	0.0116384237	0.0049884448	0.0013911627	0.0002534642

$MSE(\hat{\beta}_{\text{SPD}})$	0.0117203007	0.0051882874	0.0013938861	0.0002574235
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0886217867	0.0992888002	0.0887527063	0.0852647617
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0015400048	0.0122070183	0.0016709244	-0.0018170202
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.0176845804	0.1401787842	0.0191879908	-0.0208656760
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0097914855	0.0052945022	0.0017695885	0.0003070267
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0097938571	0.0054435135	0.0017723805	0.0003103283
$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0854331815	0.0972642360	0.0730106325	0.0740689852
$Bias(\hat{\beta}_{\text{die linke}})$	0.0085561692	0.0203872238	-0.0038663797	-0.0028080271
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	0.1112968492	0.2651927177	-0.0502930539	-0.0365262253
$Var(\hat{\beta}_{\text{die linke}})$	0.0109641478	0.0064582175	0.0017362743	0.0003051704
$MSE(\hat{\beta}_{\text{die linke}})$	0.0110373558	0.0068738564	0.0017512232	0.0003130554
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0104302379	0.0292413250	0.0063154100	0.0039102728
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	0.0030296505	0.0218407375	-0.0010851775	-0.0034903146
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	0.4093797246	2.9512167302	-0.1466339638	-0.4716267004
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0149667918	0.0065513955	0.0019074206	0.0004066915
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0149759706	0.0070284133	0.0019085982	0.0004188738
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1533585804	-0.1404249163	-0.1525769637	-0.1565240139
$Bias(\hat{\beta}_{\text{FDP}})$	0.0008170298	0.0137506939	0.0015986465	-0.0023484037
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.0052993455	0.0891885162	0.0103689975	-0.0152320055
$Var(\hat{\beta}_{\text{FDP}})$	0.0157989041	0.0077098198	0.0022374883	0.0004079139
$MSE(\hat{\beta}_{\text{FDP}})$	0.0157995717	0.0078989014	0.0022400439	0.0004134289
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1271683464	-0.1074205165	-0.1272127137	-0.1348080681
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0044205621	0.0241683921	0.0043761948	-0.0032191596
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.0335937287	0.1836658752	0.0332565627	-0.0244637606
$Var(\hat{\beta}_{\text{would not vote}})$	0.0140566303	0.0055773995	0.0021707360	0.0003321823
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0140761716	0.0061615106	0.0021898871	0.0003425453

Table C.6: Expected value, variance, bias and MSE of regression coefficient estimates in order to quantify sampling error

Categories	p-value (true)	p-value ($n = 500$)	p-Value ($n = 1000$)	p-Value ($n = 2000$)	p-Value ($n = 3000$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000133476	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000124	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.4846414563	0.5231084013	0.5852652096	0.7722003381
hochschulreife	0.6987555776	0.4845411493	0.5212755229	0.5873354924	0.6837340211
mittlere reife	0.0115117133	0.3974786005	0.3086938672	0.1650586097	0.0307352211
no graduation	0.0058831082	0.3724683448	0.2929106762	0.1230947617	0.0186274942
volks-, hauptschule	0.0065408724	0.3653650084	0.2367214013	0.0731967852	0.0198723856
half-time	0.0000000000	0.0000012303	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.3031055156	0.1154705119	0.0240584382	0.0016810460
married living apart	0.0000000011	0.1298902272	0.0185224655	0.0003474806	0.0000001691
married living together	0.0021189162	0.3271530448	0.1734636773	0.0551254078	0.0091309923
widowed	0.8397563199	0.5181479416	0.5466028711	0.6142017013	0.7487346801
CDU-CSU	0.5884698278	0.5773698947	0.5805196546	0.5966467209	0.5962387988
die gruenen	0.0762638735	0.4502705721	0.3645146319	0.2465045381	0.1268236635
die linke	0.1299242680	0.4676528694	0.3413259703	0.3478680914	0.1979834927
FDP	0.0082560314	0.3873638288	0.2867143285	0.0923837776	0.0183194722
extreme right-wing	0.8910480181	0.5336364888	0.5688526553	0.6397568897	0.7806511905
SPD	0.7770122356	0.5627509510	0.5699940150	0.6424052852	0.7245567738
would not vote	0.0103321856	0.4310941703	0.3363123418	0.1190782728	0.0213780100

Table C.7: Comparison of p-values of regression coefficients of different samples

	$n = 500$	$n = 1000$	$n = 2000$	$n = 3000$
β_0	7.0070608279	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.1154379263	7.0801986559	7.0755887289	7.0399111779
$Bias(\hat{\beta}_0)$	0.1083770984	0.0731378281	0.0685279010	0.0328503500
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0154668414	0.0104377327	0.0097798353	0.0046881782
$Var(\hat{\beta}_0)$	0.1629392190	0.0706184277	0.0194462066	0.0032271841
$MSE(\hat{\beta}_0)$	0.1746848145	0.0759675696	0.0241422798	0.0043063296
β_{Age}	0.0113720951	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{Age}})$	0.0117635838	0.0117257881	0.0117127652	0.0115600468
$Bias(\hat{\beta}_{\text{Age}})$	0.0003914887	0.0003536930	0.0003406701	0.0001879516
$Bias_{\text{rel}}(\hat{\beta}_{\text{Age}})$	0.0344253832	0.0311018349	0.0299566712	0.0165274420
$Var(\hat{\beta}_{\text{Age}})$	0.0000030547	0.0000013244	0.0000003802	0.0000000545
$MSE(\hat{\beta}_{\text{Age}})$	0.0000032080	0.0000014495	0.0000004963	0.0000000899
β_{Female}	-0.3783040793	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{Female}})$	-0.3909264185	-0.3878612432	-0.3841711729	-0.3809456803
$Bias(\hat{\beta}_{\text{Female}})$	-0.0126223392	-0.0095571639	-0.0058670936	-0.0026416010
$Bias_{\text{rel}}(\hat{\beta}_{\text{Female}})$	-0.0333655910	-0.0252631796	-0.0155089357	-0.0069827453
$Var(\hat{\beta}_{\text{Female}})$	0.0029370984	0.0009905647	0.0002482561	0.0000548808
$MSE(\hat{\beta}_{\text{Female}})$	0.0030964218	0.0010819041	0.0002826789	0.0000618589
$\beta_{\text{No Graduation}}$	0.2712024535	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{No Graduation}})$	0.2041747407	0.2072411454	0.2239710333	0.2359776345
$Bias(\hat{\beta}_{\text{No Graduation}})$	-0.0670277128	-0.0639613081	-0.0472314201	-0.0352248189

$Bias_{\text{rel}}(\hat{\beta}_{\text{No Graduation}})$	-0.2471500973	-0.2358433977	-0.1741555784	-0.1298838506
$Var(\hat{\beta}_{\text{No Graduation}})$	0.2130278055	0.0913305805	0.0211148732	0.0038252367
$MSE(\hat{\beta}_{\text{No Graduation}})$	0.2175205198	0.0954216295	0.0233456802	0.0050660246
$\beta_{\text{Volks-, Hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{Volks-, Hauptschule}})$	-0.2711525776	-0.2604853393	-0.2647295971	-0.2529969353
$Bias(\hat{\beta}_{\text{Volks-, Hauptschule}})$	-0.0451870315	-0.0345197932	-0.0387640510	-0.0270313892
$Bias_{\text{rel}}(\hat{\beta}_{\text{Volks-, Hauptschule}})$	-0.1999731031	-0.1527657371	-0.1715485023	-0.1196261539
$Var(\hat{\beta}_{\text{Volks-, Hauptschule}})$	0.1399360740	0.0679637153	0.0143599491	0.0024680627
$MSE(\hat{\beta}_{\text{Volks-, Hauptschule}})$	0.1419779418	0.0691553314	0.0158626008	0.0031987587
$\beta_{\text{Mittlere Reife}}$	0.2098521677	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{Mittlere Reife}})$	0.1558676143	0.1770297149	0.1708281882	0.1865496598
$Bias(\hat{\beta}_{\text{Mittlere Reife}})$	-0.0539845535	-0.0328224529	-0.0390239795	-0.0233025079
$Bias_{\text{rel}}(\hat{\beta}_{\text{Mittlere Reife}})$	-0.2572503970	-0.1564074998	-0.1859593823	-0.1110424932
$Var(\hat{\beta}_{\text{Mittlere Reife}})$	0.1299092479	0.0644033767	0.0148165139	0.0025691219
$MSE(\hat{\beta}_{\text{Mittlere Reife}})$	0.1328235799	0.0654806901	0.0163393849	0.0031121287
$\beta_{\text{Fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{Fachhochschulreife}})$	-0.0654172855	-0.0500391696	-0.0553047500	-0.0390041912
$Bias(\hat{\beta}_{\text{Fachhochschulreife}})$	-0.0541183686	-0.0387402527	-0.0440058331	-0.0277052743
$Bias_{\text{rel}}(\hat{\beta}_{\text{Fachhochschulreife}})$	-4.7896952532	-3.4286695897	-3.8946948242	-2.4520292174
$Var(\hat{\beta}_{\text{Fachhochschulreife}})$	0.1354824235	0.0658343444	0.0143433155	0.0025044716
$MSE(\hat{\beta}_{\text{Fachhochschulreife}})$	0.1384112213	0.0673351515	0.0162798288	0.0032720538
$\beta_{\text{Hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{Hochschulreife}})$	-0.0436979950	-0.0179163771	-0.0174262311	0.0035657540
$Bias(\hat{\beta}_{\text{Hochschulreife}})$	-0.0759257404	-0.0501441224	-0.0496539765	-0.0286619914
$Bias_{\text{rel}}(\hat{\beta}_{\text{Hochschulreife}})$	-2.3559122595	-1.5559302049	-1.5407213860	-0.8893576354
$Var(\hat{\beta}_{\text{Hochschulreife}})$	0.1324042309	0.0634406269	0.0142899246	0.0025508383
$MSE(\hat{\beta}_{\text{Hochschulreife}})$	0.1381689489	0.0659550599	0.0167554420	0.0033723480
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4506389082	-0.4552266323	-0.4572321058	-0.4554729567
$Bias(\hat{\beta}_{\text{half-time}})$	0.0068731608	0.0022854368	0.0002799632	0.0020391123
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.0150229061	0.0049953584	0.0006119254	0.0044569585
$Var(\hat{\beta}_{\text{half-time}})$	0.0029624269	0.0010298490	0.0003175442	0.0000424009
$MSE(\hat{\beta}_{\text{half-time}})$	0.0030096672	0.0010350722	0.0003176226	0.0000465588
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9463786577	-0.9435444489	-0.9496333357	-0.9408212044
$Bias(\hat{\beta}_{\text{part-time}})$	-0.0094734419	-0.0066392330	-0.0127281198	-0.0039159885
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	-0.0101114197	-0.0070863444	-0.0135852802	-0.0041797062
$Var(\hat{\beta}_{\text{part-time}})$	0.0058894577	0.0020289291	0.0004025726	0.0001273966
$MSE(\hat{\beta}_{\text{part-time}})$	0.0059792038	0.0020730085	0.0005645777	0.0001427316
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7880753230	-0.7888406589	-0.7930340586	-0.7932251606
$Bias(\hat{\beta}_{\text{not employed}})$	0.0103857969	0.0096204610	0.0054270613	0.0052359593
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.0130072669	0.0120487532	0.0067969012	0.0065575633
$Var(\hat{\beta}_{\text{not employed}})$	0.0019497542	0.0007154160	0.0002048933	0.0000503758

$MSE(\hat{\beta}_{\text{not employed}})$	0.0020576189	0.0008079692	0.0002343462	0.0000777910
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105	0.09234611105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0566766426	0.0579132093	0.0711183792	0.0847817048
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0356694680	-0.0344329012	-0.0212277314	-0.0075644057
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.3862584766	-0.3728679101	-0.2298714179	-0.0819136366
$Var(\hat{\beta}_{\text{married living together}})$	0.0046671366	0.0015574891	0.0004044986	0.0001025380
$MSE(\hat{\beta}_{\text{married living together}})$	0.0059394476	0.0027431138	0.0008551152	0.0001597583
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3247103622	0.3091225094	0.3261395349	0.3668601023
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0556760296	-0.0712638823	-0.0542468569	-0.0135262895
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.1463670384	-0.1873460352	-0.1426098779	-0.0355593412
$Var(\hat{\beta}_{\text{married living apart}})$	0.0283246760	0.0108040363	0.0021455458	0.0003498134
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0314244963	0.0158825772	0.0050882673	0.0005327739
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0618879829	-0.0535112932	-0.0422748173	-0.0213010153
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0531537027	-0.0447770130	-0.0335405371	-0.0125667351
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-6.0856420339	-5.1265830710	-3.8401031802	-1.4387831412
$Var(\hat{\beta}_{\text{widowed}})$	0.0085802491	0.0045946576	0.0010508667	0.0001783559
$MSE(\hat{\beta}_{\text{widowed}})$	0.0114055652	0.0065996384	0.0021758343	0.0003362787
β_{single}	0.1244899786	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.0924260433	0.0910865296	0.1069843653	0.1230883028
$Bias(\hat{\beta}_{\text{single}})$	-0.0320639354	-0.0334034490	-0.0175056134	-0.0014016758
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.2575623815	-0.2683223935	-0.1406186552	-0.0112593466
$Var(\hat{\beta}_{\text{single}})$	0.0056518620	0.0019645839	0.0006703066	0.0001413639
$MSE(\hat{\beta}_{\text{single}})$	0.0066799580	0.0030803743	0.0009767531	0.0001433286
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0482096328	-0.0355994707	-0.0421670019	-0.0350611259
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	-0.0231687863	-0.0105586242	-0.0171261555	-0.0100202794
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	-0.9252397398	-0.4216560402	-0.6839287751	-0.4001573738
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0081771019	0.0033155843	0.0011561434	0.0003079281
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0087138946	0.0034270689	0.0014494486	0.0004083341
β_{SPD}	-0.0133277166	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0349140965	-0.0198381719	-0.0274502355	-0.0196215408
$Bias(\hat{\beta}_{\text{SPD}})$	-0.0215863769	-0.0065104522	-0.0141225158	-0.0062938212
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	-1.6196601877	-0.4884895839	-1.0596348223	-0.4722354133
$Var(\hat{\beta}_{\text{SPD}})$	0.0094781352	0.0039399756	0.0012287754	0.0003702036
$MSE(\hat{\beta}_{\text{SPD}})$	0.0099441068	0.0039823616	0.0014282208	0.0004098157
$\beta_{\text{Die Gruenen}}$	0.0870817819	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{Die Gruenen}})$	0.0466039401	0.0674736227	0.0619250403	0.0756031311
$Bias(\hat{\beta}_{\text{Die Gruenen}})$	-0.0404778418	-0.0196081592	-0.0251567416	-0.0114786508
$Bias_{\text{rel}}(\hat{\beta}_{\text{Die Gruenen}})$	-0.4648256027	-0.2251694765	-0.2888863904	-0.1318146061
$Var(\hat{\beta}_{\text{Die Gruenen}})$	0.0080156577	0.0040112863	0.0011584808	0.0003473714
$MSE(\hat{\beta}_{\text{Die Gruenen}})$	0.0096541134	0.0043957662	0.0017913425	0.0004791308
$\beta_{\text{Die Linke}}$	0.0768770122	0.0768770122	0.0768770122	0.0768770122

$\mathbb{E}(\hat{\beta}_{\text{Die Linke}})$	0.0341241782	0.0526020173	0.0504614260	0.0644332085
$Bias(\hat{\beta}_{\text{Die Linke}})$	-0.0427528341	-0.0242749949	-0.0264155862	-0.0124438038
$Bias_{\text{rel}}(\hat{\beta}_{\text{Die Linke}})$	-0.5561198701	-0.3157640260	-0.3436083872	-0.1618663813
$Var(\hat{\beta}_{\text{Die Linke}})$	0.0110602866	0.0039829128	0.0014641559	0.0003888114
$MSE(\hat{\beta}_{\text{Die Linke}})$	0.0128880915	0.0045721882	0.0021619391	0.0005436597
$\beta_{\text{Extreme Right-Wing}}$	0.0074005875	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{Extreme Right-Wing}})$	0.0010846376	0.0128964735	-0.0066442158	-0.0103847410
$Bias(\hat{\beta}_{\text{Extreme Right-Wing}})$	-0.0063159498	0.0054958860	-0.0140448032	-0.0177853285
$Bias_{\text{rel}}(\hat{\beta}_{\text{Extreme Right-Wing}})$	-0.8534389804	0.7426283442	-1.8977957225	-2.4032319835
$Var(\hat{\beta}_{\text{Extreme Right-Wing}})$	0.0145233979	0.0058199623	0.0014233644	0.0004314324
$MSE(\hat{\beta}_{\text{Extreme Right-Wing}})$	0.0145632891	0.0058501670	0.0016206209	0.0007477504
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1956125463	-0.1799309386	-0.1777507777	-0.1606059822
$Bias(\hat{\beta}_{\text{FDP}})$	-0.0414369361	-0.0257553284	-0.0235751674	-0.0064303720
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	-0.2687645344	-0.1670518986	-0.1529111343	-0.0417081013
$Var(\hat{\beta}_{\text{FDP}})$	0.0163115187	0.0070842518	0.0020469777	0.0005364632
$MSE(\hat{\beta}_{\text{FDP}})$	0.0180285383	0.0077475888	0.0026027662	0.0005778129
$\beta_{\text{Would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{Would not vote}})$	-0.1744103120	-0.1549665057	-0.1622630339	-0.1454320909
$Bias(\hat{\beta}_{\text{Would not vote}})$	-0.0428214035	-0.0233775972	-0.0306741253	-0.0138431824
$Bias_{\text{rel}}(\hat{\beta}_{\text{Would not vote}})$	-0.3254180305	-0.1776562889	-0.2331057054	-0.1052002218
$Var(\hat{\beta}_{\text{Would not vote}})$	0.0121725903	0.0042653210	0.0015900656	0.0004180957
$MSE(\hat{\beta}_{\text{Would not vote}})$	0.0140062628	0.0048118330	0.0025309676	0.0006097294

Table C.8: Expected value, variance, bias and MSE of regression coefficient estimates in order to quantify coverage error

Categories	p-value (true)	p-value ($n = 500$)	p-Value ($n = 1000$)	p-Value ($n = 2000$)	p-Value ($n = 3000$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000062293	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000001079	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.5254398841	0.4866904587	0.5417766172	0.6430866228
hochschulreife	0.6987555776	0.5162105184	0.4914733347	0.5495164328	0.7076627534
mittlere reife	0.0115117133	0.4957747147	0.3920019472	0.3169943782	0.0974987427
no graduation	0.0058831082	0.4730777192	0.4420658532	0.2911230253	0.0814360783
volks-, hauptschule	0.0065408724	0.3930195062	0.3244654779	0.1596933291	0.0244991542
half-time	0.0000000000	0.0000004691	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.3621467593	0.2261101346	0.0381519529	0.0013772571
married living apart	0.0000000011	0.1463066881	0.0481046598	0.0004968581	0.0000001337
married living together	0.0021189162	0.4513344733	0.3719304289	0.1065161133	0.0124650985
widowed	0.8397563199	0.5042984411	0.4799692008	0.4540290796	0.6432394692
CDU-CSU	0.5884698278	0.5824024177	0.5529641618	0.4952469254	0.5079035682
die gruenen	0.0762638735	0.5567131829	0.4635259079	0.3827120727	0.1754585390
die linke	0.1299242680	0.5884244556	0.5414375475	0.4656254480	0.2667191418
FDP	0.0082560314	0.3235326917	0.1959597457	0.0475838943	0.0170843316
extreme right-wing	0.8910480181	0.5962912230	0.5925751327	0.7019790700	0.7566821879
SPD	0.7770122356	0.5747152031	0.5818393332	0.5957048115	0.6557235775
would not vote	0.0103321856	0.3126469192	0.1840822160	0.0427038986	0.0139677242

Table C.9: Comparison of p-values of regression coefficients of different samples containing coverage error

	MCAR	MAR	MNAR
β_0	7.0044034272	7.0042831638	7.0008817670
$\mathbb{E}(\hat{\beta}_0)$	7.0044034272	7.0042831638	7.0008817670
$Bias(\hat{\beta}_0)$	-0.0026574007	-0.0027776641	-0.0061790609
$Bias_{\text{rel}}(\hat{\beta}_0)$	-0.0003792461	-0.0003964093	-0.0008818335
$Var(\hat{\beta}_0)$	0.0041978230	0.0050734059	0.0044389459
$MSE(\hat{\beta}_0)$	0.0042048848	0.0050811213	0.0044771267
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0114018304	0.0113304028	0.0112812518
$Bias(\hat{\beta}_{\text{age}})$	0.0000297353	-0.0000416923	-0.0000908434
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	0.0026147616	-0.0036661908	-0.0079882682
$Var(\hat{\beta}_{\text{age}})$	0.0000001371	0.0000002409	0.0000002327
$MSE(\hat{\beta}_{\text{age}})$	0.0000001380	0.0000002427	0.0000002409
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.3796061235	-0.3790319730	-0.3779256473
$Bias(\hat{\beta}_{\text{female}})$	-0.0013020442	-0.0007278937	0.0003784320
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	-0.0034417928	-0.0019240969	0.0010003380
$Var(\hat{\beta}_{\text{female}})$	0.0000921494	0.0001521025	0.0001761813
$MSE(\hat{\beta}_{\text{female}})$	0.0000938447	0.0001526323	0.0001763245
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.2682868171	0.2895406671	0.2895335753

