A conditional minimax estimator for treating nonresponse

Gabler, Siegfried; Häder, Sabine

A Conditional Minimax Estimator for Treating Nonresponse

SIEGFRIED GABLER AND SABINE HÄDER

Abstract: *Losses due to nonresponse may lead to systematic biases in the samples which result in biased estimates. A usual way to compensate for this bias is the adjustment of the net sample to known population data. We choose another approach. If auxiliary information is available for each individual of the gross sample we adjust the net sample to the gross sample. The advantage is - in contrast to the usual post-stratification - that each element of the net sample may get an „own" weight. For the construction of these weights the conditional minimax principle is applied. It determines optimal weights conditional on the selected net sample which minimizes the maximal loss under the assumption that a certain variance of the unknown population values in the gross sample is finite. The effects of our weighting procedure are shown for data of the German General Social Survey (ALLBUS) 1996.*

Keywords: *ALLBUS, BLU estimator, conditional minimax, nonresponse, parameter space, representative estimator*

1 Introduction

The response rates in German social surveys declined during the last decades. Nowadays they have reached a level of about 60 percent. Typical for academic surveys is a wide range of nonresponse rates from about 20 to about 50 percent - depending on the conducting commercial institute and the subject of the survey (Schnell 1996). Unfortunately, we cannot assume on principle that the resulting net samples are unbiased. That means, that the losses of the gross samples are frequently not at random but systematic. A usual way to compensate for this nonresponse bias is the adjustment of the net sample to known population data (Little 1989, Deville and Särndal 1992, Deville, Särndal and Sautory 1993, Gabler and Häder 1997, Häder and Gabler 1997). Elliot (1991, 1996) discusses pros and cons of various procedures for post-stratification. We want to demonstrate another approach of correction for nonresponse bias.

The application of post-stratification means to classify the net sample into multivariate cells for which we have reference data of the population in the form of some marginal distributions.The weights are constructed to adjust the cell frequencies in the net sample to the reference data. However, this approach is connected with practical difficulties or restrictions resulting from the (non-) availability of appropriate population data (Elliot 1996, p. 2).

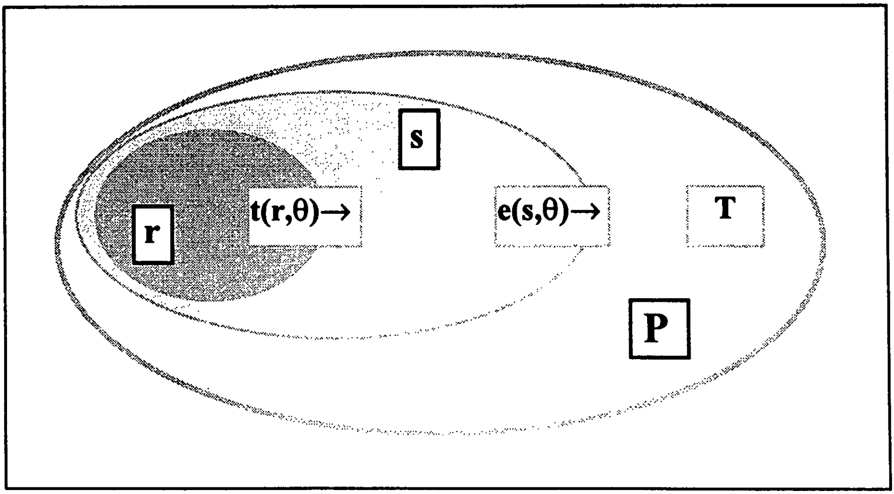
In our approach we estimate parameters of the population, for example the population total of a variable of interest, not directly from the net sample but by considering auxiliary information from the gross sample with the aim to adjust the net sample to the gross sample. The advantage is - in contrast to the usual post-stratification - that this information is available for each individual - that means each element of the net sample constitutes one cell and gets an "own" weight. For the construction of the weights the conditional minimax principle is applied. This way, some disadvantages of iterative procedures for the computation of weights can be avoided, for example the order of the variables in the adjusting process does not play a role. It is also possible to interpret our weights as BLU estimator in the framework of linear regression models.

The idea is demonstrated for data of the German General Social Survey (ALLBUS) 1996 whose response rate reached only 54%.

2 The conditional minimax principle

After having selected a sample s of a population $P=\{1,\dots,N\}$ the statistician estimates the population total $T= y_1+\dots+y_N$ of a variable of interest by an estimate $e(s,\theta)=\Sigma a_{si}y_i$ where a_{si} are real numbers with $a_{si}=0$ for $i\notin s$ and $\theta=(y_1,\dots,y_N)$. He will choose an estimate which

Figure 1: The estimation process



has good properties. Unfortunately, due to nonresponse the statistician does not get the y_i-values for all $i \in s$ but only for the units in a subset r of s. In many cases it cannot be assumed that r is a simple random sample of s. Thus the question remains how to estimate T using only the y-values in r. Often there exist x_i-values of auxiliary variables for all units in s, for example regional or demographic values. Our suggestion is to estimate T not directly but to estimate $e(s,\theta)$ by an estimate $t(r,\theta)=\Sigma b_{ri} y_i$ where b_{ri} are real numbers with $b_{ri}=0$ for $i \notin r$, and thus to estimate indirectly also T.

A decision theoretical approach for choosing an optimal estimator is given by the conditional minimax principle. This principle has been defined by Gabler (1988) and yields one possibility for computing $t(r,\theta)$. It can be also used for constructing cell weights, as Gabler (1991) shows. The conditional minimax principle says that in a given class D of estimators the optimal one in D minimizes the supremum of $L(t,\theta)$ where $L(t,\theta)$ denotes the loss of the estimate t at θ and the supremum is taken with respect to θ. We will consider only

$$L(t,\theta) = [t(r,\theta) - e(s,\theta)]^2$$

as loss function. Obviously, the loss function depends only on θ_s containing those y_i-values for which $i \in s$. Thus $L(t,\theta)=L(t,\theta_s)$, and we omit the suffix s in θ_s having in mind that θ now is an element of $\Re^n$ where n denotes the sample size of s. In general, the loss is unbounded on $\Re^n$. Therefore, we restrict ourselves to a subset Ω of $\Re^n$ known as parameter space. We consider only parameter spaces Ω on $\Re^n$ defined by quadratic forms in θ, i.e.

$$\Omega = \{\theta \in \Re^n : \theta' V \theta \leq c^2\}$$

where V is a nonnegative definite symmetric matrix and $c \neq 0$ is a given real number. Ω reflects the a-priori assumption about the y_i-values in s we have in mind. For example, Ω may be the set of all θ with bounded sample variance or more general

$$\Omega = \{\theta \in \Re^n : \sum_{i \in s} g_i (y_i - x_i \frac{ys}{xs})^2 \leq c^2 \}$$

where g_i and x_i are positive numbers and ys (xs) denotes the sum of all y-values (x-values) in s. An early paper dealing with such parameter spaces has been given by Wynn (1976). Minimax strategies are considered also by Bickel and Lehmann (1981), Chaudhuri and Stenger (1992), Stenger (1979) and Stenger and Gabler (1996).

If V is singular there exists a matrix X with VX=0. In this case the loss $L(t,\theta)$ is unbounded on Ω unless the linear estimator $t(r,\theta)$ satisfies $t(r,\xi)=e(s,\xi)$ for each column ξ

of X. Due to Hájek (1981) it means that t is a representative estimator for e with respect to X.

Thus our task is to look for a representative estimator $t(r,\theta)$ which minimizes the maximum loss on Ω.

The following Lemma can be derived from the Cauchy-Schwarz inequality and can be found in Gabler(1990, p. 111).

Lemma. Let W be a nonnegative definite symmetric n×n-matrix. For any n-dimensional vectors θ and α we have

$$(\theta' W\alpha)^2 \leq \theta' W\theta \cdot \alpha' W\alpha$$

Equality holds if and only if $A\theta$ and $ABW\alpha$ are proportional where $W=A'A$ and B is a symmetric reflexive g-inverse of W (see Rao and Mitra 1971).