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0029156364	0.0183382136	0.0183311218
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.0107507743	0.0676181700	0.0675920207
$Var(\hat{\beta}_{\text{no graduation}})$	0.0039703132	0.0044803749	0.0047509944
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0039788141	0.0048166650	0.0050870244
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2245508928	-0.2169712114	-0.2172227621
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0014146533	0.0089943348	0.0087427841
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0062604824	0.0398040097	0.0386907836
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0029822702	0.0033866079	0.0032989027
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0029842714	0.0034675059	0.0033753390
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.2106275137	0.2149192017	0.2149819088
$Bias(\hat{\beta}_{\text{mittlere reife}})$	0.0007753460	0.0050670340	0.0051297410
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	0.0036947246	0.0241457310	0.0244445462
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0030923626	0.0031790084	0.0031659124
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0030929637	0.0032046833	0.0031922267
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{Fachhochschule}})$	-0.0087201464	-0.0048945265	-0.0020178910
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	0.0025787705	0.0064043903	0.0092810259
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	0.2282316533	0.5668145371	0.8214084594
$Var(\hat{\beta}_{\text{Fachhochschule}})$	0.0033618707	0.0034005100	0.0034579845
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0033685208	0.0034415262	0.0035441220
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0344479538	0.0378994142	0.0381753413
$Bias(\hat{\beta}_{\text{hochschulreife}})$	0.0022202084	0.0056716688	0.0059475959
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	0.0688912119	0.1759871435	0.1845489299
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0031001066	0.0030995121	0.0029924109
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0031050359	0.0031316799	0.0030277848
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4557659486	-0.4590397976	-0.4578473504
$Bias(\hat{\beta}_{\text{half-time}})$	0.0017461204	-0.0015277285	-0.0003352814
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.0038165560	-0.0033392092	-0.0007328361
$Var(\hat{\beta}_{\text{half-time}})$	0.0002267748	0.0002913716	0.0003131457
$MSE(\hat{\beta}_{\text{half-time}})$	0.0002298237	0.0002937056	0.0003132581
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9364030091	-0.9333178892	-0.9288870329
$Bias(\hat{\beta}_{\text{part-time}})$	0.0005022067	0.0035873267	0.0080181829
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0005360273	0.0038289110	0.0085581581
$Var(\hat{\beta}_{\text{part-time}})$	0.0003779516	0.0004074637	0.0004748919
$MSE(\hat{\beta}_{\text{part-time}})$	0.0003782038	0.0004203327	0.0005391832
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7993882496	-0.8014593189	-0.7984637338
$Bias(\hat{\beta}_{\text{not employed}})$	-0.0009271297	-0.0029981990	-0.0000026139
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	-0.0011611457	-0.0037549718	-0.0000032737

$Var(\hat{\beta}_{\text{not employed}})$	0.0000939353	0.0001488469	0.0001497775
$MSE(\hat{\beta}_{\text{not employed}})$	0.0000947949	0.0001578360	0.0001497775
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0936079377	0.0921728141	0.0919017457
$Bias(\hat{\beta}_{\text{married living together}})$	0.0012618272	-0.0001732964	-0.0004443649
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	0.0136641076	-0.0018765968	-0.0048119499
$Var(\hat{\beta}_{\text{married living together}})$	0.0002613860	0.0003101330	0.0004057326
$MSE(\hat{\beta}_{\text{married living together}})$	0.0002629782	0.0003101631	0.0004059300
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3779813415	0.3903509189	0.3895756962
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0024050503	0.0099645271	0.0091893044
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0063226507	0.0261958033	0.0241578159
$Var(\hat{\beta}_{\text{married living apart}})$	0.0013323587	0.0013212501	0.0014968220
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0013381430	0.0014205419	0.0015812653
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0049415034	-0.0124670012	-0.0111401483
$Bias(\hat{\beta}_{\text{widowed}})$	0.0037927768	-0.0037327210	-0.0024058681
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	0.4342403374	-0.4273644706	-0.2754512147
$Var(\hat{\beta}_{\text{widowed}})$	0.0004357986	0.0008083337	0.0010387319
$MSE(\hat{\beta}_{\text{widowed}})$	0.0004501837	0.0008222669	0.0010445201
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1259458391	0.1299328407	0.1294214353
$Bias(\hat{\beta}_{\text{single}})$	0.0014558605	0.0054428621	0.0049314567
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	0.0116945997	0.0437212871	0.0396132822
$Var(\hat{\beta}_{\text{single}})$	0.0003002450	0.0004858308	0.0005644980
$MSE(\hat{\beta}_{\text{single}})$	0.0003023645	0.0005154555	0.0005888173
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0255039714	-0.0283724049	-0.0290806502
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	-0.0004631249	-0.0033315584	-0.0040398037
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0184947784	-0.1330449610	-0.1613285590
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0004779072	0.0006164845	0.0006384744
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0004781217	0.0006275838	0.0006547944
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0150842255	-0.0157617815	-0.0163736394
$Bias(\hat{\beta}_{\text{SPD}})$	-0.0017565059	-0.0024340619	-0.0030459197
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	-0.1317934297	-0.1826315354	-0.2285402020
$Var(\hat{\beta}_{\text{SPD}})$	0.0005364426	0.0006222658	0.0006232001
$MSE(\hat{\beta}_{\text{SPD}})$	0.0005395279	0.0006281904	0.0006324777
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0877530302	0.0852070906	0.0838418153
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0006712483	-0.0018746913	-0.0032399666
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.0077082518	-0.0215279388	-0.0372060207
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0004816770	0.0006472476	0.0007078549
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0004821276	0.0006507621	0.0007183523

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0743199506	0.0825931241	0.0817961542
$Bias(\hat{\beta}_{\text{die linke}})$	-0.0025570617	0.0057161118	0.0049191419
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	-0.0332617200	0.0743539799	0.0639871633
$Var(\hat{\beta}_{\text{die linke}})$	0.0006040029	0.0007439868	0.0008599807
$MSE(\hat{\beta}_{\text{die linke}})$	0.0006105415	0.0007766607	0.0008841787
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0057468268	0.0075170569	0.0076505304
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0016537607	0.0001164694	0.0002499429
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-0.2234634311	0.0157378598	0.0337733860
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0005889750	0.0009249351	0.0011073161
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0005917100	0.0009249487	0.0011073786
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1589626479	-0.1460601385	-0.1474668797
$Bias(\hat{\beta}_{\text{FDP}})$	-0.0047870377	0.0081154717	0.0067087305
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	-0.0310492543	0.0526378437	0.0435135657
$Var(\hat{\beta}_{\text{FDP}})$	0.0008853340	0.0009259273	0.0010069921
$MSE(\hat{\beta}_{\text{FDP}})$	0.0009082497	0.0009917881	0.0010519991
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1310548328	-0.1425775939	-0.1453572762
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0005340757	-0.0109886854	-0.0137683676
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.0040586684	-0.0835076869	-0.1046316728
$Var(\hat{\beta}_{\text{would not vote}})$	0.0006570435	0.0009800988	0.0009495598
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0006573288	0.0011008500	0.0011391278

Table C.10: Expected value, variance, bias and MSE of regression coefficient estimates in order to quantify nonresponse error

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.6680739711	0.6628283582	0.6562783851
hochschulreife	0.6987555776	0.6423145757	0.6299941400	0.6206511203
mittlere reife	0.0115117133	0.0494781753	0.0574688244	0.0599561792
no graduation	0.0058831082	0.0348987904	0.0297719923	0.0344701548
volks-, hauptschule	0.0065408724	0.0433796888	0.0612560170	0.0584476609
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0025815205	0.0045112674	0.0059004752
married living apart	0.0000000011	0.0000037581	0.0000169972	0.0000208887
married living together	0.0021189162	0.0139212002	0.0201877015	0.0250424959
widowed	0.8397563199	0.7309788247	0.6571680432	0.6296633355
CDU-CSU	0.5884698278	0.6130190069	0.5709998942	0.5678385077
die gruenen	0.0762638735	0.1414358283	0.1723656108	0.1858294288
die linke	0.1299242680	0.2333765248	0.2057032345	0.2193862114
FDP	0.0082560314	0.0284163765	0.0511916966	0.0525880985
extreme right-wing	0.8910480181	0.7587440687	0.7207747567	0.6786652989
SPD	0.7770122356	0.6988377426	0.6789818085	0.6799991566
would not vote	0.0103321856	0.0375766963	0.0341556432	0.0312987438

Table C.11: Comparison of p-values of regression coefficients of data containing non-response error

	RCAR	RAR	RNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0084646257	7.0089288686	7.0089286062
$Bias(\hat{\beta}_0)$	0.0014037978	0.0018680408	0.0018677784
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0002003405	0.0002665941	0.0002665566
$Var(\hat{\beta}_0)$	0.0000003189	0.0000001437	0.0000001401
$MSE(\hat{\beta}_0)$	0.0000022895	0.0000036333	0.0000036287
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0113690764	0.0113687551	0.0113687614
$Bias(\hat{\beta}_{\text{age}})$	-0.0000030187	-0.0000033400	-0.0000033337
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.0002654459	-0.0002936979	-0.0002931447
$Var(\hat{\beta}_{\text{age}})$	0.0000000000	0.0000000000	0.0000000000
$MSE(\hat{\beta}_{\text{age}})$	0.0000000000	0.0000000000	0.0000000000
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.3782736201	-0.3783088661	-0.3783078528
$Bias(\hat{\beta}_{\text{female}})$	0.0000304592	-0.0000047868	-0.0000037735
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	0.0000805150	-0.0000126534	-0.0000099748
$Var(\hat{\beta}_{\text{female}})$	0.0000000049	0.0000000023	0.0000000022
$MSE(\hat{\beta}_{\text{female}})$	0.0000000058	0.0000000023	0.0000000023
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.2712149936	0.2712518066	0.2712535837

$Bias(\hat{\beta}_{\text{no graduation}})$	0.0000125401	0.0000493531	0.0000511302
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	0.0000462389	0.0001819789	0.0001885317
$Var(\hat{\beta}_{\text{no graduation}})$	0.0000002459	0.0000000893	0.0000000882
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0000002461	0.0000000917	0.0000000908
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2259544934	-0.2260519017	-0.2260478722
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0000110527	-0.0000863556	-0.0000823260
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0000489132	-0.0003821626	-0.0003643300
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0000002078	0.0000001047	0.0000001036
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0000002079	0.0000001121	0.0000001104
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.2093436728	0.2091968139	0.2091982920
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0005084949	-0.0006553539	-0.0006538757
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.0024231101	-0.0031229311	-0.0031158873
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0000002160	0.0000000896	0.0000000881
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0000004746	0.0000005191	0.0000005157
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0116755300	-0.0117970193	-0.0117951799
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0003766131	-0.0004981024	-0.0004962630
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0333317893	-0.0440840809	-0.0439212871
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0000002267	0.0000001051	0.0000001032
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0000003686	0.0000003532	0.0000003495
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0317411205	0.0315814419	0.0315837876
$Bias(\hat{\beta}_{\text{hochschulreife}})$	-0.0004866248	-0.0006463035	-0.0006439577
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	-0.0150995619	-0.0200542573	-0.0199814704
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0000002187	0.0000000930	0.0000000917
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0000004555	0.0000005107	0.0000005064
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4573032442	-0.4572538825	-0.4572534373
$Bias(\hat{\beta}_{\text{half-time}})$	0.0002088249	0.0002581865	0.0002586318
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.0004564358	0.0005643273	0.0005653005
$Var(\hat{\beta}_{\text{half-time}})$	0.0000000083	0.0000000037	0.0000000036
$MSE(\hat{\beta}_{\text{half-time}})$	0.0000000519	0.0000000703	0.0000000705
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9365331814	-0.9364403503	-0.9364395479
$Bias(\hat{\beta}_{\text{part-time}})$	0.0003720345	0.0004648656	0.0004656679
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0003970887	0.0004961714	0.0004970278
$Var(\hat{\beta}_{\text{part-time}})$	0.0000000315	0.0000000161	0.0000000155
$MSE(\hat{\beta}_{\text{part-time}})$	0.0000001699	0.0000002322	0.0000002324
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7983260179	-0.7983004907	-0.7983000381
$Bias(\hat{\beta}_{\text{not employed}})$	0.0001351019	0.0001606292	0.0001610818
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.0001692029	0.0002011735	0.0002017403

$Var(\hat{\beta}_{\text{not employed}})$	0.0000000042	0.0000000021	0.0000000021
$MSE(\hat{\beta}_{\text{not employed}})$	0.0000000225	0.0000000279	0.0000000281
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0917948901	0.0916409130	0.0916400521
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0005512205	-0.0007051975	-0.0007060584
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.0059690709	-0.0076364615	-0.0076457843
$Var(\hat{\beta}_{\text{married living together}})$	0.0000000128	0.0000000060	0.0000000061
$MSE(\hat{\beta}_{\text{married living together}})$	0.0000003167	0.0000005033	0.0000005047
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3796755969	0.3794860530	0.3794842176
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0007107949	-0.0009003387	-0.0009021742
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0018686127	-0.0023669057	-0.0023717310
$Var(\hat{\beta}_{\text{married living apart}})$	0.0000000311	0.0000000147	0.0000000145
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0000005364	0.0000008253	0.0000008284
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0091867214	-0.0093377293	-0.0093380782
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0004524412	-0.0006034492	-0.0006037980
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.0518006287	-0.0690897408	-0.0691296794
$Var(\hat{\beta}_{\text{widowed}})$	0.0000000320	0.0000000188	0.0000000189
$MSE(\hat{\beta}_{\text{widowed}})$	0.0000002367	0.0000003829	0.0000003835
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1239477407	0.1237896387	0.1237878782
$Bias(\hat{\beta}_{\text{single}})$	-0.0005422379	-0.0007003399	-0.0007021004
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.0043556753	-0.0056256728	-0.0056398145
$Var(\hat{\beta}_{\text{single}})$	0.0000000176	0.0000000097	0.0000000097
$MSE(\hat{\beta}_{\text{single}})$	0.0000003116	0.0000005002	0.0000005027
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0253263800	-0.0254332576	-0.0254346528
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	-0.0002855335	-0.0003924111	-0.0003938063
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0114027107	-0.0156708415	-0.0157265563
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0000000213	0.0000000085	0.0000000085
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0000001028	0.0000001625	0.0000001636
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0134861602	-0.0135593763	-0.0135600462
$Bias(\hat{\beta}_{\text{SPD}})$	-0.0001584405	-0.0002316567	-0.0002323265
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	-0.0118880455	-0.0173815684	-0.0174318293
$Var(\hat{\beta}_{\text{SPD}})$	0.0000000237	0.0000000095	0.0000000094
$MSE(\hat{\beta}_{\text{SPD}})$	0.0000000488	0.0000000632	0.0000000634
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0871427152	0.0871422325	0.0871410800
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0000609333	0.0000604506	0.0000592981
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.0006997244	0.0006941819	0.0006809471
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0000000203	0.0000000143	0.0000000139
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0000001028	0.0000001625	0.0000001636

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0767100503	0.0766122897	0.0766110397
$Bias(\hat{\beta}_{\text{die linke}})$	-0.0001669620	-0.0002647225	-0.0002659726
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	-0.0021718062	-0.0034434545	-0.0034597152
$Var(\hat{\beta}_{\text{die linke}})$	0.0000000284	0.0000000171	0.0000000170
$MSE(\hat{\beta}_{\text{die linke}})$	0.0000000563	0.0000000872	0.0000000877
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0071263435	0.0069954801	0.0069957502
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0002742439	-0.0004051074	-0.0004048373
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-0.0370570481	-0.0547398981	-0.0547034011
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0000000272	0.0000000186	0.0000000183
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0000001024	0.0000001827	0.0000001822
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1544285033	-0.1545585934	-0.1545585614
$Bias(\hat{\beta}_{\text{FDP}})$	-0.0002528930	-0.0003829831	-0.0003829512
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	-0.0016402922	-0.0024840709	-0.0024838637
$Var(\hat{\beta}_{\text{FDP}})$	0.0000000345	0.0000000196	0.0000000199
$MSE(\hat{\beta}_{\text{FDP}})$	0.0000000985	0.0000001663	0.0000001665
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1317802848	-0.1319280960	-0.1319293203
$Bias(\hat{\beta}_{\text{would not vote}})$	-0.0001913763	-0.0003391875	-0.0003404118
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	-0.0014543495	-0.0025776298	-0.0025869339
$Var(\hat{\beta}_{\text{would not vote}})$	0.0000000312	0.0000000152	0.0000000151
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0000000679	0.0000001302	0.0000001310

Table C.12: Expected value, variance, bias and MSE of regression coefficient estimates in order to quantify measurement error

Categories	p-value (true)	p-value ($RCAR$)	p-Value (RAR)	p-Value ($RNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.8936605867	0.8925557327	0.8925722817
hochschulreife	0.6987555776	0.7030149574	0.7044247131	0.7044035283
mittlere reife	0.0115117133	0.0116925703	0.0117473348	0.0117466619
no graduation	0.0058831082	0.0058681772	0.0058591353	0.0058587681
volks-, hauptschule	0.0065408724	0.0065298936	0.0065042684	0.0065051724
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0002602571	0.0002648488	0.0002648990
married living apart	0.0000000011	0.0000000015	0.0000000012	0.0000000012
married living together	0.0021189162	0.0022467306	0.0022843726	0.0022845720
widowed	0.8397563199	0.8315338577	0.8287999411	0.8287935089
CDU-CSU	0.5884698278	0.5841258142	0.5825241427	0.5825031425
die gruenen	0.0762638735	0.0759810973	0.0759706929	0.0759742993
die linke	0.1299242680	0.1306585777	0.1311333785	0.1311393296
FDP	0.0082560314	0.0081343814	0.0080783306	0.0080782921
extreme right-wing	0.8910480181	0.8950340999	0.8969465540	0.8969425007
SPD	0.7770122356	0.7743750082	0.7731740612	0.7731629743
would not vote	0.0103321856	0.0102019246	0.0101142810	0.0101135198

Table C.13: Comparison of p-values of regression coefficients of data containing measurement error

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0081995563	7.0019378414	7.0008817670
$Bias(\hat{\beta}_0)$	0.0011387285	-0.0051229864	-0.0061790609
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0001625116	-0.0007311177	-0.0008818335
$Var(\hat{\beta}_0)$	0.0064756808	0.0054974618	0.0044389459
$MSE(\hat{\beta}_0)$	0.0064769775	0.0055237068	0.0044771267
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0113709927	0.0113461843	0.0112812518
$Bias(\hat{\beta}_{\text{age}})$	-0.0000011025	-0.0000259108	-0.0000908433
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.0000969439	-0.0022784523	-0.0079882682
$Var(\hat{\beta}_{\text{age}})$	0.0000002507	0.0000002293	0.0000002327
$MSE(\hat{\beta}_{\text{age}})$	0.0000002507	0.0000002300	0.0000002409
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.3794787168	-0.3787471535	-0.3779256473
$Bias(\hat{\beta}_{\text{female}})$	-0.0011746375	-0.0004430742	0.0003784320
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	-0.0031050089	-0.0011712119	0.0010003380
$Var(\hat{\beta}_{\text{female}})$	0.0001151273	0.0001570960	0.0001761813
$MSE(\hat{\beta}_{\text{female}})$	0.0001165071	0.0001572923	0.0001763245
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.2695864866	0.2897015507	0.2895335753

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0016159668	0.0184990972	0.0183311218
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.0059585259	0.0682113931	0.0675920207
$Var(\hat{\beta}_{\text{no graduation}})$	0.0053068899	0.0046190022	0.0047509944
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0053095012	0.0049612188	0.0050870244
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2249441841	-0.2161513518	-0.2172227621
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0010213620	0.0098141944	0.0087427841
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0045199901	0.0434322600	0.0386907836
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0036265945	0.0035737851	0.0032989027
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0036276377	0.0036701035	0.0033753390
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.2096176875	0.2146907912	0.2149819088
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0002344803	0.0048386234	0.0051297410
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.0011173592	0.0230572955	0.0244445462
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0039347857	0.0033175964	0.0031659124
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0039348407	0.0033410087	0.0031922267
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0119741818	-0.0047312402	-0.0020178910
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0006752650	0.0065676766	0.0092810259
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0597636899	0.5812660373	0.8214084594
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0039840894	0.0034939024	0.0034579845
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0039845454	0.0035370368	0.0035441220
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0323420176	0.0376000071	0.0381753413
$Bias(\hat{\beta}_{\text{hochschulreife}})$	0.0001142723	0.0053722618	0.0059475959
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	0.0035457727	0.1666967922	0.1845489299
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0036409305	0.0032117118	0.0029924109
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0036409435	0.0032405730	0.0030277848
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4580362653	-0.4586041192	-0.4578473504
$Bias(\hat{\beta}_{\text{half-time}})$	-0.0005241962	-0.0010920502	-0.0003352814
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	-0.0011457539	-0.0023869320	-0.0007328361
$Var(\hat{\beta}_{\text{half-time}})$	0.0003025670	0.0003149202	0.0003131457
$MSE(\hat{\beta}_{\text{half-time}})$	0.0003028418	0.0003161128	0.0003132581
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9362579212	-0.9309718946	-0.9288870329
$Bias(\hat{\beta}_{\text{part-time}})$	0.0006472946	0.0059333212	0.0080181829
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0006908859	0.0063328938	0.0085581581
$Var(\hat{\beta}_{\text{part-time}})$	0.0006219461	0.0003691899	0.0004748919
$MSE(\hat{\beta}_{\text{part-time}})$	0.0006223650	0.0004043942	0.0005391832
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7993284670	-0.8020121035	-0.7984637338
$Bias(\hat{\beta}_{\text{not employed}})$	-0.0008673471	-0.0035509836	-0.0000026139
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	-0.0010862735	-0.0044472843	-0.0000032737