Theorem. Let V be a nonnegative definite symmetric n×n-matrix of rank n-H and VQ=0, Q a n×H-matrix of rank H. We assume that R is a n×H matrix of rank H with $R'Q=I$, the identity matrix. Let U be the symmetric reflexive g-inverse of V with UR=0. The conditional minimax solution $t^c(r,\theta)=\Sigma b_{ri}y_i$ for $e(s,\theta)=\Sigma a_{si}y_i$ on $\Omega=\{\theta\in\Re^n: \theta' V\theta\leq c^2\}$ is given by

$$b^c = U_{rr}^{-1}U_{rs}a + U_{rr}^{-1}Q_r(Q_r' U_{rr}^{-1}Q_r)^{-1}(Q'-Q_r' U_{rr}^{-1}U_{rs})a$$

Proof. For the n-dimensional vector $\alpha = (b_{si} - a_{si})_{i\in s}$ we must have $\alpha' Q=0$ as shown above to get finite maximal loss. We define $\zeta=(I - QR')\theta$. Then $R'\zeta=0$ and

$$L(t,\theta) = (\alpha'\theta)^2 = (\alpha'\zeta)^2$$

For any such ζ there exists η with $\zeta=U\eta$. We get

$$(\alpha'\theta)^2 = (\alpha'\zeta)^2=(\alpha' U\eta)^2$$

From the lemma it follows

$$(\alpha'\theta)^2 = (\alpha' U\eta)^2 \leq (\alpha' U\alpha)(\eta' U\eta) = (\alpha' U\alpha)(\eta' UVU\eta) = (\alpha' U\alpha)(\zeta' V\zeta) = (\alpha' U\alpha)(\theta' V\theta)$$

In order to obtain a conditional minimax solution we have to minimize

$\alpha' U\alpha$ with the constraint $Q'\alpha=0$.

Defining the n-dimensional vector $a = (a_{si,})_{i\in s}$ and the m-dimensional vector $b=(b_{ri})_{i\in r}$ where m is the size of the sample r, we have $\alpha'U\alpha = a'Ua - 2a'U_{sr}b + b'U_{rr}b$ and the minimum of $\alpha'U\alpha$ under the constraint $a'Q=b'Q_r$ is attained at

$$b^c = U_{rr}^{-1}U_{rs}a + U_{rr}^{-1}Q_r(Q_r'U_{rr}^{-1}Q_r)^{-1}(Q'-Q_r'U_{rr}^{-1}U_{rs})a$$

provided that all inverses exist.
In the case $U_{rs}a=0$, for example if $a=Rx$, the above expression simplifies to

$$b^c = U_{rr}^{-1}Q_r(Q_r'U_{rr}^{-1}Q_r)^{-1}Q'a$$

Examples for b^c such as the ratio estimator can be found in Gabler (1990, pp. 114-116).

Remark 1. Let $Y=(Y_i)_{i\in s}$ be a superpopulation model with expectation $Q\beta$, Q a n×H matrix of rank H and β a n-dimensional parameter vector, and variance-covariance matrix U, where U is a nonnegative definite symmetric matrix of rank n-H and R a n×H matrix with $R'Q=I$ and $UR=0$. Analogous to chapter 5.5 in Gabler (1990) it can be shown that

b^c is a BLU estimator with respect to the above model.

From $E[R'Y]=R'Q\beta=\beta$ and $var(R'Y)=R'UR=0$ we conclude $R'Y=\beta$ with probability one. Estimation of β is the same as estimation of $R'Y$.

Remark 2. Let $Y=(Y_i)_{i\in s}$ be singular multivariate normally distributed with expectation $Q\beta$, Q a n×H matrix of rank H, β a n-dimensional parameter vector, and variance-covariance matrix U, where U is a nonnegative definite symmetric matrix of rank n-H and R a n×H matrix with $R'Q=I$ and $UR=0$. Analogous to chapter 5.5 in Gabler (1990), after a suitable truncation of the distribution of Y to $\Omega = \{\theta\in\Re^n: \theta'V\theta\le c^2\}$ and choosing

$$f(\beta) = 1 \text{ for all } \beta\in\Re^H$$

as prior distribution of β, it can be shown that

b^c is a Bayes estimator with respect to the above model.

If our viewpoint is the minimax principle we will start the computations with V and Q. If our viewpoint is the superpopulation approach we will start the computations with U and R.

3 Special cases

For real computations there is a drawback in the formula of b^c. If the sizes of the samples r and s are large, say 2000 and 4000, the matrices U_{rr} and U_{rs} are huge and b^c cannot be computed on a PC because of memory limitations. Fortunately, there exist special cases for which this problem can be eliminated.
If

$$V=(I-RQ')G(I-QR')$$

where G is a positive definite matrix, then

$$U = G^{-1} - G^{-1}R(R'G^{-1}R)^{-1}R'G^{-1}.$$

U depends on Q only via R'Q=I.