$Var(\hat{\beta}_{\text{not employed}})$	0.0001523276	0.0001651746	0.0001497775
$MSE(\hat{\beta}_{\text{not employed}})$	0.0001530799	0.0001777841	0.0001497775
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0895602125	0.0931500427	0.0919017457
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0027858980	0.0008039322	-0.0004443649
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.0301680059	0.0087056419	-0.0048119499
$Var(\hat{\beta}_{\text{married living together}})$	0.0003828379	0.0004305779	0.0004057326
$MSE(\hat{\beta}_{\text{married living together}})$	0.0003905991	0.0004312242	0.0004059300
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3708022001	0.3908839383	0.3895756962
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0095841917	0.0104975465	0.0091893044
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0251959373	0.0275970611	0.0241578159
$Var(\hat{\beta}_{\text{married living apart}})$	0.0018670342	0.0015549275	0.0014968220
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0019588909	0.0016651260	0.0015812653
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0146992145	-0.0099389825	-0.0111401483
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0059649343	-0.0012047023	-0.0024058681
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.6829337053	-0.1379280554	-0.2754512147
$Var(\hat{\beta}_{\text{widowed}})$	0.0010062375	0.0009134068	0.0010387319
$MSE(\hat{\beta}_{\text{widowed}})$	0.0010418179	0.0009148581	0.0010445201
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1226603587	0.1313542041	0.1294214353
$Bias(\hat{\beta}_{\text{single}})$	-0.0018296199	0.0068642255	0.0049314567
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.0146969255	0.0551387793	0.0396132822
$Var(\hat{\beta}_{\text{single}})$	0.0004751713	0.0006569137	0.0005644980
$MSE(\hat{\beta}_{\text{single}})$	0.0004785188	0.0007040313	0.0005888173
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0229227370	-0.0281686920	-0.0290806502
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0021181095	-0.0031278455	-0.0040398037
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.0845861772	-0.1249097349	-0.1613285590
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0008838622	0.0006359230	0.0006384744
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0008883486	0.0006457064	0.0006547944
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0111179779	-0.0146596849	-0.0163736394
$Bias(\hat{\beta}_{\text{SPD}})$	0.0022097417	-0.0013319652	-0.0030459197
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.1658004374	-0.0999394697	-0.2285402020
$Var(\hat{\beta}_{\text{SPD}})$	0.0009382316	0.0006576627	0.0006232001
$MSE(\hat{\beta}_{\text{SPD}})$	0.0009431145	0.0006594369	0.0006324777
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0881926195	0.0845359245	0.0838418153
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0011108376	-0.0025458574	-0.0032399666
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.0127562566	-0.0292352467	-0.0372060207
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0008967191	0.0006976425	0.0007078549
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0008979530	0.0007041239	0.0007183523

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0789574135	0.0834896350	0.0817961542
$Bias(\hat{\beta}_{\text{die linke}})$	0.0020804013	0.0066126228	0.0049191419
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	0.0270614225	0.0860156061	0.0639871633
$Var(\hat{\beta}_{\text{die linke}})$	0.0011795031	0.0008204864	0.0008599807
$MSE(\hat{\beta}_{\text{die linke}})$	0.0011838311	0.0008642132	0.0008841787
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0061681585	0.0091739226	0.0076505304
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0012324290	0.0017733352	0.0002499429
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-0.1665312345	0.2396208666	0.0337733860
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0013543188	0.0010179893	0.0011073161
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0013558376	0.0010211341	0.0011073786
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1498549707	-0.1454490427	-0.1474668797
$Bias(\hat{\beta}_{\text{FDP}})$	0.0043206395	0.0087265675	0.0067087305
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.0280241439	0.0566014788	0.0435135657
$Var(\hat{\beta}_{\text{FDP}})$	0.0013007362	0.0010745877	0.0010069921
$MSE(\hat{\beta}_{\text{FDP}})$	0.0013194041	0.0011507407	0.0010519991
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1272123571	-0.1404105203	-0.1453572762
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0043765514	-0.0088216118	-0.0137683676
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.0332592728	-0.0670391742	-0.1046316728
$Var(\hat{\beta}_{\text{would not vote}})$	0.0012538397	0.0010772416	0.0009495598
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0012729939	0.0011550624	0.0011391278

Table C.14: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of ad-hoc methods

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.6445291302	0.6554488778	0.6562783851
hochschulreife	0.6987555776	0.6303954281	0.6187477014	0.6206511203
mittlere reife	0.0115117133	0.0698781808	0.0626901130	0.0599561792
no graduation	0.0058831082	0.0482099834	0.0334830984	0.0344701548
volks-, hauptschule	0.0065408724	0.0569530744	0.0638656609	0.0584476609
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0095848912	0.0064235815	0.0059004752
married living apart	0.0000000011	0.0000533622	0.0000258073	0.0000208887
married living together	0.0021189162	0.0296002826	0.0249147367	0.0250424959
widowed	0.8397563199	0.6275834745	0.6469005211	0.6296633355
CDU-CSU	0.5884698278	0.6251904460	0.5821696423	0.5678385077
die gruenen	0.0762638735	0.1843375322	0.1823569994	0.1858294288
die linke	0.1299242680	0.2524824519	0.2088113320	0.2193862114
FDP	0.0082560314	0.0581553261	0.0580694110	0.0525880985
extreme right-wing	0.8910480181	0.6687090111	0.6979252222	0.6786652989
SPD	0.7770122356	0.6938616739	0.6807508161	0.6799991566
would not vote	0.0103321856	0.0751394752	0.0408673281	0.0312987438

Table C.15: Comparison of p-values of regression coefficients of data after applying Ad-Hoc-method

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0057853420	6.9931510922	6.9915674169
$Bias(\hat{\beta}_0)$	-0.0012754858	-0.0139097357	-0.0154934110
$Bias_{\mathrm{rel}}(\hat{\beta}_0)$	-0.0001820287	-0.0019851027	-0.0022111141
$Var(\hat{\beta}_0)$	0.0065647616	0.0067587870	0.0053946008
$MSE(\hat{\beta}_0)$	0.0065663884	0.0069522677	0.0056346465
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\mathrm{age}})$	0.0113762177	0.0113572996	0.0112944010
$Bias(\hat{\beta}_{\mathrm{age}})$	-0.0000041226	-0.0000147955	-0.0000776941
$Bias_{\mathrm{rel}}(\hat{\beta}_{\mathrm{age}})$	0.0003625195	-0.0013010398	-0.0068319958
$Var(\hat{\beta}_{\mathrm{age}})$	0.0000002576	0.0000003057	0.0000003032
$MSE(\hat{\beta}_{\mathrm{age}})$	0.0000002576	0.0000003059	0.0000003092
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\mathrm{female}})$	-0.3795536141	-0.3779141134	-0.3769325825
$Bias(\hat{\beta}_{\mathrm{female}})$	-0.0012495348	0.0003899658	0.0013714968
$Bias_{\mathrm{rel}}(\hat{\beta}_{\mathrm{female}})$	-0.0033029905	0.0010308264	0.0036253821
$Var(\hat{\beta}_{\mathrm{female}})$	0.0001131290	0.0001792657	0.0002048872
$MSE(\hat{\beta}_{\mathrm{female}})$	0.0001146904	0.0001794178	0.0002067682
$\beta_{\mathrm{no\ graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\mathrm{no\ graduation}})$	0.2709065301	0.2927205845	0.2924085169

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0002959233	0.0215181310	0.0212060635
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.0010911529	0.0793434230	0.0781927419
$Var(\hat{\beta}_{\text{no graduation}})$	0.0047008557	0.0057206290	0.0056515917
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0047009433	0.0061836590	0.0061012889
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2231618974	-0.2146172012	-0.2161925306
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0028036488	0.0113483449	0.0097730155
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0124074170	0.0502215718	0.0432500249
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0032866132	0.0043936704	0.0042182380
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0032944737	0.0045224553	0.0043137499
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.2115621276	0.2201239075	0.2201507355
$Bias(\hat{\beta}_{\text{mittlere reife}})$	0.0017099598	0.0102717397	0.0102985677
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	0.0081484019	0.0489475035	0.0490753459
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0035019225	0.0042565626	0.0040364663
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0035048464	0.0043620712	0.0041425268
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0098676490	-0.0020180259	0.0004142535
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	0.0014312679	0.0092808909	0.0117131704
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	0.1266730184	0.8213965143	1.0366631188
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0035136676	0.0043029875	0.0042715973
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0035157161	0.0043891224	0.0044087957
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0342294864	0.0420630800	0.0423885005
$Bias(\hat{\beta}_{\text{hochschulreife}})$	0.0020017410	0.0098353346	0.0101607552
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	0.0621123497	0.3051822124	0.3152797392
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0032755085	0.0040755453	0.0038429074
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0032795155	0.0041722792	0.0039461484
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4581721329	-0.4567754569	-0.4565432897
$Bias(\hat{\beta}_{\text{half-time}})$	-0.0006600638	0.0007366121	0.0009687794
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	-0.0014427244	0.0016100387	0.0021174947
$Var(\hat{\beta}_{\text{half-time}})$	0.0002954338	0.0004096300	0.0004195491
$MSE(\hat{\beta}_{\text{half-time}})$	0.0002958695	0.0004101726	0.0004204876
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9366520792	-0.9311068627	-0.9295477489
$Bias(\hat{\beta}_{\text{part-time}})$	0.0002531366	0.0057983531	0.0073574669
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0002701838	0.0061888364	0.0078529469
$Var(\hat{\beta}_{\text{part-time}})$	0.0006362979	0.0004432856	0.0005807795
$MSE(\hat{\beta}_{\text{part-time}})$	0.0006363620	0.0004769065	0.0006349118
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7991471478	-0.7984506651	-0.7947150817
$Bias(\hat{\beta}_{\text{not employed}})$	-0.0006860279	0.0000104548	0.0037460381
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	-0.0008591877	0.0000130937	0.0046915724

$Var(\hat{\beta}_{\text{not employed}})$	0.0001552199	0.0001932457	0.0001737632
$MSE(\hat{\beta}_{\text{not employed}})$	0.0001556905	0.0001932458	0.0001877960
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0896059009	0.0915921453	0.0907677049
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0027402097	-0.0007539652	-0.0015784057
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.0296732548	-0.0081645588	-0.0170922809
$Var(\hat{\beta}_{\text{married living together}})$	0.0003918660	0.0005518262	0.0004969672
$MSE(\hat{\beta}_{\text{married living together}})$	0.0003993748	0.0005523947	0.0004994585
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3709924769	0.3900583221	0.3888872968
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0093939149	0.0096719303	0.0085009050
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0246957176	0.0254265939	0.0223480787
$Var(\hat{\beta}_{\text{married living apart}})$	0.0017798980	0.0016533198	0.0014864898
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0018681436	0.0017468660	0.0015587552
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0146119331	-0.0066119262	-0.0079358866
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0058776529	0.0021223540	0.0007983936
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.6729407299	0.2429912849	0.0914091990
$Var(\hat{\beta}_{\text{widowed}})$	0.0010452791	0.0011471575	0.0012696086
$MSE(\hat{\beta}_{\text{widowed}})$	0.0010798259	0.0011516619	0.0012702460
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1232113751	0.1262409271	0.1246388629
$Bias(\hat{\beta}_{\text{single}})$	-0.0012786036	0.0017509485	0.0001488843
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.0102707348	0.0140649752	0.0011959537
$Var(\hat{\beta}_{\text{single}})$	0.0004805417	0.0008211014	0.0006738044
$MSE(\hat{\beta}_{\text{single}})$	0.0004821765	0.0008241672	0.0006738265
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0226876789	-0.0227008792	-0.0230829463
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0023531676	0.0023399673	0.0019579002
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.0939731637	0.0934460145	0.0781882602
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0008650205	0.0008579751	0.0008452580
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0008705579	0.0008634506	0.0008490913
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0109561801	-0.0086302910	-0.0104691108
$Bias(\hat{\beta}_{\text{SPD}})$	0.0023715395	0.0046974286	0.0028586088
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.1779403814	0.3524555417	0.2144859643
$Var(\hat{\beta}_{\text{SPD}})$	0.0009452663	0.0008096609	0.0008156126
$MSE(\hat{\beta}_{\text{SPD}})$	0.0009508905	0.0008317268	0.0008237843
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0879929205	0.0877341580	0.0876076677
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0009111386	0.0006523761	0.0005258858
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.0104630219	0.0074915331	0.0060389877
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0009196651	0.0008182131	0.0008566458
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0009204953	0.0008186387	0.0008569223

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0787886738	0.0820954908	0.0806667722
$Bias(\hat{\beta}_{\text{die linke}})$	0.0019116616	0.0052184786	0.0037897600
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	0.0248664914	0.0678808711	0.0492964003
$Var(\hat{\beta}_{\text{die linke}})$	0.0011795057	0.0010354550	0.0011080306
$MSE(\hat{\beta}_{\text{die linke}})$	0.0011831602	0.0010626875	0.0011223928
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0052295234	0.0078583917	0.0064161182
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0021710641	0.0004578043	-0.0009844693
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-0.2933637469	0.0618605327	-0.1330258296
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0013412333	0.0012979108	0.0014165283
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0013459468	0.0012981204	0.0014174975
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1496984451	-0.1490770849	-0.1510528343
$Bias(\hat{\beta}_{\text{FDP}})$	0.0044771651	0.0050985253	0.0031227759
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.0290393865	0.0330695969	0.0202546688
$Var(\hat{\beta}_{\text{FDP}})$	0.0012737080	0.0012618450	0.0012017686
$MSE(\hat{\beta}_{\text{FDP}})$	0.0012937530	0.0012878399	0.0012115204
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1270221107	-0.1282189568	-0.1339405286
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0045667978	0.0033699518	-0.0023516200
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.0347050361	0.0256096946	-0.0178709593
$Var(\hat{\beta}_{\text{would not vote}})$	0.0012475757	0.0013878276	0.0011829998
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0012684314	0.0013991842	0.0011885299

Table C.16: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of weighting (weighting variables: gender, education, willingness)

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.6628130938	0.6265360772	0.6273936118
hochschulreife	0.6987555776	0.6490401106	0.5893711959	0.5893751235
mittlere reife	0.0115117133	0.0654566475	0.0653789597	0.0640529538
no graduation	0.0058831082	0.0464346358	0.0346369720	0.0358644735
volks-, hauptschule	0.0065408724	0.0562988515	0.0773878113	0.0711195063
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0090253643	0.0106423775	0.0099881229
married living apart	0.0000000011	0.0000339231	0.0000129958	0.0000096964
married living together	0.0021189162	0.0296786812	0.0319305361	0.0303792919
widowed	0.8397563199	0.6297938847	0.6297093380	0.5996434606
CDU-CSU	0.5884698278	0.6301511197	0.6268472622	0.6215203554
die gruenen	0.0762638735	0.1861385254	0.1806647273	0.1827968136
die linke	0.1299242680	0.2552582761	0.2338623258	0.2449733688
FDP	0.0082560314	0.0583311162	0.0583445423	0.0531175566
extreme right-wing	0.8910480181	0.6738404756	0.6797352299	0.6604988714
SPD	0.7770122356	0.6931576780	0.6896856278	0.6922530015
would not vote	0.0103321856	0.0752282818	0.0716199238	0.0532133519

Table C.17: Comparison of p-values of regression coefficients of data after applying weighting (weighting variables: gender, education, willingness)

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0079091604	7.0149466076	7.0507168759
$Bias(\hat{\beta}_0)$	0.0008483325	0.0078857797	0.0436560480
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0001210682	0.0011254048	0.0062302939
$Var(\hat{\beta}_0)$	0.0176946852	0.0194463769	0.0294004475
$MSE(\hat{\beta}_0)$	0.0176954049	0.0195085624	0.0313062980
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0079325310	0.0078346622	0.0075269152
$Bias(\hat{\beta}_{\text{age}})$	-0.0034395641	-0.0035374329	-0.0038451799
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.3024565013	-0.3110625521	-0.3381241450
$Var(\hat{\beta}_{\text{age}})$	0.0000008141	0.0000007479	0.0000011306
$MSE(\hat{\beta}_{\text{age}})$	0.0000126447	0.0000132614	0.0000159160
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.2591333668	-0.2515740557	-0.2486061153
$Bias(\hat{\beta}_{\text{female}})$	0.1191707125	0.1267300236	0.1296979640
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	0.3150130252	0.3349951284	0.3428405115
$Var(\hat{\beta}_{\text{female}})$	0.0006321753	0.0005698467	0.0005775539
$MSE(\hat{\beta}_{\text{female}})$	0.0148338340	0.0166303456	0.0173991157
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.2041262533	0.1985733102	0.1800071633

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0670762001	-0.0726291433	-0.0911952902
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.2473288840	-0.2678041528	-0.3362627773
$Var(\hat{\beta}_{\text{no graduation}})$	0.0131634344	0.0215871152	0.0249151620
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0176626510	0.0268621077	0.0332317430
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.1359889283	-0.0976197920	-0.1168516529
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0899766178	0.1283457541	0.1091138932
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.3981873314	0.5679881571	0.4828784526
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0117095715	0.0132977714	0.0164938284
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0198053632	0.0297704040	0.0283996701
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.1611680133	0.1666653160	0.1443079182
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0486841545	-0.0431868517	-0.0655442495
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.2319926212	-0.2057965481	-0.3123353464
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0116600068	0.0131535902	0.0154196768
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0140301537	0.0150186944	0.0197157254
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	0.0053407359	0.0079178542	-0.0068642310
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	0.0166396528	0.0192167711	0.0044346859
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	1.4726767998	1.7007622310	0.3924877010
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0132994359	0.0142789071	0.0160025402
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0135763139	0.0146481914	0.0160222067
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0426735159	0.0332458088	0.0124240767
$Bias(\hat{\beta}_{\text{hochschulreife}})$	0.0104457706	0.0010180634	-0.0198036686
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	0.3241235291	0.0315896574	-0.6144912844
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0116431803	0.0127664919	0.0156421517
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0117522944	0.0127675284	0.0160343370
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.3197371974	-0.3357684843	-0.3278696036
$Bias(\hat{\beta}_{\text{half-time}})$	0.1377748717	0.1217435848	0.1296424654
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.3011393163	0.2660991764	0.2833640338
$Var(\hat{\beta}_{\text{half-time}})$	0.0012089522	0.0013104430	0.0016315176
$MSE(\hat{\beta}_{\text{half-time}})$	0.0201908675	0.0161319434	0.0184386864
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.6421666437	-0.6385922985	-0.6412814155
$Bias(\hat{\beta}_{\text{part-time}})$	0.2947385721	0.2983129173	0.2956238004
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.3145873960	0.3184024513	0.3155322389
$Var(\hat{\beta}_{\text{part-time}})$	0.0029684210	0.0017916466	0.0023296296
$MSE(\hat{\beta}_{\text{part-time}})$	0.0898392469	0.0907822433	0.0897230609
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.5484727666	-0.5535223015	-0.5499253918
$Bias(\hat{\beta}_{\text{not employed}})$	0.2499883533	0.2449388184	0.2485357280
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.3130876972	0.3067636136	0.3112684160

$Var(\hat{\beta}_{\text{not employed}})$	0.0006740632	0.0005606559	0.0006292088
$MSE(\hat{\beta}_{\text{not employed}})$	0.0631682399	0.0605556807	0.0623992169
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0614862549	0.0684973558	0.0683043797
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0308598556	-0.0238487547	-0.0240417308
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.3341760195	-0.2582540249	-0.2603437297
$Var(\hat{\beta}_{\text{married living together}})$	0.0014443448	0.0017124588	0.0018217687
$MSE(\hat{\beta}_{\text{married living together}})$	0.0023966755	0.0022812219	0.0023997736
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.2502177982	0.2866516142	0.2732822246
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.1301685935	-0.0937347775	-0.1071041672
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.3422009734	-0.2464199023	-0.2815667687
$Var(\hat{\beta}_{\text{married living apart}})$	0.0078986687	0.0060099148	0.0067889362
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0248425314	0.0147961233	0.0182602388
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0141979077	0.0093201879	0.0269289788
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0054636275	0.0180544681	0.0356632590
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.6255383829	2.0670813981	4.0831365854
$Var(\hat{\beta}_{\text{widowed}})$	0.0029692292	0.0042783261	0.0045825278
$MSE(\hat{\beta}_{\text{widowed}})$	0.0029990804	0.0046042900	0.0058543959
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.0845893029	0.0899864555	0.0853501105
$Bias(\hat{\beta}_{\text{single}})$	-0.0399006757	-0.0345035231	-0.0391398681
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.3205131539	-0.2771590413	-0.3144017580
$Var(\hat{\beta}_{\text{single}})$	0.0017877686	0.0024682028	0.0024892996
$MSE(\hat{\beta}_{\text{single}})$	0.0033798325	0.0036586959	0.0040212289
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0219175949	-0.0118842296	-0.0250669767
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0031232515	0.0131566168	-0.0000261302
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.1247262766	0.5254062335	-0.0010435021
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0033253085	0.0034311031	0.0038890756
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0033350632	0.0036041996	0.0038890763
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0136609925	-0.0108006495	-0.0175932914
$Bias(\hat{\beta}_{\text{SPD}})$	-0.0003332728	0.0025270702	-0.0042655718
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	-0.0250059897	0.1896100950	-0.3200526330
$Var(\hat{\beta}_{\text{SPD}})$	0.0031839195	0.0031321851	0.0037099411
$MSE(\hat{\beta}_{\text{SPD}})$	0.0031840306	0.0031385712	0.0037281362
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0563627654	0.0645259994	0.0558080719
$Bias(\hat{\beta}_{\text{die gruenen}})$	-0.0307190165	-0.0225557825	-0.0312737100
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	-0.3527605415	-0.2590183847	-0.3591303412
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0035334211	0.0041609600	0.0037587510
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0044770791	0.0046697234	0.0047367960

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0418090430	0.0661425030	0.0573229780
$Bias(\hat{\beta}_{\text{die linke}})$	-0.0350679692	-0.0107345092	-0.0195540342
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	-0.4561567649	-0.1396322377	-0.2543547627
$Var(\hat{\beta}_{\text{die linke}})$	0.0038953183	0.0049451755	0.0047566340
$MSE(\hat{\beta}_{\text{die linke}})$	0.0051250808	0.0050604052	0.0051389942
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	-0.0004936485	-0.0029691947	-0.0124448989
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0078942360	-0.0103697821	-0.0198454864
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-1.0667039657	-1.4012106698	-2.6816095990
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0035815393	0.0048259429	0.0050288343
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0036438582	0.0049334753	0.0054226776
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1063717987	-0.1015080148	-0.1142014888
$Bias(\hat{\beta}_{\text{FDP}})$	0.0478038115	0.0526675954	0.0399741215
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.3100607899	0.3416078287	0.2592765574
$Var(\hat{\beta}_{\text{FDP}})$	0.0056797511	0.0073388433	0.0045968928
$MSE(\hat{\beta}_{\text{FDP}})$	0.0079649554	0.0101127189	0.0061948231
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.0915003317	-0.0880047470	-0.1027010327
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0400885768	0.0435841616	0.0288878759
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.3046501203	0.3312145532	0.2195312372
$Var(\hat{\beta}_{\text{would not vote}})$	0.0045893355	0.0045644900	0.0053588812
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0061964295	0.0064640692	0.0061933906

Table C.18: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of single imputation (random imputation)

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000151	0.0000000021	0.0000002102
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.5589734669	0.5319106327	0.5376397931
hochschulreife	0.6987555776	0.5512264683	0.5308565575	0.5023756288
mittlere reife	0.0115117133	0.2954909112	0.2945165731	0.3562208723
no graduation	0.0058831082	0.2662710266	0.2648009435	0.3228959409
volks-, hauptschule	0.0065408724	0.3315332461	0.4199307174	0.3719831710
half-time	0.0000000000	0.0000000034	0.0000000021	0.0000000007
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.2001905706	0.2059830417	0.2110769909
married living apart	0.0000000011	0.0608715150	0.0157956356	0.0317432884
married living together	0.0021189162	0.2939261257	0.2525118834	0.2501057366
widowed	0.8397563199	0.5338076119	0.4964576419	0.4900424133
CDU-CSU	0.5884698278	0.5421398763	0.5553650681	0.4921492236
die gruenen	0.0762638735	0.4662242380	0.4177351953	0.4522021390
die linke	0.1299242680	0.5026731479	0.4112355027	0.4702838623
FDP	0.0082560314	0.3047518820	0.3420028910	0.2741642317
extreme right-wing	0.8910480181	0.5849721906	0.5673165819	0.5491120005
SPD	0.7770122356	0.5817985845	0.5730759447	0.5477687711
would not vote	0.0103321856	0.3409417395	0.3347758117	0.2828691201

Table C.19: Comparison of p-values of regression coefficients of data after applying single imputation (random imputation)

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0916369070	7.1001636991	7.0955100453
$Bias(\hat{\beta}_0)$	0.0845760791	0.0931028712	0.0884492175
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0120701220	0.0132870077	0.0126228699
$Var(\hat{\beta}_0)$	0.0054594981	0.0077466948	0.0065872425
$MSE(\hat{\beta}_0)$	0.0126126112	0.0164148395	0.0144105065
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0078482392	0.0077676642	0.0077232479
$Bias(\hat{\beta}_{\text{age}})$	-0.0035238559	-0.0036044309	-0.0036488472
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.3098686623	-0.3169539877	-0.3208597164
$Var(\hat{\beta}_{\text{age}})$	0.0000003152	0.0000002671	0.0000002588
$MSE(\hat{\beta}_{\text{age}})$	0.0000127328	0.0000132591	0.0000135729
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.2619632223	-0.2339841705	-0.2344401869
$Bias(\hat{\beta}_{\text{female}})$	0.1163408570	0.1443199088	0.1438638924
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	0.3075326525	0.3814918122	0.3802863893
$Var(\hat{\beta}_{\text{female}})$	0.0001211282	0.0003151214	0.0003088420
$MSE(\hat{\beta}_{\text{female}})$	0.0136563232	0.0211433575	0.0210056615
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.1915116433	0.1982282710	0.1994620402

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0796908102	-0.0729741825	-0.0717404133
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.2938425120	-0.2690764098	-0.2645271545
$Var(\hat{\beta}_{\text{no graduation}})$	0.0047131969	0.0059308767	0.0052811917
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0110638221	0.0112561080	0.0104278786
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.1503682301	-0.0545398883	-0.0571272534
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0755973160	0.1714256578	0.1688382927
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.3345524010	0.7586362644	0.7471860006
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0031562267	0.0062893322	0.0057083319
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0088711809	0.0356760883	0.0342147009
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.1493249843	0.1426126369	0.1428180040
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0605271834	-0.0672395308	-0.0670341638
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.2884277254	-0.3204138015	-0.3194351742
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0031946339	0.0052511315	0.0046176226
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0068581739	0.0097722860	0.0091112018
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0051701984	-0.0100109886	-0.0075955023
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	0.0061287185	0.0012879282	0.0037034146
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	0.5424164602	0.1139868764	0.3277672205
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0031395911	0.0049240515	0.0045151911
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0031771523	0.0049257103	0.0045289063
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0276762420	0.0171925909	0.0179904895
$Bias(\hat{\beta}_{\text{hochschulreife}})$	-0.0045515034	-0.0150351545	-0.0142372559
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	-0.1412293454	-0.4665282756	-0.4417701485
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0030535896	0.0048005163	0.0041836912
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0030743058	0.0050265722	0.0043863906
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.3154235979	-0.3392545596	-0.3372660380
$Bias(\hat{\beta}_{\text{half-time}})$	0.1420884711	0.1182575094	0.1202460311
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.3105677003	0.2584795406	0.2628259214
$Var(\hat{\beta}_{\text{half-time}})$	0.0002075695	0.0003199617	0.0002922067
$MSE(\hat{\beta}_{\text{half-time}})$	0.0203967032	0.0143048002	0.0147513147
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.6438839013	-0.6443465394	-0.6457184072
$Bias(\hat{\beta}_{\text{part-time}})$	0.2930213145	0.2925586765	0.2911868086
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.3127544917	0.3122606978	0.3107964431
$Var(\hat{\beta}_{\text{part-time}})$	0.0011465790	0.0011896584	0.0011516977
$MSE(\hat{\beta}_{\text{part-time}})$	0.0870080698	0.0867802376	0.0859414552
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.5504355624	-0.5591140542	-0.5576560739
$Bias(\hat{\beta}_{\text{not employed}})$	0.2480255575	0.2393470657	0.2408050460
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.3106294737	0.2997604513	0.3015864392

$Var(\hat{\beta}_{\text{not employed}})$	0.0001288135	0.0001807492	0.0001760937
$MSE(\hat{\beta}_{\text{not employed}})$	0.0616454906	0.0574677671	0.0581631639
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0619353427	0.0685831442	0.0670299892
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0304107678	-0.0237629663	-0.0253161213
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.3293129253	-0.2573250370	-0.2741438828
$Var(\hat{\beta}_{\text{married living together}})$	0.0004903544	0.0004124792	0.0003580178
$MSE(\hat{\beta}_{\text{married living together}})$	0.0014151692	0.0009771578	0.0009989238
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.2563606678	0.2815109917	0.2806793047
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.1240257239	-0.0988754000	-0.0997070871
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.3260519477	-0.2599341148	-0.2621205417
$Var(\hat{\beta}_{\text{married living apart}})$	0.0017439917	0.0013075731	0.0013272677
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0171263719	0.0110839178	0.0112687709
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0094457299	0.0224692893	0.0213696895
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0007114497	0.0312035695	0.0301039697
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.0814548729	3.5725404762	3.4466457488
$Var(\hat{\beta}_{\text{widowed}})$	0.0011676968	0.0009417863	0.0009171123
$MSE(\hat{\beta}_{\text{widowed}})$	0.0011682030	0.0019154491	0.0018233613
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.0856785500	0.0872207027	0.0853986034
$Bias(\hat{\beta}_{\text{single}})$	-0.0388114287	-0.0372692760	-0.0390913752
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.3117634775	-0.2993757118	-0.3140122253
$Var(\hat{\beta}_{\text{single}})$	0.0006845503	0.0006056384	0.0005656167
$MSE(\hat{\beta}_{\text{single}})$	0.0021908773	0.0019946373	0.0020937523
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0122580716	-0.0115961874	-0.0140827234
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0127827749	0.0134446591	0.0109581230
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.5104769465	0.5369091298	0.4376099288
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0006006097	0.0007371277	0.0007097073
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0007640091	0.0009178866	0.0008297878
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0043832393	-0.0115167450	-0.0130923744
$Bias(\hat{\beta}_{\text{SPD}})$	0.0089444804	0.0018109747	0.0002353452
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.6711185876	0.1358803097	0.0176583265
$Var(\hat{\beta}_{\text{SPD}})$	0.0008197159	0.0007230128	0.0006687580
$MSE(\hat{\beta}_{\text{SPD}})$	0.0008997196	0.0007262924	0.0006688134
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0654742647	0.0629766232	0.0617840610
$Bias(\hat{\beta}_{\text{die gruenen}})$	-0.0216075172	-0.0241051587	-0.0252977209
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	-0.2481290202	-0.2768105812	-0.2905053196
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0007943764	0.0008109716	0.0008419022
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0012612612	0.0013920302	0.0014818768

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0562145152	0.0651355885	0.0623257607
$Bias(\hat{\beta}_{\text{die linke}})$	-0.0206624970	-0.0117414237	-0.0145512515
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	-0.2687734142	-0.1527299693	-0.1892796182
$Var(\hat{\beta}_{\text{die linke}})$	0.0010660424	0.0010415293	0.0009414680
$MSE(\hat{\beta}_{\text{die linke}})$	0.0014929812	0.0011793903	0.0011532070
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0078749123	0.0028155209	0.0004005251
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	0.0004743249	-0.0045850666	-0.0070000624
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	0.0640928672	-0.6195544104	-0.9458792843
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0010753537	0.0009660003	0.0009418810
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0010755787	0.0009870231	0.0009908819
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1006390216	-0.1051570072	-0.1071460863
$Bias(\hat{\beta}_{\text{FDP}})$	0.0535365886	0.0490186030	0.0470295239
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.3472442141	0.3179400614	0.3050386756
$Var(\hat{\beta}_{\text{FDP}})$	0.0012651733	0.0014500104	0.0012769865
$MSE(\hat{\beta}_{\text{FDP}})$	0.0041313396	0.0038528338	0.0034887627
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.0838594143	-0.0892299719	-0.0937393737
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0477294942	0.0423589367	0.0378495348
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.3627166968	0.3219035490	0.2876346894
$Var(\hat{\beta}_{\text{would not vote}})$	0.0013716638	0.0011243847	0.0011097192
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0036497684	0.0029186642	0.0025423065

Table C.20: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of single imputation (mean imputation)

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.6161782378	0.5570837368	0.5811524383
hochschulreife	0.6987555776	0.5718562273	0.5435633583	0.5587214976
mittlere reife	0.0115117133	0.1375629590	0.1829518173	0.1701233600
no graduation	0.0058831082	0.1132670462	0.1190364275	0.1075621702
volks-, hauptschule	0.0065408724	0.1253318724	0.4754763144	0.4789033809
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0481189936	0.0369086329	0.0381008661
married living apart	0.0000000011	0.0006289558	0.0001738168	0.0001383354
married living together	0.0021189162	0.0946385613	0.0620802913	0.0623668348
widowed	0.8397563199	0.5657039981	0.5142883575	0.5256782552
CDU-CSU	0.5884698278	0.6449102239	0.6315266488	0.6249251443
die gruenen	0.0762638735	0.2440754312	0.2697124228	0.2802151233
die linke	0.1299242680	0.3296128605	0.2766576101	0.2895697650
FDP	0.0082560314	0.1397141571	0.1329125596	0.1199749997
extreme right-wing	0.8910480181	0.6323018550	0.6548695565	0.6599262446
SPD	0.7770122356	0.6334676974	0.6526662188	0.6559587817
would not vote	0.0103321856	0.1821107197	0.1466938452	0.1246360307

Table C.21: Comparison of p-values of regression coefficients of data after applying single imputation (mean imputation)

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	6.9174937780	6.8667165945	6.8694878168
$Bias(\hat{\beta}_0)$	-0.0895670499	-0.1403442334	-0.1375730111
$Bias_{\text{rel}}(\hat{\beta}_0)$	-0.0127823994	-0.0200289732	-0.0196334832
$Var(\hat{\beta}_0)$	0.0148186904	0.0127187798	0.0143960762
$MSE(\hat{\beta}_0)$	0.0228409468	0.0324152837	0.0333224096
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0128148553	0.0136551293	0.0134839992
$Bias(\hat{\beta}_{\text{age}})$	0.0014427602	0.0022830342	0.0021119041
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	0.1268684624	0.2007575694	0.1857093215
$Var(\hat{\beta}_{\text{age}})$	0.0000004051	0.0000004267	0.0000004557
$MSE(\hat{\beta}_{\text{age}})$	0.0000024866	0.0000056390	0.0000049158
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.4185481911	-0.4336206341	-0.4307167670
$Bias(\hat{\beta}_{\text{female}})$	-0.0402441118	-0.0553165548	-0.0524126878
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	-0.1063803274	-0.1462224645	-0.1385464515
$Var(\hat{\beta}_{\text{female}})$	0.0001812870	0.0003264131	0.0002946685
$MSE(\hat{\beta}_{\text{female}})$	0.0018008755	0.0033863343	0.0030417583
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.3285529909	0.3688765028	0.3700097038

$Bias(\hat{\beta}_{\text{no graduation}})$	0.0573505374	0.0976740493	0.0988072503
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	0.2114676201	0.3601517907	0.3643302228
$Var(\hat{\beta}_{\text{no graduation}})$	0.0116811792	0.0119122165	0.0142889414
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0149702633	0.0214524364	0.0240518141
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2197814102	-0.2149688899	-0.2101629293
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0061841359	0.0109966562	0.0158026169
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0273676056	0.0486651898	0.0699337449
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0095650371	0.0104890529	0.0115388316
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0096032806	0.0106099793	0.0117885543
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.2657630531	0.2688739597	0.2703469967
$Bias(\hat{\beta}_{\text{mittlere reife}})$	0.0559108853	0.0590217920	0.0604948290
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	0.2664298678	0.2812541448	0.2882735482
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0098718323	0.0096965323	0.0118469708
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0129978594	0.0131801042	0.0155065951
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	0.0063823563	0.0131245830	0.0189483735
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	0.0176812732	0.0244234999	0.0302472903
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	1.5648644331	2.1615788610	2.6770079521
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0098469136	0.0096170372	0.0119383498
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0101595410	0.0102135445	0.0128532483
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0752421860	0.0813653972	0.0840635838
$Bias(\hat{\beta}_{\text{hochschulreife}})$	0.0430144406	0.0491376518	0.0518358384
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	1.3347021383	1.5247002627	1.6084227370
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0095302676	0.0097461013	0.0118446625
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0113805097	0.0121606102	0.0145316166
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4527266137	-0.4568787688	-0.4565167217
$Bias(\hat{\beta}_{\text{half-time}})$	0.0047854553	0.0006333003	0.0009953473
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.0104597357	0.0013842264	0.0021755652
$Var(\hat{\beta}_{\text{half-time}})$	0.0003959847	0.0004215359	0.0004250038
$MSE(\hat{\beta}_{\text{half-time}})$	0.0004188853	0.0004219370	0.0004259945
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9456496746	-0.9504569490	-0.9482769057
$Bias(\hat{\beta}_{\text{part-time}})$	-0.0087444588	-0.0135517332	-0.0113716899
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	-0.0093333441	-0.0144643588	-0.0121375030
$Var(\hat{\beta}_{\text{part-time}})$	0.0008628088	0.0006519396	0.0006957363
$MSE(\hat{\beta}_{\text{part-time}})$	0.0009392744	0.0008355891	0.0008250517
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.8222905931	-0.8507521089	-0.8470741776
$Bias(\hat{\beta}_{\text{not employed}})$	-0.0238294733	-0.0522909890	-0.0486130577
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	-0.0298442500	-0.0654897123	-0.0608834375

$Var(\hat{\beta}_{\text{not employed}})$	0.0002060908	0.0002086444	0.0001990145
$MSE(\hat{\beta}_{\text{not employed}})$	0.0007739346	0.0029429920	0.0025622438
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.1197587361	0.1366606357	0.1341215575
$Bias(\hat{\beta}_{\text{married living together}})$	0.0274126255	0.0443145252	0.0417754470
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	0.2968465633	0.4798743000	0.4523790635
$Var(\hat{\beta}_{\text{married living together}})$	0.0006126933	0.0008067530	0.0008049296
$MSE(\hat{\beta}_{\text{married living together}})$	0.0013641453	0.0027705302	0.0025501175
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.4147830230	0.4618571167	0.4589009891
$Bias(\hat{\beta}_{\text{married living apart}})$	0.0343966312	0.0814707250	0.0785145974
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	0.0904255041	0.2141788632	0.2064074821
$Var(\hat{\beta}_{\text{married living apart}})$	0.0023670686	0.0020648256	0.0019197398
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0035501969	0.0087023046	0.0080842818
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0016621413	0.0054700428	0.0043346813
$Bias(\hat{\beta}_{\text{widowed}})$	0.0070721389	0.0142043230	0.0130689615
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	0.8096991093	1.6262728804	1.4962837507
$Var(\hat{\beta}_{\text{widowed}})$	0.0016779374	0.0019574283	0.0018997222
$MSE(\hat{\beta}_{\text{widowed}})$	0.0017279526	0.0021591911	0.0020705199
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1612721931	0.1911036717	0.1875867603
$Bias(\hat{\beta}_{\text{single}})$	0.0367822144	0.0666136930	0.0630967816
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	0.2954632561	0.5350928144	0.5068422562
$Var(\hat{\beta}_{\text{single}})$	0.0008392990	0.0011122343	0.0010600854
$MSE(\hat{\beta}_{\text{single}})$	0.0021922303	0.0055496184	0.0050412892
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0150763041	-0.0276500753	-0.0292530088
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0099645423	-0.0026092288	-0.0042121623
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.3979315295	-0.1041989048	-0.1682116598
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0012484707	0.0009369331	0.0009210290
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0013477628	0.0009437412	0.0009387713
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0034397439	-0.0113950220	-0.0123505366
$Bias(\hat{\beta}_{\text{SPD}})$	0.0098879758	0.0019326976	0.0009771831
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.7419105474	0.1450133778	0.0733196014
$Var(\hat{\beta}_{\text{SPD}})$	0.0014432855	0.0010530442	0.0010033010
$MSE(\hat{\beta}_{\text{SPD}})$	0.0015410575	0.0010567795	0.0010042559
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.1061627704	0.1019690850	0.1007561729
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0190809885	0.0148873031	0.0136743910
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.2191157332	0.1709577224	0.1570292966
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0012505375	0.0011905797	0.0011336250
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0016146216	0.0014122115	0.0013206140

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0993693344	0.1072459347	0.1052577625
$Bias(\hat{\beta}_{\text{die linke}})$	0.0224923222	0.0303689225	0.0283807503
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	0.2925753947	0.3950325537	0.3691708282
$Var(\hat{\beta}_{\text{die linke}})$	0.0017985402	0.0012517383	0.0012559278
$MSE(\hat{\beta}_{\text{die linke}})$	0.0023044448	0.0021740097	0.0020613948
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0151881461	0.0187848461	0.0162848435
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	0.0077875586	0.0113842587	0.0088842561
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	1.0522892450	1.5382912119	1.2004798396
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0017086314	0.0013582304	0.0014149696
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0017692775	0.0014878318	0.0014938996
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1642862562	-0.1587615575	-0.1623867071
$Bias(\hat{\beta}_{\text{FDP}})$	-0.0101106460	-0.0045859473	-0.0082110968
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	-0.0655787643	-0.0297449594	-0.0532580791
$Var(\hat{\beta}_{\text{FDP}})$	0.0023180681	0.0020388320	0.0021844570
$MSE(\hat{\beta}_{\text{FDP}})$	0.0024202933	0.0020598629	0.0022518792
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1279730239	-0.1536533128	-0.1587981495
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0036158846	-0.0220644042	-0.0272092409
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.0274786428	-0.1676767781	-0.2067745771
$Var(\hat{\beta}_{\text{would not vote}})$	0.0016551290	0.0015962070	0.0014305641
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0016682036	0.0020830450	0.0021709069

Table C.22: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of single imputation (deterministic regression imputation)

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.4226237873	0.4221740554	0.4546937984
hochschulreife	0.6987555776	0.4110095082	0.3755749819	0.3729700561
mittlere reife	0.0115117133	0.0116743841	0.0188229177	0.0188860940
no graduation	0.0058831082	0.0067637241	0.0062963946	0.0081791465
volks-, hauptschule	0.0065408724	0.0651977419	0.0837397955	0.0763236244
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0004187104	0.0000043959	0.0000020202
married living apart	0.0000000011	0.0000000056	0.0000000001	0.0000000001
married living together	0.0021189162	0.0006941404	0.0000844594	0.0001408856
widowed	0.8397563199	0.4738770185	0.4799905063	0.4635606240
CDU-CSU	0.5884698278	0.5353602905	0.4728381180	0.4543794018
die gruenen	0.0762638735	0.0623533347	0.0611655859	0.0622420341
die linke	0.1299242680	0.0973795749	0.0596744141	0.0637489214
FDP	0.0082560314	0.0164168074	0.0216775264	0.0210066029
extreme right-wing	0.8910480181	0.5151358984	0.5718777707	0.5688330435
SPD	0.7770122356	0.5572434153	0.5585187752	0.5780114339
would not vote	0.0103321856	0.0387381707	0.0122286226	0.0079922933

Table C.23: Comparison of p-values of regression coefficients of data after applying single imputation (deterministic regression imputation)

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	6.9924669484	7.0049230861	7.0014843795
$Bias(\hat{\beta}_0)$	-0.0145938795	-0.0021377418	-0.0055764484
$Bias_{\text{rel}}(\hat{\beta}_0)$	-0.0020827391	-0.0003050840	-0.0007958327
$Var(\hat{\beta}_0)$	0.0158732165	0.0139817396	0.0124702533
$MSE(\hat{\beta}_0)$	0.0160861978	0.0139863095	0.0125013501
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0113188148	0.0113286479	0.0112105894
$Bias(\hat{\beta}_{\text{age}})$	-0.0000532803	-0.0000434472	-0.0001615057
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.0046851813	-0.0038205116	-0.0142019330
$Var(\hat{\beta}_{\text{age}})$	0.0000005893	0.0000007770	0.0000006649
$MSE(\hat{\beta}_{\text{age}})$	0.0000005922	0.0000007788	0.0000006910
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.3703613127	-0.3689728951	-0.3672654350
$Bias(\hat{\beta}_{\text{female}})$	0.0079427666	0.0093311842	0.0110386443
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	0.0209957202	0.0246658302	0.0291792896
$Var(\hat{\beta}_{\text{female}})$	0.0004188609	0.0003741238	0.0003924366
$MSE(\hat{\beta}_{\text{female}})$	0.0004819484	0.0004611948	0.0005142883
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.2650360611	0.2759969684	0.2797034142

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0061663924	0.0047945149	0.0085009607
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.0227372294	0.0176787299	0.0313454418
$Var(\hat{\beta}_{\text{no graduation}})$	0.0156203545	0.0098634925	0.0101005677
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0156583789	0.0098864799	0.0101728340
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2086815722	-0.1974003455	-0.1898886982
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0172839740	0.0285652006	0.0360768479
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0764894217	0.1264139650	0.1596564101
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0093085654	0.0071663851	0.0072587364
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0096073012	0.0079823558	0.0085602753
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.2169000164	0.1995132025	0.2058591081
$Bias(\hat{\beta}_{\text{mittlere reife}})$	0.0070478487	-0.0103389653	-0.0039930596
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	0.0335848266	-0.0492678507	-0.0190279647
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0098049371	0.0074466952	0.0077073874
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0098546093	0.0075535894	0.0077233320
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0161393806	-0.0233642927	-0.0131602402
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0048404637	-0.0120653758	-0.0018613233
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.4284006817	-1.0678347233	-0.1647346693
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0102855414	0.0069456463	0.0074951694
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0103089715	0.0070912196	0.0074986339
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0431358094	0.0252421536	0.0324632961
$Bias(\hat{\beta}_{\text{hochschulreife}})$	0.0109080640	-0.0069855917	0.0002355507
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	0.3384681082	-0.2167570726	0.0073089427
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0098601629	0.0069235578	0.0072353435
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0099791488	0.0069723563	0.0072353990
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4676966667	-0.4691279944	-0.4693772845
$Bias(\hat{\beta}_{\text{half-time}})$	-0.0101845977	-0.0116159253	-0.0118652154
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	-0.0222608284	-0.0253893309	-0.0259342130
$Var(\hat{\beta}_{\text{half-time}})$	0.0008828500	0.0008748458	0.0008458249
$MSE(\hat{\beta}_{\text{half-time}})$	0.0009865760	0.0010097755	0.0009866082
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9110910208	-0.9025778178	-0.9002725498
$Bias(\hat{\beta}_{\text{part-time}})$	0.0258141950	0.0343273980	0.0366326661
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0275526217	0.0366391364	0.0390996500
$Var(\hat{\beta}_{\text{part-time}})$	0.0013966292	0.0015786659	0.0018079508
$MSE(\hat{\beta}_{\text{part-time}})$	0.0020630019	0.0027570362	0.0031499030
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7924577023	-0.8011304583	-0.7985572706
$Bias(\hat{\beta}_{\text{not employed}})$	0.0060034175	-0.0026693384	-0.0000961507
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.0075187350	-0.0033431038	-0.0001204201

$Var(\hat{\beta}_{\text{not employed}})$	0.0004048379	0.0004261165	0.0004092040
$MSE(\hat{\beta}_{\text{not employed}})$	0.0004408790	0.0004332419	0.0004092133
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.1007691812	0.1105200537	0.1086877763
$Bias(\hat{\beta}_{\text{married living together}})$	0.0084230706	0.0181739432	0.0163416657
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	0.0912119696	0.1968024760	0.1769610615
$Var(\hat{\beta}_{\text{married living together}})$	0.0011384559	0.0010728279	0.0010385665
$MSE(\hat{\beta}_{\text{married living together}})$	0.0012094040	0.0014031201	0.0013056165
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3743968450	0.4037653753	0.4018423171
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0059895468	0.0233789835	0.0214559254
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0157459545	0.0614611458	0.0564056071
$Var(\hat{\beta}_{\text{married living apart}})$	0.0047300397	0.0038490158	0.0039881477
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0047659144	0.0043955927	0.0044485045
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0070793382	0.0084118229	0.0047610645
$Bias(\hat{\beta}_{\text{widowed}})$	0.0016549420	0.0171461031	0.0134953447
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	0.1894766314	1.9630814122	1.5451009561
$Var(\hat{\beta}_{\text{widowed}})$	0.0034309817	0.0028740638	0.0028599179
$MSE(\hat{\beta}_{\text{widowed}})$	0.0034337205	0.0031680527	0.0030420422
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1351255706	0.1460764270	0.1433908702
$Bias(\hat{\beta}_{\text{single}})$	0.0106355919	0.0215864484	0.0189008916
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	0.0854333181	0.1733990852	0.1518266114
$Var(\hat{\beta}_{\text{single}})$	0.0014463874	0.0017664979	0.0015213023
$MSE(\hat{\beta}_{\text{single}})$	0.0015595032	0.0022324727	0.0018785460
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0116623010	-0.0220447104	-0.0251116744
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0133785455	0.0029961360	-0.0000708279
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.5342688979	0.1196499507	-0.0028284954
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0026162784	0.0017208177	0.0016823296
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0027952639	0.0017297946	0.0016823346
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0056884557	-0.0140624267	-0.0174610966
$Bias(\hat{\beta}_{\text{SPD}})$	0.0076392639	-0.0007347071	-0.0041333769
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.5731861226	-0.0551262388	-0.3101338441
$Var(\hat{\beta}_{\text{SPD}})$	0.0023880839	0.0019604526	0.0019357178
$MSE(\hat{\beta}_{\text{SPD}})$	0.0024464422	0.0019609924	0.0019528026
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0949947999	0.0898268628	0.0882972790
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0079130180	0.0027450809	0.0012154971
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.0908688113	0.0315230214	0.0139581101
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0024147683	0.0017788079	0.0017979293
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0024773841	0.0017863433	0.0017994067

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0923304132	0.0939683186	0.0912436215
$Bias(\hat{\beta}_{\text{die linke}})$	0.0154534010	0.0170913064	0.0143666092
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	0.2010145885	0.2223201178	0.1868778300
$Var(\hat{\beta}_{\text{die linke}})$	0.0032338845	0.0025796105	0.0024643090
$MSE(\hat{\beta}_{\text{die linke}})$	0.0034726921	0.0028717232	0.0026707085
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0133899445	0.0087813303	0.0051564181
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	0.0059893570	0.0013807428	-0.0022441694
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	0.8093083253	0.1865720536	-0.3032420549
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0030938072	0.0027880167	0.0024630031
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0031296796	0.0027899232	0.0024680394
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1334978570	-0.1277133363	-0.1305586961
$Bias(\hat{\beta}_{\text{FDP}})$	0.0206777533	0.0264622739	0.0236169141
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.1341181865	0.1716372250	0.1531819079
$Var(\hat{\beta}_{\text{FDP}})$	0.0042775195	0.0030329900	0.0029541386
$MSE(\hat{\beta}_{\text{FDP}})$	0.0047050890	0.0037332420	0.0035118972
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1106927186	-0.1315588054	-0.1357813289
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0208961900	0.0000301032	-0.0041924204
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.1587990221	0.0002287667	-0.0318599831
$Var(\hat{\beta}_{\text{would not vote}})$	0.0030221607	0.0027353053	0.0020504382
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0034588115	0.0027353062	0.0020680146

Table C.24: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of single imputation (stochastic regression imputation)

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.4920189787	0.5591207507	0.5706634684
hochschulreife	0.6987555776	0.5201901479	0.5449572572	0.5449636729
mittlere reife	0.0115117133	0.1030286977	0.1325360975	0.1149679774
no graduation	0.0058831082	0.1038775302	0.0784402486	0.0720871047
volks-, hauptschule	0.0065408724	0.1301237346	0.1285667186	0.1257073255
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0210063888	0.0117042983	0.0094436602
married living apart	0.0000000011	0.0003359741	0.0000988969	0.0000457599
married living together	0.0021189162	0.0436741858	0.0262694311	0.0213384083
widowed	0.8397563199	0.5076716817	0.4973691588	0.5069096466
CDU-CSU	0.5884698278	0.4942816767	0.5282857918	0.5268545533
die gruenen	0.0762638735	0.2007899367	0.2119679992	0.2128950398
die linke	0.1299242680	0.2328135646	0.2309966270	0.2394614802
FDP	0.0082560314	0.1516394072	0.1571872850	0.1392307206
extreme right-wing	0.8910480181	0.5267447909	0.5817940980	0.5892774586
SPD	0.7770122356	0.5438360966	0.5618724223	0.5582356997
would not vote	0.0103321856	0.1709116795	0.1088359625	0.0777123613

Table C.25: Comparison of p-values of regression coefficients of data after applying single imputation (stochastic regression imputation)

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0494776542	7.0408830685	7.0450210944
$Bias(\hat{\beta}_0)$	0.0424168263	0.0338222406	0.0379602666
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0060534406	0.0048268798	0.0054174307
$Var(\hat{\beta}_0)$	0.0071095715	0.0126071385	0.0088393449
$MSE(\hat{\beta}_0)$	0.0089087587	0.0137510825	0.0102803267
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0091326469	0.0092390049	0.0091646857
$Bias(\hat{\beta}_{\text{age}})$	-0.0022394482	-0.0021330902	-0.0022074094
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.1969248566	-0.1875723125	-0.1941075412
$Var(\hat{\beta}_{\text{age}})$	0.0000007705	0.0000007909	0.0000009550
$MSE(\hat{\beta}_{\text{age}})$	0.0000057856	0.0000053410	0.0000058276
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.4247887735	-0.4061528801	-0.4061373292
$Bias(\hat{\beta}_{\text{female}})$	-0.0464846942	-0.0278488008	-0.0278332499
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	-0.1228765344	-0.0736148574	-0.0735737503
$Var(\hat{\beta}_{\text{female}})$	0.0005218836	0.0004676553	0.0005292038
$MSE(\hat{\beta}_{\text{female}})$	0.0026827104	0.0012432110	0.0013038936
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.1677342398	0.2042129735	0.2055898701

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.1034682137	-0.0669894800	-0.0656125834
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.3815165106	-0.2470091222	-0.2419321159
$Var(\hat{\beta}_{\text{no graduation}})$	0.0070747121	0.0099471654	0.0081157308
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0177803833	0.0144347558	0.0124207419
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.1716589631	-0.0934276054	-0.0996740891
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0543065831	0.1325379407	0.1262914570
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.2403312540	0.5865404836	0.5588969610
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0043512814	0.0080698468	0.0062832830
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0073004864	0.0256361525	0.0222328151
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.1329857827	0.1634524289	0.1623932981
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0768663851	-0.0463997389	-0.0474588696
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.3662882585	-0.2211067885	-0.2261538212
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0040331967	0.0081575312	0.0058592252
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0099416379	0.0103104670	0.0081115695
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0180976061	0.0004326025	-0.0001943028
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0067986892	0.0117315194	0.0111046140
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.6017115851	1.0382870752	0.9828034099
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0040748725	0.0083011426	0.0064044045
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0041210947	0.0084387712	0.0065277170
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0049741602	0.0261944947	0.0246253683
$Bias(\hat{\beta}_{\text{hochschulreife}})$	-0.0272535852	-0.0060332507	-0.0076023771
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	-0.8456559669	-0.1872067249	-0.2358954057
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0040406640	0.0077068282	0.0052007154
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0047834219	0.0077432283	0.0052585115
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.3157945608	-0.3347498916	-0.3300330091
$Bias(\hat{\beta}_{\text{half-time}})$	0.1417175082	0.1227621774	0.1274790600
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.3097568737	0.2683255497	0.2786354035
$Var(\hat{\beta}_{\text{half-time}})$	0.0006306537	0.0005848341	0.0005972894
$MSE(\hat{\beta}_{\text{half-time}})$	0.0207145058	0.0156553863	0.0168482001
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.6415573143	-0.6478827692	-0.6492496255
$Bias(\hat{\beta}_{\text{part-time}})$	0.2953479016	0.2890224466	0.2876555903
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.3152377600	0.3084863247	0.3070274191
$Var(\hat{\beta}_{\text{part-time}})$	0.0014864460	0.0012319582	0.0013315287
$MSE(\hat{\beta}_{\text{part-time}})$	0.0887168289	0.0847659328	0.0840772673
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.5498451329	-0.5616477939	-0.5566890113
$Bias(\hat{\beta}_{\text{not employed}})$	0.2486159870	0.2368133260	0.2417721086
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.3113689330	0.2965871726	0.3027975972

$Var(\hat{\beta}_{\text{not employed}})$	0.0002539160	0.0004182574	0.0003278248
$MSE(\hat{\beta}_{\text{not employed}})$	0.0620638250	0.0564988088	0.0587815773
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0642833402	0.0669095609	0.0631583282
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0280627703	-0.0254365496	-0.0291877823
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.3038868683	-0.2754479799	-0.3160694282
$Var(\hat{\beta}_{\text{married living together}})$	0.0008624842	0.0009214818	0.0010482174
$MSE(\hat{\beta}_{\text{married living together}})$	0.0016500033	0.0015684998	0.0019001440
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.2549339005	0.2762374027	0.2782826229
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.1254524912	-0.1041489890	-0.1021037689
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.3298027846	-0.2737978837	-0.2684211926
$Var(\hat{\beta}_{\text{married living apart}})$	0.0039316996	0.0039979985	0.0033955234
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0196700271	0.0148450104	0.0138207031
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0211557442	-0.0080528667	-0.0107990250
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0124214640	0.0006814135	-0.0020647448
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-1.4221508476	0.0780159878	-0.2363955297
$Var(\hat{\beta}_{\text{widowed}})$	0.0018637022	0.0024182237	0.0022245076
$MSE(\hat{\beta}_{\text{widowed}})$	0.0020179950	0.0024186881	0.0022287708
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.0783119073	0.0811967945	0.0782538131
$Bias(\hat{\beta}_{\text{single}})$	-0.0461780713	-0.0432931841	-0.0462361655
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.3709380611	-0.3477644113	-0.3714047190
$Var(\hat{\beta}_{\text{single}})$	0.0010762664	0.0013103352	0.0009633723
$MSE(\hat{\beta}_{\text{single}})$	0.0032086806	0.0031846350	0.0031011553
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0130530538	-0.0190651568	-0.0271278137
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0119877927	0.0059756896	-0.0020869672
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.4787295293	0.2386376854	-0.0833425182
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0019144041	0.0012918184	0.0011524239
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0020581113	0.0013275273	0.0011567793
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0044864923	-0.0126224498	-0.0218395277
$Bias(\hat{\beta}_{\text{SPD}})$	0.0088412273	0.0007052699	-0.0085118080
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.6633713475	0.0529175196	-0.6386544938
$Var(\hat{\beta}_{\text{SPD}})$	0.0020549853	0.0013517016	0.0013859872
$MSE(\hat{\beta}_{\text{SPD}})$	0.0021331526	0.0013521990	0.0014584381
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0660676941	0.0583156080	0.0517312225
$Bias(\hat{\beta}_{\text{die gruenen}})$	-0.0210140878	-0.0287661740	-0.0353505594
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	-0.2413143978	-0.3303351553	-0.4059466701
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0021848640	0.0014978830	0.0016470591
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0026264558	0.0023253758	0.0028967212

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0578004861	0.0562560561	0.0505793133
$Bias(\hat{\beta}_{\text{die linke}})$	-0.0190765261	-0.0206209561	-0.0262976989
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	-0.2481434377	-0.2682330581	-0.3420749346
$Var(\hat{\beta}_{\text{die linke}})$	0.0024766654	0.0017369308	0.0021951226
$MSE(\hat{\beta}_{\text{die linke}})$	0.0028405793	0.0021621546	0.0028866916
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0078413132	-0.0034247502	-0.0097730668
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	0.0004407257	-0.0108253377	-0.0171736542
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	0.0595527980	-1.4627673472	-2.3205798603
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0030602009	0.0018926953	0.0017081045
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0030603951	0.0020098832	0.0020030389
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1019620002	-0.1079193662	-0.1128967409
$Bias(\hat{\beta}_{\text{FDP}})$	0.0522136100	0.0462562441	0.0412788693
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.3386632292	0.3000230970	0.2677392959
$Var(\hat{\beta}_{\text{FDP}})$	0.0031629442	0.0026723944	0.0022523103
$MSE(\hat{\beta}_{\text{FDP}})$	0.0058892053	0.0048120345	0.0039562553
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.0787161741	-0.0998042047	-0.1065442648
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0528727344	0.0317847038	0.0250446437
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.4018023634	0.2415454628	0.1903248840
$Var(\hat{\beta}_{\text{would not vote}})$	0.0025935667	0.0020418090	0.0021692650
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0053890928	0.0030520764	0.0027964992

Table C.26: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of single imputation (predictive mean matching imputation)

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.6400005982	0.5399199138	0.6086731009
hochschulreife	0.6987555776	0.6349616699	0.4925989037	0.5654655681
mittlere reife	0.0115117133	0.2486846538	0.2170570885	0.1824569755
no graduation	0.0058831082	0.2287891427	0.1774745246	0.1546078998
volks-, hauptschule	0.0065408724	0.1475836995	0.4055208345	0.4010211847
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.1286878923	0.1316494789	0.1175462553
married living apart	0.0000000011	0.0071902190	0.0062479159	0.0028858066
married living together	0.0021189162	0.1667238664	0.1524504781	0.1756666698
widowed	0.8397563199	0.5123104258	0.5045034245	0.5182500918
CDU-CSU	0.5884698278	0.5567962846	0.6057837903	0.5689031509
die gruenen	0.0762638735	0.3510551149	0.3868637834	0.4046235260
die linke	0.1299242680	0.3989921776	0.4019088634	0.4196312988
FDP	0.0082560314	0.2337141191	0.2025330333	0.1659040058
extreme right-wing	0.8910480181	0.5623540037	0.6411866054	0.6509667624
SPD	0.7770122356	0.5409498943	0.6133663123	0.5924279154
would not vote	0.0103321856	0.2959809154	0.1897439868	0.1595271700

Table C.27: Comparison of p-values of regression coefficients of data after applying single imputation (predictive mean matching imputation)

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0304375975	7.0227453235	7.0201799753
$Bias(\hat{\beta}_0)$	0.0233767696	0.0156844956	0.0131191474
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0033361734	0.0022383844	0.0018722754
$Var(\hat{\beta}_0)$	0.0066085489	0.0076386731	0.0064853808
$MSE(\hat{\beta}_0)$	0.0071550222	0.0078846765	0.0066574928
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0091982955	0.0094539807	0.0093667623
$Bias(\hat{\beta}_{\text{age}})$	-0.0021737996	-0.0019181144	-0.0020053328
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.1911520812	-0.1686685154	-0.1763380266
$Var(\hat{\beta}_{\text{age}})$	0.0000004960	0.0000004294	0.0000004959
$MSE(\hat{\beta}_{\text{age}})$	0.0000052214	0.0000041086	0.0000045173
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.4223858971	-0.4069108182	-0.4049339175
$Bias(\hat{\beta}_{\text{female}})$	-0.0440818179	-0.0286067389	-0.0266298382
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	-0.1165248282	-0.0756183728	-0.0703926807
$Var(\hat{\beta}_{\text{female}})$	0.0003597579	0.0003495922	0.0003410396
$MSE(\hat{\beta}_{\text{female}})$	0.0023029646	0.0011679377	0.0010501879
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.1857595460	0.1976396053	0.1986974346

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0854429075	-0.0735628481	-0.0725050189
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.3150521183	-0.2712469861	-0.2673464711
$Var(\hat{\beta}_{\text{no graduation}})$	0.0054086424	0.0054753861	0.0047737040
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0127091329	0.0108868787	0.0100306818
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.1513945408	-0.0907309803	-0.0934995985
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0745710053	0.1352345659	0.1324659476
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.3300105108	0.5984742726	0.5862218816
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0036496485	0.0043421428	0.0040007527
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0092104833	0.0226305306	0.0215479800
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.1469955552	0.1648891310	0.1636506371
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0628566126	-0.0449630368	-0.0462015307
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.2995280595	-0.2142605303	-0.2201622750
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0036630024	0.0041877563	0.0037351461
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0076139562	0.0062094310	0.0058697275
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0059493570	0.0010892655	0.0026308631
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	0.0053495598	0.0123881824	0.0139297800
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	0.4734577563	1.0964044176	1.2328420658
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0037095116	0.0042564512	0.0041181875
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0037381294	0.0044099182	0.0043122262
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0245303942	0.0290398316	0.0286231517
$Bias(\hat{\beta}_{\text{hochschulreife}})$	-0.0076973511	-0.0031879137	-0.0036045937
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	-0.2388423716	-0.0989182987	-0.1118475292
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0037212123	0.0039453541	0.0032355358
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0037804616	0.0039555169	0.0032485289
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.3175143890	-0.3331350775	-0.3290284637
$Bias(\hat{\beta}_{\text{half-time}})$	0.1399976800	0.1243769916	0.1284836053
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.3059977856	0.2718551051	0.2808310731
$Var(\hat{\beta}_{\text{half-time}})$	0.0003187605	0.0003371107	0.0002903866
$MSE(\hat{\beta}_{\text{half-time}})$	0.0199181109	0.0158067467	0.0167984234
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.6446238265	-0.6484716261	-0.6462384951
$Bias(\hat{\beta}_{\text{part-time}})$	0.2922813893	0.2884335898	0.2906667208
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.3119647371	0.3078578120	0.3102413306
$Var(\hat{\beta}_{\text{part-time}})$	0.0010485426	0.0009126054	0.0009077884
$MSE(\hat{\beta}_{\text{part-time}})$	0.0864769532	0.0841065411	0.0853949310
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.5503924963	-0.5634440975	-0.5586716201
$Bias(\hat{\beta}_{\text{not employed}})$	0.2480686236	0.2350170224	0.2397894998
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.3106834101	0.2943374656	0.3003145599

$Var(\hat{\beta}_{\text{not employed}})$	0.0001525664	0.0002294034	0.0002200869
$MSE(\hat{\beta}_{\text{not employed}})$	0.0616906084	0.0554624043	0.0577190912
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0617954945	0.0680940216	0.0666510199
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0305506160	-0.0242520889	-0.0256950906
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.3308273171	-0.2626216610	-0.2782476754
$Var(\hat{\beta}_{\text{married living together}})$	0.0005368237	0.0005530569	0.0005633311
$MSE(\hat{\beta}_{\text{married living together}})$	0.0014701638	0.0011412208	0.0012235688
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.2559657478	0.2753782859	0.2749596514
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.1244206440	-0.1050081059	-0.1054267403
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.3270901553	-0.2760564209	-0.2771569714
$Var(\hat{\beta}_{\text{married living apart}})$	0.0024055946	0.0020984026	0.0021298894
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0178860912	0.0131251049	0.0132446870
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0124619103	0.0171887527	0.0138083466
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0037276301	0.0259230329	0.0225426268
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.4267815984	2.9679644244	2.5809369813
$Var(\hat{\beta}_{\text{widowed}})$	0.0014415766	0.0012663882	0.0014277869
$MSE(\hat{\beta}_{\text{widowed}})$	0.0014554718	0.0019383918	0.0019359569
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.0856290608	0.0888376073	0.0867252156
$Bias(\hat{\beta}_{\text{single}})$	-0.0388609179	-0.0356523714	-0.0377647630
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.3121610132	-0.2863874809	-0.3033558478
$Var(\hat{\beta}_{\text{single}})$	0.0007172217	0.0007666266	0.0007149538
$MSE(\hat{\beta}_{\text{single}})$	0.0022273926	0.0020377182	0.0021411312
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0144995422	-0.0187811008	-0.0226619711
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0105413043	0.0062597457	0.0023788754
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.4209643740	0.2499813943	0.0949997997
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0009226194	0.0010736507	0.0007813625
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0010337385	0.0011128351	0.0007870216
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0067656866	-0.0151146102	-0.0167876313
$Bias(\hat{\beta}_{\text{SPD}})$	0.0065620331	-0.0017868905	-0.0034599116
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.4923597754	-0.1340732373	-0.2596026726
$Var(\hat{\beta}_{\text{SPD}})$	0.0010713243	0.0010374051	0.0007809085
$MSE(\hat{\beta}_{\text{SPD}})$	0.0011143846	0.0010405981	0.0007928795
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0625971002	0.0566515794	0.0553477211
$Bias(\hat{\beta}_{\text{die gruenen}})$	-0.0244846817	-0.0304302025	-0.0317340608
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	-0.2811688183	-0.3494439571	-0.3644167603
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0012776388	0.0013218623	0.0012089592
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0018771384	0.0022478595	0.0022160098

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0580046007	0.0568721251	0.0545660796
$Bias(\hat{\beta}_{\text{die linke}})$	-0.0188724116	-0.0200048872	-0.0223109327
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	-0.2454883590	-0.2602193633	-0.2902159180
$Var(\hat{\beta}_{\text{die linke}})$	0.0014182819	0.0013710877	0.0011999339
$MSE(\hat{\beta}_{\text{die linke}})$	0.0017744498	0.0017712832	0.0016977116
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0062185058	-0.0016995261	-0.0040513574
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0011820816	-0.0091001135	-0.0114519449
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-0.1597280829	-1.2296474538	-1.5474372699
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0017571573	0.0014264336	0.0011825196
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0017585546	0.0015092457	0.0013136666
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1018881194	-0.1083880902	-0.1105198424
$Bias(\hat{\beta}_{\text{FDP}})$	0.0522874908	0.0457875200	0.0436557678
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.3391424284	0.2969829011	0.2831561215
$Var(\hat{\beta}_{\text{FDP}})$	0.0015734462	0.0020665268	0.0013821526
$MSE(\hat{\beta}_{\text{FDP}})$	0.0043074279	0.0041630238	0.0032879786
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.0856286759	-0.0961600628	-0.0998539300
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0459602326	0.0354288458	0.0317349786
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.3492713262	0.2692388453	0.2411675796
$Var(\hat{\beta}_{\text{would not vote}})$	0.0012186681	0.0017384579	0.0013342601
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0033310111	0.0029936610	0.0023413689

Table C.28: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of multiple imputation (multiple regression imputation)

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.5994647325	0.5801884234	0.5737944285
hochschulreife	0.6987555776	0.5783644362	0.5437774614	0.5616631821
mittlere reife	0.0115117133	0.2219338895	0.1858094533	0.1837016251
no graduation	0.0058831082	0.2027647458	0.1830303517	0.1760417760
volks-, hauptschule	0.0065408724	0.2177620097	0.4196402886	0.4078887655
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.1019152067	0.0930320665	0.0985495813
married living apart	0.0000000011	0.0090269279	0.0058347665	0.0042860935
married living together	0.0021189162	0.1797767016	0.1354031875	0.1489965740
widowed	0.8397563199	0.5228485069	0.5183474771	0.5222649473
CDU-CSU	0.5884698278	0.5825001070	0.5624122191	0.5700192553
die gruenen	0.0762638735	0.3515163817	0.3856015698	0.3892167913
die linke	0.1299242680	0.3903278862	0.3919376004	0.4078895787
FDP	0.0082560314	0.2350742862	0.2163759545	0.1920589408
extreme right-wing	0.8910480181	0.5656204232	0.5734056150	0.5982574121
SPD	0.7770122356	0.5918735767	0.5831713991	0.5844653145
would not vote	0.0103321856	0.2552678578	0.2230892614	0.1888909768

Table C.29: Comparison of p-values of regression coefficients of data after applying multiple imputation (multiple regression imputation)

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0479753633	7.0527274941	7.0485608950
$Bias(\hat{\beta}_0)$	0.0409145354	0.0456666662	0.0415000671
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0058390438	0.0065172356	0.0059226069
$Var(\hat{\beta}_0)$	0.0047759294	0.0078564731	0.0058959749
$MSE(\hat{\beta}_0)$	0.0064499286	0.0099419175	0.0076182305
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0090959404	0.0091906686	0.0090592364
$Bias(\hat{\beta}_{\text{age}})$	-0.0022761547	-0.0021814265	-0.0023128587
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.2001526236	-0.1918227470	-0.2033801788
$Var(\hat{\beta}_{\text{age}})$	0.0000004427	0.0000005533	0.0000006105
$MSE(\hat{\beta}_{\text{age}})$	0.0000056236	0.0000053120	0.0000059598
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.4295438431	-0.4072829241	-0.4069773886
$Bias(\hat{\beta}_{\text{female}})$	-0.0512397638	-0.0289788448	-0.0286733093
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	-0.1354459722	-0.0766019888	-0.0757943434
$Var(\hat{\beta}_{\text{female}})$	0.0003446942	0.0004484273	0.0004291362
$MSE(\hat{\beta}_{\text{female}})$	0.0029702076	0.0012882007	0.0012512949
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.1860053805	0.1946788602	0.1968830028

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0851970730	-0.0765235933	-0.0743194507
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.3141456571	-0.2821640892	-0.2740367934
$Var(\hat{\beta}_{\text{no graduation}})$	0.0048848602	0.0052069161	0.0046307030
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0121434014	0.0110627764	0.0101540837
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.1562166769	-0.1015131878	-0.1019488777
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0697488692	0.1244523583	0.1240166685
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.3086703720	0.5507581154	0.5488299901
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0030458065	0.0046417908	0.0034915439
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0079107112	0.0201301803	0.0188716780
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.1470232589	0.1568863576	0.1583693239
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0628289089	-0.0529658101	-0.0514828439
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.2993960442	-0.2523958207	-0.2453291021
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0030469519	0.0045953831	0.0032510927
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0069944237	0.0074007602	0.0059015760
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0053367290	-0.0050327385	-0.0011566704
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	0.0059621878	0.0062661784	0.0101422465
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	0.5276778207	0.5545822195	0.8976299769
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0030944070	0.0045683221	0.0032882950
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0031299547	0.0046075871	0.0033911602
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0220616741	0.0194647636	0.0221554458
$Bias(\hat{\beta}_{\text{hochschulreife}})$	-0.0101660713	-0.0127629818	-0.0100722996
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	-0.3154446938	-0.3960246570	-0.3125350369
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0029015399	0.0042336084	0.0027584192
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0030048889	0.0043965021	0.0028598704
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.3128395956	-0.3335735226	-0.3291317217
$Bias(\hat{\beta}_{\text{half-time}})$	0.1446724734	0.1239385464	0.1283803474
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.3162156437	0.2708967803	0.2806053787
$Var(\hat{\beta}_{\text{half-time}})$	0.0003603710	0.0003800793	0.0003399733
$MSE(\hat{\beta}_{\text{half-time}})$	0.0212904955	0.0157408426	0.0168214869
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.6442524444	-0.6483268946	-0.6469725228
$Bias(\hat{\beta}_{\text{part-time}})$	0.2926527714	0.2885783213	0.2899326931
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.3123611295	0.3080122902	0.3094578706
$Var(\hat{\beta}_{\text{part-time}})$	0.0009682370	0.0007665213	0.0008257132
$MSE(\hat{\beta}_{\text{part-time}})$	0.0866138816	0.0840439688	0.0848866797
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.5490472282	-0.5606969251	-0.5586710578
$Bias(\hat{\beta}_{\text{not employed}})$	0.2494138917	0.2377641948	0.2397900621
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.3123682362	0.2977780493	0.3003152641

$Var(\hat{\beta}_{\text{not employed}})$	0.0001735267	0.0002286350	0.0001659435
$MSE(\hat{\beta}_{\text{not employed}})$	0.0623808161	0.0567604473	0.0576652174
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0597840001	0.0632633017	0.0621539683
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0325621104	-0.0290828088	-0.0301921422
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.3526094414	-0.3149326877	-0.3269454666
$Var(\hat{\beta}_{\text{married living together}})$	0.0006447246	0.0005258368	0.0004727026
$MSE(\hat{\beta}_{\text{married living together}})$	0.0017050156	0.0013716466	0.0013842680
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.2593834803	0.2750007781	0.2738188757
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.1210029115	-0.1053856137	-0.1065675161
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.3181052585	-0.2770488534	-0.2801559635
$Var(\hat{\beta}_{\text{married living apart}})$	0.0025382517	0.0017435499	0.0022580535
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0171799563	0.0128496774	0.0136146890
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0257275224	-0.0079821598	-0.0058172456
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0169932422	0.0007521204	0.0029170346
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-1.9455801520	0.0861113197	0.3339753814
$Var(\hat{\beta}_{\text{widowed}})$	0.0015187474	0.0011586957	0.0013120043
$MSE(\hat{\beta}_{\text{widowed}})$	0.0018075177	0.0011592614	0.0013205134
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.0746969216	0.0774915551	0.0742377339
$Bias(\hat{\beta}_{\text{single}})$	-0.0497930570	-0.0469984235	-0.0502522447
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.3999764283	-0.3775277659	-0.4036649796
$Var(\hat{\beta}_{\text{single}})$	0.0008510463	0.0006809461	0.0007014291
$MSE(\hat{\beta}_{\text{single}})$	0.0033303948	0.0028897979	0.0032267172
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0137988214	-0.0198680260	-0.0195260953
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0112420250	0.0051728205	0.0055147512
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.4489474846	0.2065753051	0.2202302212
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0012256805	0.0007424932	0.0007644546
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0013520636	0.0007692512	0.0007948671
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0065945480	-0.0149082343	-0.0150163242
$Bias(\hat{\beta}_{\text{SPD}})$	0.0067331717	-0.0015805146	-0.0016886046
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.5052005768	-0.1185885248	-0.1266986877
$Var(\hat{\beta}_{\text{SPD}})$	0.0013812952	0.0007141519	0.0008271319
$MSE(\hat{\beta}_{\text{SPD}})$	0.0014266308	0.0007166499	0.0008299833
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0636589292	0.0569019466	0.0581024717
$Bias(\hat{\beta}_{\text{die gruenen}})$	-0.0234228527	-0.0301798353	-0.0289793102
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	-0.2689753488	-0.3465688766	-0.3327826969
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0013922381	0.0010271111	0.0011266517
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0019408682	0.0019379336	0.0019664521

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0577643376	0.0574324503	0.0568385977
$Bias(\hat{\beta}_{\text{die linke}})$	-0.0191126746	-0.0194445620	-0.0200384146
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	-0.2486136504	-0.2529307709	-0.2606554805
$Var(\hat{\beta}_{\text{die linke}})$	0.0015760195	0.0010229370	0.0012018279
$MSE(\hat{\beta}_{\text{die linke}})$	0.0019413139	0.0014010280	0.0016033660
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0044177056	-0.0009898385	-0.0007729467
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0029828819	-0.0083904259	-0.0081735341
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-0.4030601447	-1.1337513381	-1.1044439607
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0020324780	0.0011152774	0.0012955244
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0020413755	0.0011856766	0.0013623310
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1010161432	-0.1087642301	-0.1071360934
$Bias(\hat{\beta}_{\text{FDP}})$	0.0531594670	0.0454113801	0.0470395168
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.3447981620	0.2945432163	0.3051034904
$Var(\hat{\beta}_{\text{FDP}})$	0.0018827197	0.0015760649	0.0015003281
$MSE(\hat{\beta}_{\text{FDP}})$	0.0047086486	0.0036382583	0.0037130443
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.0844763004	-0.0981149769	-0.1003528149
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0471126082	0.0334739317	0.0312360937
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.3580287177	0.2543826226	0.2373763413
$Var(\hat{\beta}_{\text{would not vote}})$	0.0015186376	0.0011469008	0.0013551405
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0037382354	0.0022674049	0.0023308341

Table C.30: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of multiple imputation (multiple predictive mean matching imputation)

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.5973631634	0.5749133070	0.5960097816
hochschulreife	0.6987555776	0.5850609355	0.5580515150	0.5813748015
mittlere reife	0.0115117133	0.2282753330	0.2056709823	0.1934122980
no graduation	0.0058831082	0.2010059764	0.1839338006	0.1777433748
volks-, hauptschule	0.0065408724	0.2000747387	0.3814970953	0.3677233650
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.1546002110	0.1323561847	0.1478754958
married living apart	0.0000000011	0.0070647308	0.0033749580	0.0038586424
married living together	0.0021189162	0.1998793216	0.1629709411	0.1659786957
widowed	0.8397563199	0.5001510531	0.5538005062	0.5309622083
CDU-CSU	0.5884698278	0.5547134063	0.5804400734	0.5737841177
die gruenen	0.0762638735	0.3510601111	0.3905773088	0.3729623383
die linke	0.1299242680	0.3936395851	0.3899180032	0.3839137222
FDP	0.0082560314	0.2434722322	0.2006518668	0.2124844894
extreme right-wing	0.8910480181	0.5735225073	0.5991118436	0.6009447252
SPD	0.7770122356	0.5611069360	0.5990477795	0.5865844138
would not vote	0.0103321856	0.2577724352	0.1891104110	0.1916535476

Table C.31: Comparison of p-values of regression coefficients of data after applying multiple imputation (multiple predictive mean matching imputation)

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0082006725	7.0019379895	7.0008819677
$Bias(\hat{\beta}_0)$	0.0011398447	-0.0051228384	-0.0061788601
$Bias_{\mathrm{rel}}(\hat{\beta}_0)$	0.0001626709	-0.0007310966	-0.0008818048
$Var(\hat{\beta}_0)$	0.0064756822	0.0054974366	0.0044389675
$MSE(\hat{\beta}_0)$	0.0064769814	0.0055236800	0.0044771458
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\mathrm{age}})$	0.0113709996	0.0113461831	0.0112812501
$Bias(\hat{\beta}_{\mathrm{age}})$	-0.0000010955	-0.0000259120	-0.0000908451
$Bias_{\mathrm{rel}}(\hat{\beta}_{\mathrm{age}})$	-0.0000963355	-0.0022785640	-0.0079884181
$Var(\hat{\beta}_{\mathrm{age}})$	0.0000002507	0.0000002293	0.0000002327
$MSE(\hat{\beta}_{\mathrm{age}})$	0.0000002507	0.0000002300	0.0000002409
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\mathrm{female}})$	-0.3794783867	-0.3787470020	-0.3779255335
$Bias(\hat{\beta}_{\mathrm{female}})$	-0.0011743074	-0.0004429227	0.0003785458
$Bias_{\mathrm{rel}}(\hat{\beta}_{\mathrm{female}})$	-0.0031041363	-0.0011708112	0.0010006389
$Var(\hat{\beta}_{\mathrm{female}})$	0.0001151285	0.0001570986	0.0001761844
$MSE(\hat{\beta}_{\mathrm{female}})$	0.0001165075	0.0001572948	0.0001763277
$\beta_{\mathrm{no\ graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\mathrm{no\ graduation}})$	0.2695856493	0.2897015242	0.2895336148

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0016168041	0.0184990707	0.0183311613
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.0059616132	0.0682112955	0.0675921661
$Var(\hat{\beta}_{\text{no graduation}})$	0.0053068376	0.0046189885	0.0047510328
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0053094516	0.0049612042	0.0050870643
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2249460156	-0.2161513885	-0.2172227195
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0010195305	0.0098141576	0.0087428266
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0045118848	0.0434320976	0.0386909720
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0036265675	0.0035737742	0.0032989182
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0036276070	0.0036700919	0.0033753553
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.2096165246	0.2146906578	0.2149819032
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0002356432	0.0048384900	0.0051297355
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.0011229007	0.0230566597	0.0244445198
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0039347398	0.0033175769	0.0031659354
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0039347953	0.0033409879	0.0031922496
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0119754636	-0.0047311822	-0.0020179014
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0006765467	0.0065677347	0.0092810155
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0598771283	0.5812711741	0.8214075408
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0039841335	0.0034939148	0.0034580206
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0039845912	0.0035370499	0.0035441578
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0323411840	0.0376000452	0.0381754814
$Bias(\hat{\beta}_{\text{hochschulreife}})$	0.0001134386	0.0053722999	0.0059477361
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	0.0035199047	0.1666979743	0.1845532785
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0036408900	0.0032117008	0.0029924392
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0036409029	0.0032405624	0.0030278148
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4580362468	-0.4586040284	-0.4578473190
$Bias(\hat{\beta}_{\text{half-time}})$	-0.0005241777	-0.0010919594	-0.0003352499
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	-0.0011457135	-0.0023867335	-0.0007327674
$Var(\hat{\beta}_{\text{half-time}})$	0.0003025553	0.0003149230	0.0003131423
$MSE(\hat{\beta}_{\text{half-time}})$	0.0003028301	0.0003161154	0.0003132547
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9362594891	-0.9309719975	-0.9288871775
$Bias(\hat{\beta}_{\text{part-time}})$	0.0006457267	0.0059332183	0.0080180383
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0006892124	0.0063327839	0.0085580037
$Var(\hat{\beta}_{\text{part-time}})$	0.0006219581	0.0003691974	0.0004748899
$MSE(\hat{\beta}_{\text{part-time}})$	0.0006223751	0.0004044005	0.0005391788
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7993285844	-0.8020121437	-0.7984637943
$Bias(\hat{\beta}_{\text{not employed}})$	-0.0008674645	-0.0035510238	-0.0000026744
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	-0.0010864204	-0.0044473346	-0.0000033494

$Var(\hat{\beta}_{\text{not employed}})$	0.0001523255	0.0001651738	0.0001497769
$MSE(\hat{\beta}_{\text{not employed}})$	0.0001530780	0.0001777836	0.0001497769
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0895600627	0.0931499853	0.0919016204
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0027860478	0.0008038747	-0.0004444901
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.0301696279	0.0087050199	-0.0048133066
$Var(\hat{\beta}_{\text{married living together}})$	0.0003828481	0.0004305771	0.0004057323
$MSE(\hat{\beta}_{\text{married living together}})$	0.0003906102	0.0004312233	0.0004059299
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3708024216	0.3908839923	0.3895757877
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0095839701	0.0104976005	0.0091893959
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0251953549	0.0275972032	0.0241580564
$Var(\hat{\beta}_{\text{married living apart}})$	0.0018670211	0.0015549336	0.0014968323
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0019588735	0.0016651332	0.0015812773
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0146995036	-0.0099390145	-0.0111401219
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0059652234	-0.0012047344	-0.0024058417
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.6829667965	-0.1379317277	-0.2754481903
$Var(\hat{\beta}_{\text{widowed}})$	0.0010062054	0.0009134237	0.0010387443
$MSE(\hat{\beta}_{\text{widowed}})$	0.0010417893	0.0009148751	0.0010445324
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1226597222	0.1313541629	0.1294212611
$Bias(\hat{\beta}_{\text{single}})$	-0.0018302565	0.0068641843	0.0049312825
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.0147020385	0.0551384485	0.0396118833
$Var(\hat{\beta}_{\text{single}})$	0.0004751916	0.0006569141	0.0005645000
$MSE(\hat{\beta}_{\text{single}})$	0.0004785414	0.0007040311	0.0005888176
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0229226850	-0.0281686923	-0.0290806615
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0021181615	-0.0031278458	-0.0040398150
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.0845882528	-0.1249097458	-0.1613290108
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0008838660	0.0006359254	0.0006384745
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0008883526	0.0006457088	0.0006547946
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0111177713	-0.0146595850	-0.0163735220
$Bias(\hat{\beta}_{\text{SPD}})$	0.0022099484	-0.0013318653	-0.0030458023
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.1658159416	-0.0999319723	-0.2285313927
$Var(\hat{\beta}_{\text{SPD}})$	0.0009382327	0.0006576646	0.0006231971
$MSE(\hat{\beta}_{\text{SPD}})$	0.0009431166	0.0006594385	0.0006324740
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0881925791	0.0845355867	0.0838415176
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0011107972	-0.0025461952	-0.0032402643
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.0127557933	-0.0292391264	-0.0372094401
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0008967430	0.0006976514	0.0007078592
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0008979769	0.0007041345	0.0007183585

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0789568318	0.0834889137	0.0817953652
$Bias(\hat{\beta}_{\text{die linke}})$	0.0020798196	0.0066119015	0.0049183530
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	0.0270538557	0.0860062235	0.0639769004
$Var(\hat{\beta}_{\text{die linke}})$	0.0011795173	0.0008204973	0.0008599850
$MSE(\hat{\beta}_{\text{die linke}})$	0.0011838430	0.0008642145	0.0008841752
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0061701985	0.0091748804	0.0076514834
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0012303889	0.0017742930	0.0002508959
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-0.1662555741	0.2397502880	0.0339021587
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0013543502	0.0010179899	0.0011073169
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0013558641	0.0010211380	0.0011073799
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1498552484	-0.1454489169	-0.1474668477
$Bias(\hat{\beta}_{\text{FDP}})$	0.0043203618	0.0087266933	0.0067087626
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.0280223430	0.0566022947	0.0435137733
$Var(\hat{\beta}_{\text{FDP}})$	0.0013007446	0.0010745885	0.0010069947
$MSE(\hat{\beta}_{\text{FDP}})$	0.0013194101	0.0011507436	0.0010520022
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1272129708	-0.1404108702	-0.1453577318
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0043759377	-0.0088219616	-0.0137688233
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.0332546093	-0.0670418330	-0.1046351353
$Var(\hat{\beta}_{\text{would not vote}})$	0.0012538460	0.0010772562	0.0009495644
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0012729949	0.0011550832	0.0011391449

Table C.32: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of FIML estimation in Structural Equation Model

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0000000000	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0000000000	0.0000000000	0.0000000000
fachhochschulreife	0.8970973124	0.6431313377	0.6540574572	0.6548852435
hochschulreife	0.6987555776	0.6290300897	0.6173246478	0.6192200303
mittlere reife	0.0115117133	0.0688016396	0.0617183768	0.0590082235
no graduation	0.0058831082	0.0473303319	0.0328208752	0.0338225382
volks-, hauptschule	0.0065408724	0.0560892931	0.0628885620	0.0574913732
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0093147861	0.0062200256	0.0056948058
married living apart	0.0000000011	0.0000500471	0.0000241613	0.0000195033
married living together	0.0021189162	0.0289521266	0.0243283972	0.0244443670
widowed	0.8397563199	0.6262046769	0.6455010993	0.6282393301
CDU-CSU	0.5884698278	0.6237853252	0.5805858172	0.5662050469
die gruenen	0.0762638735	0.1826297074	0.1805680202	0.1840284867
die linke	0.1299242680	0.2506685616	0.2069838301	0.2175458977
FDP	0.0082560314	0.0571466444	0.0570480192	0.0516080365
extreme right-wing	0.8910480181	0.6673976205	0.6966994215	0.6773654621
SPD	0.7770122356	0.6926824343	0.6794546011	0.6787065292
would not vote	0.0103321856	0.0740468315	0.0400409580	0.0305802699

Table C.33: Comparison of p-values of regression coefficients of data after applying FIML estimation in Structural Equation Model

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	7.0494388122	7.0495207785	7.0508538892
$Bias(\hat{\beta}_0)$	0.0423779844	0.0424599506	0.0437930613
$Bias_{\text{rel}}(\hat{\beta}_0)$	0.0060478973	0.0060595950	0.0062498475
$Var(\hat{\beta}_0)$	0.1941571854	0.0080593441	0.0061632909
$MSE(\hat{\beta}_0)$	0.1959530789	0.0098621915	0.0080811231
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0113818574	0.0113588759	0.0112945684
$Bias(\hat{\beta}_{\text{age}})$	0.0000097623	-0.0000132192	-0.0000775267
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	0.0008584451	-0.0011624271	-0.0068172763
$Var(\hat{\beta}_{\text{age}})$	0.0000002537	0.0000002301	0.0000002337
$MSE(\hat{\beta}_{\text{age}})$	0.0000002538	0.0000002302	0.0000002397
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.3795339603	-0.3694406688	-0.3687174745
$Bias(\hat{\beta}_{\text{female}})$	-0.0012298810	0.0088634105	0.0095866048
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	-0.0032510381	0.0234293285	0.0253410029
$Var(\hat{\beta}_{\text{female}})$	0.0005586641	0.0002927325	0.0002814348
$MSE(\hat{\beta}_{\text{female}})$	0.0005601767	0.0003712926	0.0003733378
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.2650379093	0.2899976605	0.2911961245

$Bias(\hat{\beta}_{\text{no graduation}})$	-0.0061645442	0.0187952070	0.0199936710
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	-0.0227304145	0.0693032338	0.0737223086
$Var(\hat{\beta}_{\text{no graduation}})$	0.0181734056	0.0050486988	0.0050542453
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0182114072	0.0054019586	0.0054539921
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2388959845	-0.1867167094	-0.1857569155
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.0129304384	0.0392488367	0.0402086307
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.0572230529	0.1736938989	0.1779414223
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0097122021	0.0043970103	0.0039964727
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0098793983	0.0059374815	0.0056132067
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.1983209491	0.2042737402	0.2051583893
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0115312186	-0.0055784276	-0.0046937785
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.0549492472	-0.0265826540	-0.0223670716
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0110834440	0.0038929472	0.0035189214
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0112164130	0.0039240661	0.0035409530
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0199826086	-0.0106500412	-0.0075365304
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0086836917	0.0006488757	0.0037623865
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.7685419577	0.0574281300	0.3329864721
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0088755446	0.0038470485	0.0036276895
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0089509511	0.0038474696	0.0036418451
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0136365233	0.0320735224	0.0330603550
$Bias(\hat{\beta}_{\text{hochschulreife}})$	-0.0185912220	-0.0001542230	0.0008326096
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	-0.5768700792	-0.0047854093	0.0258351808
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0084314426	0.0035123786	0.0031410435
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0087770761	0.0035124023	0.0031417367
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4596891038	-0.4614798642	-0.4609249636
$Bias(\hat{\beta}_{\text{half-time}})$	-0.0021770347	-0.0039677952	-0.0034128945
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	-0.0047584203	-0.0086725476	-0.0074596819
$Var(\hat{\beta}_{\text{half-time}})$	0.0003162875	0.0003243474	0.0003258443
$MSE(\hat{\beta}_{\text{half-time}})$	0.0003210269	0.0003400908	0.0003374921
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9366676103	-0.9298827839	-0.9277169769
$Bias(\hat{\beta}_{\text{part-time}})$	0.0002376055	0.0070224319	0.0091882389
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0002536068	0.0074953493	0.0098070101
$Var(\hat{\beta}_{\text{part-time}})$	0.0006292143	0.0003661642	0.0004789937
$MSE(\hat{\beta}_{\text{part-time}})$	0.0006292708	0.0004154788	0.0005634175
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7997651802	-0.8023967964	-0.7988986017
$Bias(\hat{\beta}_{\text{not employed}})$	-0.0013040603	-0.0039356765	-0.0004374818
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	-0.0016332170	-0.0049290772	-0.0005479062

$Var(\hat{\beta}_{\text{not employed}})$	0.0001517673	0.0001656733	0.0001512944
$MSE(\hat{\beta}_{\text{not employed}})$	0.0001534679	0.0001811629	0.0001514858
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0891764947	0.0927380361	0.0914321051
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0031696158	0.0003919255	-0.0009140055
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.0343232197	0.0042440937	-0.0098976065
$Var(\hat{\beta}_{\text{married living together}})$	0.0003819491	0.0004309455	0.0004050571
$MSE(\hat{\beta}_{\text{married living together}})$	0.0003919956	0.0004310991	0.0004058925
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3700006976	0.3919466378	0.3906627652
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0103856942	0.0115602460	0.0102763734
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0273030120	0.0303907980	0.0270156179
$Var(\hat{\beta}_{\text{married living apart}})$	0.0018943565	0.0015591504	0.0015007033
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0020022191	0.0016927897	0.0016063072
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0159792338	-0.0100061489	-0.0112879080
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0072449536	-0.0012718687	-0.0025536278
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.8294849039	-0.1456180355	-0.2923684286
$Var(\hat{\beta}_{\text{widowed}})$	0.0010128528	0.0009158236	0.0010397592
$MSE(\hat{\beta}_{\text{widowed}})$	0.0010653422	0.0009174413	0.0010462802
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1223887101	0.1314968013	0.1295730806
$Bias(\hat{\beta}_{\text{single}})$	-0.0021012686	0.0070068227	0.0050831019
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.0168790177	0.0562842307	0.0408314145
$Var(\hat{\beta}_{\text{single}})$	0.0004763235	0.0006578228	0.0005629970
$MSE(\hat{\beta}_{\text{single}})$	0.0004807388	0.0007069184	0.0005888349
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0226815482	-0.0269710395	-0.0278177898
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0023592983	-0.0019301931	-0.0027769433
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.0942179943	-0.0770817813	-0.1108965441
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0008874832	0.0006330925	0.0006395904
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0008930495	0.0006368181	0.0006473018
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0110937646	-0.0143766652	-0.0160777473
$Bias(\hat{\beta}_{\text{SPD}})$	0.0022339550	-0.0010489456	-0.0027500276
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.1676171962	-0.0787040532	-0.2063389445
$Var(\hat{\beta}_{\text{SPD}})$	0.0009395381	0.0006547554	0.0006195972
$MSE(\hat{\beta}_{\text{SPD}})$	0.0009445287	0.0006558557	0.0006271599
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0885562473	0.0850851670	0.0843948863
$Bias(\hat{\beta}_{\text{die gruenen}})$	0.0014744654	-0.0019966149	-0.0026868956
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	0.0169319615	-0.0229280441	-0.0308548536
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0008994452	0.0006959405	0.0007036311
$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0009016193	0.0006999270	0.0007108505

$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0792275056	0.0844070548	0.0827418166
$Bias(\hat{\beta}_{\text{die linke}})$	0.0023504933	0.0075300425	0.0058648044
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	0.0305747229	0.0979492091	0.0762881415
$Var(\hat{\beta}_{\text{die linke}})$	0.0011784078	0.0008287520	0.0008705497
$MSE(\hat{\beta}_{\text{die linke}})$	0.0011839327	0.0008854535	0.0009049457
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0069942755	0.0101026426	0.0086002997
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	-0.0004063120	0.0027020551	0.0011997123
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	-0.0549026668	0.3651136007	0.1621104095
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0013602063	0.0010172043	0.0011107843
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0013603714	0.0010245054	0.0011122236
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1496225303	-0.1447740630	-0.1467602229
$Bias(\hat{\beta}_{\text{FDP}})$	0.0045530799	0.0094015472	0.0074153873
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.0295317779	0.0609794714	0.0480970196
$Var(\hat{\beta}_{\text{FDP}})$	0.0012923336	0.0010865925	0.0010159988
$MSE(\hat{\beta}_{\text{FDP}})$	0.0013130642	0.0011749816	0.0010709867
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1271475983	-0.1398976436	-0.1448236034
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0044413102	-0.0083087351	-0.0132346948
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.0337514027	-0.0631416066	-0.1005760667
$Var(\hat{\beta}_{\text{would not vote}})$	0.0012493865	0.0010749996	0.0009472228
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0012691118	0.0011440346	0.0011223800
$\mathbb{E}(\hat{\beta}_{\text{inverse Mills Ratio}})$	-0.053555283	-0.10946707	-0.11611338

Table C.34: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of Heckman Selection Model

Categories	p-value (true)	p-value ($MCAR$)	p-Value (MAR)	p-Value ($MNAR$)
intercept	0.0000000000	0.0045172246	0.0000000000	0.0000000000
age	0.0000000000	0.0000000000	0.0000000000	0.0000000000
female	0.0000000000	0.0045649891	0.0000000000	0.0000352578
fachhochschulreife	0.8970973124	0.6532193841	0.6448402038	0.6506492894
hochschulreife	0.6987555776	0.6552577803	0.6191582112	0.6202356523
mittlere reife	0.0115117133	0.2474829469	0.0940506648	0.0865741985
no graduation	0.0058831082	0.2190823041	0.0404076521	0.0401417402
volks-, hauptschule	0.0065408724	0.1622149659	0.1513897890	0.1459644911
half-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
part-time	0.0000000000	0.0000000000	0.0000000000	0.0000000000
not employed	0.0000000000	0.0000000000	0.0000000000	0.0000000000
single	0.0002454705	0.0094845585	0.0061449957	0.0055757978
married living apart	0.0000000011	0.0000550833	0.0000238799	0.0000186561
married living together	0.0021189162	0.0296743795	0.0249400632	0.0249902817
widowed	0.8397563199	0.6251807755	0.6442564422	0.6276495384
CDU-CSU	0.5884698278	0.6252349796	0.5909060181	0.5774236976
die gruenen	0.0762638735	0.1809665236	0.1780286508	0.1812774324
die linke	0.1299242680	0.2498115420	0.2027639535	0.2131372781
FDP	0.0082560314	0.0573113299	0.0587094104	0.0531167967
extreme right-wing	0.8910480181	0.6668136499	0.6932754949	0.6743930104
SPD	0.7770122356	0.6934082013	0.6820188447	0.6816754506
would not vote	0.0103321856	0.0741749792	0.0408084532	0.0311853836
inverse Mills Ratio		0.4934248428	0.4742324275	0.4560026174

Table C.35: Comparison of p-values of regression coefficients of data after applying Heckman Selection Model

	MCAR	MAR	MNAR
β_0	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	6.9353794756	6.9359497475	6.9396062838
$Bias(\hat{\beta}_0)$	-0.0716813522	-0.0711110804	-0.0674545440
$Bias_{\text{rel}}(\hat{\beta}_0)$	-0.0102298744	-0.0101484891	-0.0096266531
$Var(\hat{\beta}_0)$	0.0001630726	0.0001623701	0.0002817138
$MSE(\hat{\beta}_0)$	0.0053012889	0.0052191559	0.0048318293
β_{age}	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{age}})$	0.0112606772	0.0112689298	0.0112546516
$Bias(\hat{\beta}_{\text{age}})$	-0.0001114179	-0.0001031653	-0.0001174435
$Bias_{\text{rel}}(\hat{\beta}_{\text{age}})$	-0.0097974839	-0.0090717940	-0.0103273377
$Var(\hat{\beta}_{\text{age}})$	0.0000000004	0.0000000031	0.0000000164
$MSE(\hat{\beta}_{\text{age}})$	0.0000000128	0.0000000137	0.0000000302
β_{female}	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{female}})$	-0.3745803622	-0.3734668906	-0.3761678289
$Bias(\hat{\beta}_{\text{female}})$	0.0037237171	0.0048371886	0.0021362504
$Bias_{\text{rel}}(\hat{\beta}_{\text{female}})$	0.0098431852	0.0127865093	0.0056469136
$Var(\hat{\beta}_{\text{female}})$	0.0000032638	0.0000055671	0.0000091463
$MSE(\hat{\beta}_{\text{female}})$	0.0000171299	0.0000289655	0.0000137099
$\beta_{\text{no graduation}}$	0.2712024535	0.2712024535	0.2712024535

$\mathbb{E}(\hat{\beta}_{\text{no graduation}})$	0.2719901761	0.2692168259	0.2692370833
$Bias(\hat{\beta}_{\text{no graduation}})$	0.0007877226	-0.0019856276	-0.0019653702
$Bias_{\text{rel}}(\hat{\beta}_{\text{no graduation}})$	0.0029045556	-0.0073215696	-0.0072468748
$Var(\hat{\beta}_{\text{no graduation}})$	0.0000569120	0.0006860499	0.0006741099
$MSE(\hat{\beta}_{\text{no graduation}})$	0.0000575325	0.0006899926	0.0006779726
$\beta_{\text{volks-, hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{volks-, hauptschule}})$	-0.2216296616	-0.2225730455	-0.2253109683
$Bias(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0043358845	0.0033925006	0.0006545778
$Bias_{\text{rel}}(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0191882550	0.0150133536	0.0028968037
$Var(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0000027416	0.0001621821	0.0002327469
$MSE(\hat{\beta}_{\text{volks-, hauptschule}})$	0.0000215415	0.0001736912	0.0002331753
$\beta_{\text{mittlere reife}}$	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{mittlere reife}})$	0.2097463802	0.2088985554	0.2103246853
$Bias(\hat{\beta}_{\text{mittlere reife}})$	-0.0001057875	-0.0009536124	0.0004725176
$Bias_{\text{rel}}(\hat{\beta}_{\text{mittlere reife}})$	-0.0005041049	-0.0045442103	0.0022516688
$Var(\hat{\beta}_{\text{mittlere reife}})$	0.0000108323	0.0000298284	0.0000143050
$MSE(\hat{\beta}_{\text{mittlere reife}})$	0.0000108435	0.0000307378	0.0000145283
$\beta_{\text{fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{fachhochschulreife}})$	-0.0088630453	-0.0111873933	-0.0107039252
$Bias(\hat{\beta}_{\text{fachhochschulreife}})$	0.0024358716	0.0001115235	0.0005949917
$Bias_{\text{rel}}(\hat{\beta}_{\text{fachhochschulreife}})$	0.2155845241	0.0098702817	0.0526591768
$Var(\hat{\beta}_{\text{fachhochschulreife}})$	0.0000197259	0.0000934677	0.0001681482
$MSE(\hat{\beta}_{\text{fachhochschulreife}})$	0.0000256594	0.0000934801	0.0001685022
$\beta_{\text{hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{hochschulreife}})$	0.0341212258	0.0323089720	0.0329768221
$Bias(\hat{\beta}_{\text{hochschulreife}})$	0.0018934804	0.0000812267	0.0007490768
$Bias_{\text{rel}}(\hat{\beta}_{\text{hochschulreife}})$	0.0587531153	0.0025203952	0.0232432258
$Var(\hat{\beta}_{\text{hochschulreife}})$	0.0000097127	0.0000534746	0.0000515527
$MSE(\hat{\beta}_{\text{hochschulreife}})$	0.0000132979	0.0000534812	0.0000521138
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4528287017	-0.4527216017	-0.4526157309
$Bias(\hat{\beta}_{\text{half-time}})$	0.0046833673	0.0047904674	0.0048963381
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.0102365984	0.0104706908	0.0107020960
$Var(\hat{\beta}_{\text{half-time}})$	0.0000015784	0.0000045577	0.0000016636
$MSE(\hat{\beta}_{\text{half-time}})$	0.0000235123	0.0000275063	0.0000256378
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158
$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.9278924803	-0.9274115424	-0.9276524044
$Bias(\hat{\beta}_{\text{part-time}})$	0.0090127355	0.0094936735	0.0092528114
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0096196877	0.0101330138	0.0098759312
$Var(\hat{\beta}_{\text{part-time}})$	0.0000042723	0.0000732580	0.0001431972
$MSE(\hat{\beta}_{\text{part-time}})$	0.0000855017	0.0001633879	0.0002288117
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7906728556	-0.7906918926	-0.7903331052
$Bias(\hat{\beta}_{\text{not employed}})$	0.0077882642	0.0077692273	0.0081280147

$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.0097540933	0.0097302512	0.0101795998
$Var(\hat{\beta}_{\text{not employed}})$	0.0000029898	0.0000269951	0.0000018948
$MSE(\hat{\beta}_{\text{not employed}})$	0.0000636468	0.0000873559	0.0000679594
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0916191571	0.0914350659	0.0913243118
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0007269535	-0.0009110447	-0.0010217987
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.0078720529	-0.0098655445	-0.0110648812
$Var(\hat{\beta}_{\text{married living together}})$	0.0000191400	0.0000014118	0.0000002695
$MSE(\hat{\beta}_{\text{married living together}})$	0.0000196684	0.0000022418	0.0000013135
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3766125841	0.3755791081	0.3757808917
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0037738076	-0.0048072837	-0.0046055000
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.0099209849	-0.0126378960	-0.0121074259
$Var(\hat{\beta}_{\text{married living apart}})$	0.0001984753	0.0002893352	0.0002194420
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0002127169	0.0003124452	0.0002406526
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	- 0.0089948773	-0.0103507732	-0.0100399303
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0002605971	-0.0016164930	-0.0013056501
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-0.0298361283	-0.1850745519	-0.1494857176
$Var(\hat{\beta}_{\text{widowed}})$	0.0000680625	0.0000005757	0.0000021741
$MSE(\hat{\beta}_{\text{widowed}})$	0.0000681304	0.0000031887	0.0000038788
β_{single}	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1232049703	0.1228625795	0.1227410484
$Bias(\hat{\beta}_{\text{single}})$	-0.0012850083	-0.0016273991	-0.0017489302
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.0103221830	-0.0130725312	-0.0140487631
$Var(\hat{\beta}_{\text{single}})$	0.0000068619	0.0001143245	0.0000647004
$MSE(\hat{\beta}_{\text{single}})$	0.0000085131	0.0001169729	0.0000677591
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0247896313	-0.0250603947	-0.0254334381
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	0.0002512152	-0.0000195482	-0.0003925917
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	0.0100322182	-0.0007806542	-0.0156780504
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0000096594	0.0000176945	0.0000250355
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0000097225	0.0000176949	0.0000251896
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0129489153	-0.0131990711	-0.0134861572
$Bias(\hat{\beta}_{\text{SPD}})$	0.0003788043	0.0001286486	-0.0001584376
$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	0.0284222899	0.0096527070	-0.0118878245
$Var(\hat{\beta}_{\text{SPD}})$	0.0000082994	0.0000038249	0.0000164779
$MSE(\hat{\beta}_{\text{SPD}})$	0.0000084429	0.0000038415	0.0000165030
$\beta_{\text{die gruenen}}$	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{die gruenen}})$	0.0861594196	0.0861721909	0.0858419315
$Bias(\hat{\beta}_{\text{die gruenen}})$	-0.0009223623	-0.0009095910	-0.0012398504
$Bias_{\text{rel}}(\hat{\beta}_{\text{die gruenen}})$	-0.0105919086	-0.0104452499	-0.0142377706
$Var(\hat{\beta}_{\text{die gruenen}})$	0.0000029839	0.0000135436	0.0000177944

$MSE(\hat{\beta}_{\text{die gruenen}})$	0.0000038346	0.0000143709	0.0000193316
$\beta_{\text{die linke}}$	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{die linke}})$	0.0766285053	0.0765386126	0.0759818674
$Bias(\hat{\beta}_{\text{die linke}})$	-0.0002485069	-0.0003383996	-0.0008951449
$Bias_{\text{rel}}(\hat{\beta}_{\text{die linke}})$	-0.0032325255	-0.0044018309	-0.0116438560
$Var(\hat{\beta}_{\text{die linke}})$	0.0000052878	0.0000829108	0.0000555711
$MSE(\hat{\beta}_{\text{die linke}})$	0.0000053495	0.0000830253	0.0000563724
$\beta_{\text{extreme right-wing}}$	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{extreme right-wing}})$	0.0074446783	0.0069021312	0.0064076423
$Bias(\hat{\beta}_{\text{extreme right-wing}})$	0.0000440908	-0.0004984563	-0.0009929452
$Bias_{\text{rel}}(\hat{\beta}_{\text{extreme right-wing}})$	0.0059577427	-0.0673536127	-0.1341711303
$Var(\hat{\beta}_{\text{extreme right-wing}})$	0.0000039412	0.0000105129	0.0000030220
$MSE(\hat{\beta}_{\text{extreme right-wing}})$	0.0000039431	0.0000107613	0.0000040080
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1521668250	-0.1524688386	-0.1527640536
$Bias(\hat{\beta}_{\text{FDP}})$	0.0020087852	0.0017067716	0.0014115566
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	0.0130292021	0.0110703085	0.0091555115
$Var(\hat{\beta}_{\text{FDP}})$	0.0000318057	0.0001582649	0.0001013749
$MSE(\hat{\beta}_{\text{FDP}})$	0.0000358410	0.0001611779	0.0001033674
$\beta_{\text{would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{would not vote}})$	-0.1296837496	-0.1309278597	-0.1315006747
$Bias(\hat{\beta}_{\text{would not vote}})$	0.0019051589	0.0006610488	0.0000882339
$Bias_{\text{rel}}(\hat{\beta}_{\text{would not vote}})$	0.0144781119	0.0050235908	0.0006705268
$Var(\hat{\beta}_{\text{would not vote}})$	0.0000307147	0.0001447759	0.0003420967
$MSE(\hat{\beta}_{\text{would not vote}})$	0.0000343444	0.0001452129	0.0003421045

Table C.36: Expected value of regression coefficient estimation on income in order to evaluate the performance of Pattern Mixture Model

	ad-hoc	FIML	Heckman SM	PMM
β_0	7.0070608279	7.0070608279	7.0070608279	7.0070608279
$\mathbb{E}(\hat{\beta}_0)$	6.9918990979	6.9918994242	6.9834369706	6.9599825424
$Bias(\hat{\beta}_0)$	-0.0151617300	-0.0151614034	-0.0236238572	-0.0470782855
$Bias_{\text{rel}}(\hat{\beta}_0)$	-0.0021637788	-0.0021637322	-0.0033714360	-0.0067186923
$Var(\hat{\beta}_0)$	0.0040586057	0.0040586304	0.5002207476	0.0032153760
$MSE(\hat{\beta}_0)$	0.0042884838	0.0042884985	0.5007788343	0.0054317409
β_{Age}	0.0113720951	0.0113720951	0.0113720951	0.0113720951
$\mathbb{E}(\hat{\beta}_{\text{Age}})$	0.0104158533	0.0104158529	0.0103290286	0.0107637804
$Bias(\hat{\beta}_{\text{Age}})$	-0.0009562418	-0.0009562423	-0.0010430665	-0.0006083147
$Bias_{\text{rel}}(\hat{\beta}_{\text{Age}})$	-0.0840866856	-0.0840867264	-0.0917215807	-0.0534918725
$Var(\hat{\beta}_{\text{Age}})$	0.0000001725	0.0000001725	0.0000012543	0.0000004496
$MSE(\hat{\beta}_{\text{Age}})$	0.0000010869	0.0000010869	0.0000023423	0.0000008197

β_{Female}	-0.3783040793	-0.3783040793	-0.3783040793	-0.3783040793
$\mathbb{E}(\hat{\beta}_{\text{Female}})$	-0.3603941931	-0.3603937707	-0.3646385773	-0.3752790816
$Bias(\hat{\beta}_{\text{Female}})$	0.0179098862	0.0179103086	0.0136655020	0.0030249977
$Bias_{\text{rel}}(\hat{\beta}_{\text{Female}})$	0.0473425670	0.0473436836	0.0361230627	0.0079962069
$Var(\hat{\beta}_{\text{Female}})$	0.0001560970	0.0001560997	0.0025572657	0.0007523674
$MSE(\hat{\beta}_{\text{Female}})$	0.0004768611	0.0004768789	0.0027440117	0.0007615181
$\beta_{\text{No Graduation}}$	0.2712024535	0.2712024535	0.2712024535	0.2712024535
$\mathbb{E}(\hat{\beta}_{\text{No Graduation}})$	0.2665583578	0.2665584630	0.2781516623	0.2519122283
$Bias(\hat{\beta}_{\text{No Graduation}})$	-0.0046440957	-0.0046439905	0.0069492089	-0.0192902252
$Bias_{\text{rel}}(\hat{\beta}_{\text{No Graduation}})$	-0.0171240916	-0.0171237037	0.0256236947	-0.0711285053
$Var(\hat{\beta}_{\text{No Graduation}})$	0.0047502795	0.0047502859	0.0059502660	0.0003184682
$MSE(\hat{\beta}_{\text{No Graduation}})$	0.0047718471	0.0047718526	0.0059985575	0.0006905810
$\beta_{\text{Volks-, Hauptschule}}$	-0.2259655461	-0.2259655461	-0.2259655461	-0.2259655461
$\mathbb{E}(\hat{\beta}_{\text{Volks-, Hauptschule}})$	-0.2178593224	-0.2178595112	-0.1758143470	-0.2476348846
$Bias(\hat{\beta}_{\text{Volks-, Hauptschule}})$	0.0081062237	0.0081060349	0.0501511991	-0.0216693385
$Bias_{\text{rel}}(\hat{\beta}_{\text{Volks-, Hauptschule}})$	0.0358737155	0.0358728799	0.2219417960	-0.0958966481
$Var(\hat{\beta}_{\text{Volks-, Hauptschule}})$	0.0032971926	0.0032971667	0.0041650294	0.0022517878
$MSE(\hat{\beta}_{\text{Volks-, Hauptschule}})$	0.0033629035	0.0033628745	0.0066801721	0.0027213480
$\beta_{\text{Mittlere Reife}}$	0.2098521677	0.2098521677	0.2098521677	0.2098521677
$\mathbb{E}(\hat{\beta}_{\text{Mittlere Reife}})$	0.2039732445	0.2039727839	0.2035139941	0.2131636660
$Bias(\hat{\beta}_{\text{Mittlere Reife}})$	-0.0058789232	-0.0058793838	-0.0063381737	0.0033114983
$Bias_{\text{rel}}(\hat{\beta}_{\text{Mittlere Reife}})$	-0.0280145938	-0.0280167885	-0.0302030412	0.0157801481
$Var(\hat{\beta}_{\text{Mittlere Reife}})$	0.0028012188	0.0028012103	0.0035348148	0.0002776289
$MSE(\hat{\beta}_{\text{Mittlere Reife}})$	0.0028357805	0.0028357774	0.0035749873	0.0002885949
$\beta_{\text{Fachhochschulreife}}$	-0.0112989169	-0.0112989169	-0.0112989169	-0.0112989169
$\mathbb{E}(\hat{\beta}_{\text{Fachhochschulreife}})$	0.0124622146	0.0124619936	0.0121641972	-0.0077323281
$Bias(\hat{\beta}_{\text{Fachhochschulreife}})$	0.0237611314	0.0237609104	0.0234631141	0.0035665888
$Bias_{\text{rel}}(\hat{\beta}_{\text{Fachhochschulreife}})$	2.1029565657	2.1029370048	2.0765808216	0.3156575863
$Var(\hat{\beta}_{\text{Fachhochschulreife}})$	0.0035444498	0.0035444588	0.0039850419	0.0008895124
$MSE(\hat{\beta}_{\text{Fachhochschulreife}})$	0.0041090411	0.0041090397	0.0045355600	0.0009022330
$\beta_{\text{Hochschulreife}}$	0.0322277454	0.0322277454	0.0322277454	0.0322277454
$\mathbb{E}(\hat{\beta}_{\text{Hochschulreife}})$	0.0332303614	0.0332302411	0.0353930436	0.0348303933
$Bias(\hat{\beta}_{\text{Hochschulreife}})$	0.0010026160	0.0010024957	0.0031652982	0.0026026480
$Bias_{\text{rel}}(\hat{\beta}_{\text{Hochschulreife}})$	0.0311103379	0.0311066049	0.0982165576	0.0807579909
$Var(\hat{\beta}_{\text{Hochschulreife}})$	0.0029257449	0.0029257450	0.0034905884	0.0000082209
$MSE(\hat{\beta}_{\text{Hochschulreife}})$	0.0029267550	0.0029267500	0.0035006075	0.0000149947
$\beta_{\text{half-time}}$	-0.4575120691	-0.4575120691	-0.4575120691	-0.4575120691
$\mathbb{E}(\hat{\beta}_{\text{half-time}})$	-0.4237710004	-0.4237707342	-0.4236278362	-0.4289041088
$Bias(\hat{\beta}_{\text{half-time}})$	0.0337410686	0.0337413349	0.0338842329	0.0286079602
$Bias_{\text{rel}}(\hat{\beta}_{\text{half-time}})$	0.0737490242	0.0737496061	0.0740619432	0.0625294110
$Var(\hat{\beta}_{\text{half-time}})$	0.0003608333	0.0003608351	0.0022109048	0.0001927363
$MSE(\hat{\beta}_{\text{half-time}})$	0.0014992930	0.0014993128	0.0033590460	0.0010111517
$\beta_{\text{part-time}}$	-0.9369052158	-0.9369052158	-0.9369052158	-0.9369052158

$\mathbb{E}(\hat{\beta}_{\text{part-time}})$	-0.8839084275	-0.8839094428	-0.8734872473	-0.9012367235
$Bias(\hat{\beta}_{\text{part-time}})$	0.0529967883	0.0529957730	0.0634179685	0.0356684924
$Bias_{\text{rel}}(\hat{\beta}_{\text{part-time}})$	0.0565657949	0.0565647113	0.0676887773	0.0380705452
$Var(\hat{\beta}_{\text{part-time}})$	0.0004545761	0.0004545780	0.0081682492	0.0015044339
$MSE(\hat{\beta}_{\text{part-time}})$	0.0032632356	0.0032631300	0.0121900880	0.0027766753
$\beta_{\text{not employed}}$	-0.7984611199	-0.7984611199	-0.7984611199	-0.7984611199
$\mathbb{E}(\hat{\beta}_{\text{not employed}})$	-0.7481781999	-0.7481785426	-0.7413875504	-0.7612613519
$Bias(\hat{\beta}_{\text{not employed}})$	0.0502829200	0.0502825773	0.0570735695	0.0371997680
$Bias_{\text{rel}}(\hat{\beta}_{\text{not employed}})$	0.0629747883	0.0629743591	0.0714794597	0.0465893292
$Var(\hat{\beta}_{\text{not employed}})$	0.0001957666	0.0001957666	0.0057548007	0.0009289524
$MSE(\hat{\beta}_{\text{not employed}})$	0.0027241386	0.0027241041	0.0090121931	0.0023127752
$\beta_{\text{married living together}}$	0.0923461105	0.0923461105	0.0923461105	0.0923461105
$\mathbb{E}(\hat{\beta}_{\text{married living together}})$	0.0730691859	0.0730691705	0.0715470348	0.0859618026
$Bias(\hat{\beta}_{\text{married living together}})$	-0.0192769247	-0.0192769400	-0.0207990758	-0.0063843079
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living together}})$	-0.2087464708	-0.2087466368	-0.2252295808	-0.0691345625
$Var(\hat{\beta}_{\text{married living together}})$	0.0004259310	0.0004259335	0.0004758751	0.0004094003
$MSE(\hat{\beta}_{\text{married living together}})$	0.0007975309	0.0007975339	0.0009084767	0.0004501597
$\beta_{\text{married living apart}}$	0.3803863918	0.3803863918	0.3803863918	0.3803863918
$\mathbb{E}(\hat{\beta}_{\text{married living apart}})$	0.3389024844	0.3389021418	0.3366858969	0.3573671872
$Bias(\hat{\beta}_{\text{married living apart}})$	-0.0414839074	-0.0414842500	-0.0437004949	-0.0230192045
$Bias_{\text{rel}}(\hat{\beta}_{\text{married living apart}})$	-0.1090572857	-0.1090581863	-0.1148844854	-0.0605153208
$Var(\hat{\beta}_{\text{married living apart}})$	0.0023498808	0.0023498773	0.0034733274	0.0010517536
$MSE(\hat{\beta}_{\text{married living apart}})$	0.0040707955	0.0040708203	0.0053830606	0.0015816373
β_{widowed}	-0.0087342802	-0.0087342802	-0.0087342802	-0.0087342802
$\mathbb{E}(\hat{\beta}_{\text{widowed}})$	-0.0180390876	-0.0180390740	-0.0185558936	0.0077166306
$Bias(\hat{\beta}_{\text{widowed}})$	-0.0093048075	-0.0093047938	-0.0098216134	0.0010176496
$Bias_{\text{rel}}(\hat{\beta}_{\text{widowed}})$	-1.0653204674	-1.0653189039	-1.1244903081	0.1165121310
$Var(\hat{\beta}_{\text{widowed}})$	0.0011846509	0.0011846387	0.0011579968	0.0002272012
$MSE(\hat{\beta}_{\text{widowed}})$	0.0012712304	0.0012712179	0.0012544609	0.0002282368
β_{single}	0.1244899786	0.1244899786	0.1244899786	0.1244899786
$\mathbb{E}(\hat{\beta}_{\text{single}})$	0.1008751057	0.1008749109	0.0999261900	0.1147122833
$Bias(\hat{\beta}_{\text{single}})$	-0.0236148729	0.0236150677	-0.0245637886	-0.0097776953
$Bias_{\text{rel}}(\hat{\beta}_{\text{single}})$	-0.1896929631	-0.1896945275	-0.1973153895	-0.0785420271
$Var(\hat{\beta}_{\text{single}})$	0.0004794570	0.0004794683	0.0005752253	0.0004860593
$MSE(\hat{\beta}_{\text{single}})$	0.0010371192	0.0010371397	0.0011786050	0.0005816626
$\beta_{\text{CDU-CSU}}$	-0.0250408465	-0.0250408465	-0.0250408465	-0.0250408465
$\mathbb{E}(\hat{\beta}_{\text{CDU-CSU}})$	-0.0413015937	-0.0413014679	-0.0393838344	-0.0258640952
$Bias(\hat{\beta}_{\text{CDU-CSU}})$	-0.0162607472	-0.0162606201	-0.0143429879	-0.0008232487
$Bias_{\text{rel}}(\hat{\beta}_{\text{CDU-CSU}})$	-0.6493689117	-0.6493638346	-0.5727836695	-0.0328762343
$Var(\hat{\beta}_{\text{CDU-CSU}})$	0.0007178734	0.0007178844	0.0007071371	0.0004986997
$MSE(\hat{\beta}_{\text{CDU-CSU}})$	0.0009822853	0.0009822922	0.0009128584	0.0004993774
β_{SPD}	-0.0133277196	-0.0133277196	-0.0133277196	-0.0133277196
$\mathbb{E}(\hat{\beta}_{\text{SPD}})$	-0.0269776734	-0.0269774210	-0.0263829368	-0.0132190147
$Bias(\hat{\beta}_{\text{SPD}})$	-0.0136499538	-0.0136497013	-0.0130552172	0.0001087050

$Bias_{\text{rel}}(\hat{\beta}_{\text{SPD}})$	-1.0241777382	-1.0241587972	-0.9795537048	0.0081563082
$Var(\hat{\beta}_{\text{SPD}})$	0.0006879219	0.0006879397	0.0006736271	0.0004019095
$MSE(\hat{\beta}_{\text{SPD}})$	0.0008742432	0.0008742540	0.0008440658	0.0004019213
$\beta_{\text{Die Gruenen}}$	0.0870817819	0.0870817819	0.0870817819	0.0870817819
$\mathbb{E}(\hat{\beta}_{\text{Die Gruenen}})$	0.0699375324	0.0699368792	0.0698348263	0.0814774337
$Bias(\hat{\beta}_{\text{Die Gruenen}})$	-0.0171442495	-0.0171449028	-0.0172469556	-0.0056043482
$Bias_{\text{rel}}(\hat{\beta}_{\text{Die Gruenen}})$	-0.1968752725	-0.1968827736	-0.1980546938	-0.0643572985
$Var(\hat{\beta}_{\text{Die Gruenen}})$	0.0008226594	0.0008226664	0.0008476972	0.0003304210
$MSE(\hat{\beta}_{\text{Die Gruenen}})$	0.0011165847	0.0011166141	0.0011451547	0.0003618297
$\beta_{\text{Die Linke}}$	0.0768770122	0.0768770122	0.0768770122	0.0768770122
$\mathbb{E}(\hat{\beta}_{\text{Die Linke}})$	0.0600895876	0.0600886500	0.0605699254	0.0709440868
$Bias(\hat{\beta}_{\text{Die Linke}})$	-0.0167874246	-0.0167883626	-0.0163070869	-0.0059329254
$Bias_{\text{rel}}(\hat{\beta}_{\text{Die Linke}})$	-0.2183672873	-0.2183794883	-0.2121191547	-0.0771742453
$Var(\hat{\beta}_{\text{Die Linke}})$	0.0010044975	0.0010045144	0.0010124707	0.0002894757
$MSE(\hat{\beta}_{\text{Die Linke}})$	0.0012863151	0.0012863635	0.0012783918	0.0003246753
$\beta_{\text{Extreme Right-Wing}}$	0.0074005875	0.0074005875	0.0074005875	0.0074005875
$\mathbb{E}(\hat{\beta}_{\text{Extreme Right-Wing}})$	-0.0037889836	-0.0037863123	-0.0025346386	0.0062591849
$Bias(\hat{\beta}_{\text{Extreme Right-Wing}})$	-0.0111895711	-0.0111868997	-0.0099352260	-0.0011414025
$Bias_{\text{rel}}(\hat{\beta}_{\text{Extreme Right-Wing}})$	-1.5119841687	-1.5116232045	-1.3424915372	-0.1542313416
$Var(\hat{\beta}_{\text{Extreme Right-Wing}})$	0.0010832206	0.0010831891	0.0010464568	0.0002212753
$MSE(\hat{\beta}_{\text{Extreme Right-Wing}})$	0.0012084271	0.0012083358	0.0011451655	0.0002225781
β_{FDP}	-0.1541756102	-0.1541756102	-0.1541756102	-0.1541756102
$\mathbb{E}(\hat{\beta}_{\text{FDP}})$	-0.1567216987	-0.1567219824	-0.1543218537	-0.1503988556
$Bias(\hat{\beta}_{\text{FDP}})$	-0.0025460885	-0.0025463721	-0.0001462434	0.0037767546
$Bias_{\text{rel}}(\hat{\beta}_{\text{FDP}})$	-0.0165142104	-0.0165160504	-0.0009485511	0.0244964463
$Var(\hat{\beta}_{\text{FDP}})$	0.0010356435	0.0010356503	0.0012532831	0.0000503189
$MSE(\hat{\beta}_{\text{FDP}})$	0.0010421261	0.0010421343	0.0012533045	0.0000645828
$\beta_{\text{Would not vote}}$	-0.1315889085	-0.1315889085	-0.1315889085	-0.1315889085
$\mathbb{E}(\hat{\beta}_{\text{Would not vote}})$	-0.1567925656	-0.1567930100	-0.1549267899	-0.1267838888
$Bias(\hat{\beta}_{\text{Would not vote}})$	-0.0252036571	-0.0252041011	-0.0233378814	0.0048050197
$Bias_{\text{rel}}(\hat{\beta}_{\text{Would not vote}})$	-0.1915332940	-0.1915366682	-0.1773544717	0.0365153854
$Var(\hat{\beta}_{\text{Would not vote}})$	0.0009945485	0.0009945598	0.0012078546	0.0017879792
$MSE(\hat{\beta}_{\text{Would not vote}})$	0.0016297728	0.0016298065	0.0017525113	0.0018110674

Table C.37: Expected value, variance, bias and MSE of regression coefficient estimates on income in order to evaluate the performance of different correction methods on alternative MNAR model

D Electronic Appendix

The electronic appendix consists of the 3 folders “Data”, “Code” and “Thesis”.

- “Data” consists of the subfolders “Allbus2014” and “SimulatedData”. The subfolder “Allbus2014” includes the provided Allbus data set as downloaded, converted and shortened version. The folder “SimulatedData” is further subdivided in folders that contain the simulated data set (“Data Basis”), as well as all its modifications due to application of the different TSE error models (“Error Models”) and the nonresponse correction methods (“Nonresponse”).
- “Code” contains the subfolders “1.Data Basis”, “2.Error Models”, “3.Nonresponse” and “4.Other”. Each of these subfolders starts with a **R** file “0.Content” explaining the structure and content of the folder before the respective R code follows in chronological ordered subfiles. Only the subfolder “4.Other” does not have a ccertain order, since it contains additional R code for purposes of graphical illustration and statistical testing.
- “Thesis” contains the formulated thesis in PDF format.

Affidavit

I hereby declare that the present master's thesis has been written only by the undersigned and without any assistance from third parties. Furthermore, I confirm that no sources have been used in the preparation of this thesis other than those indicated in the thesis itself.

Munich, 15.11.2016

....
(Nina Scharrer)