We assume in addition that G_{rt} is a matrix of zeros. Then the inverse of U_{rr} is

$$U_{rr}^{-1} = G_{rr} + R_r(R_t'G_{tt}^{-1}R_t)^{-1}R_r'$$

where R_r consists of the rows of R belonging to r.

Proof. From $U=G^{-1}-G^{-1}R(R'G^{-1}R)^{-1}R'G^{-1}$ it follows that

$$U_{rr} = G_{rr}^{-1} - G_{rr}^{-1}R_r(R'G^{-1}R)^{-1}R_r'G_{rr}^{-1}$$

since G_{rt} is a matrix of zeros (which implies $(G_{rr})^{-1} = (G^{-1})_{rr} =: G_{rr}^{-1}$).

The above formula for U_{rr}^{-1} results from the fact that

$$(A - BB')^{-1} = A^{-1} + A^{-1}B(I - B'A^{-1}B)^{-1}B'A^{-1}.$$

Defining t as complementary set of r in s and

$$C = R'G^{-1}R - R_r'G_{rr}^{-1}R_r = R_t'G_{tt}^{-1}R_t$$

we have

$$U_{rr}^{-1}U_{rs}a = a_r - R_rC^{-1}R_t'G_{tt}^{-1}a_t$$

$$U_{rr}^{-1}Q_r = G_{rr}Q_r + R_rC^{-1}R_r'Q_r$$

$$(Q'-Q_r'U_{rr}^{-1}U_{rs})a = Q_t'a_t + Q_r'R_rC^{-1}R_t'G_{tt}^{-1}a_t.$$

and it follows

$$b^c = a_r - R_r C^{-1} R_t' G_{tt}^{-1} a_t +$$

$$\left(G_{rr}Q_r + R_r C^{-1} R_r' Q_r\right)\left(Q_r' G_{rr} Q_r + Q_r' R_r C^{-1} R_r' Q_r\right)^{-1}\left(Q_t' a_t + Q_r' R_r C^{-1} R_t' G_{tt}^{-1} a_t\right)$$

If G is a diagonal matrix this formula helps to save memory for software packages, such as GAUSS, with matrix language and elementwise multiplication of matrices. For example, $G_{rr}Q_r$ may be programmed as $g_r .{*}Q_r$ where g_r is the vector of the diagonal elements of G belonging to r, and .* is the elementwise multiplication operator. Thus all terms involved in b^c may be computed by linkage of relative small components.

The question is how restrictive the representation of $V=(I-RQ')G(I-QR')$ with G as diagonal matrix is. Our feeling is that this restriction is not a real one. We will demonstrate it for the case H=1 and $a=R\eta$.

Let w be a positive vector with $w'Q_r = a'Q = \eta$. Then there exists always a positive diagonal matrix G yielding w as conditional minimax solution.

Proof. Let G_{tt} be an arbitrary diagonal matrix and let g_r denote the vector consisting of the diagonal elements of G_{rr}. We define

$$g_r = [k \cdot w - R_r(R_t' G_{tt}^{-1} R_t)^{-1} R_r' Q_r]./Q_r$$

where k can be always chosen so large that g_r is a positive vector. The operator ./ means elementwise division. It follows

$$\begin{aligned}
b^c &= U_{rr}^{-1} Q_r \left(Q_r' U_{rr}^{-1} Q_r\right)^{-1} \eta \\
&= \left(g_r .{*}Q_r + R_r(R_t' G_{tt}^{-1} R_t)^{-1} R_r' Q_r\right)\left(Q_r'[g_r .{*}Q_r + R_r(R_t' G_{tt}^{-1} R_t)^{-1} R_r' Q_r]\right)^{-1} \eta \\
&= \frac{k \cdot w}{Q_r'(k \cdot w)} \eta \\
&= w
\end{aligned}$$

Thus the restriction to parameter spaces defined by

$$\Omega = \{\theta \in \Re^n : \theta' V \theta \leq c^2\}$$

where

$$V=(I-RQ')G(I-QR')$$

with diagonal matrix G is not as restrictive as it seemed at first.

4 Application to the ALLBUS 1996

Now we will apply the above results to the data of the ALLBUS 1996. The ALLBUS is a bi-annually General Social Survey similar to the American General Social Survey and the British Social Attitudes Survey (see Davis et al. 1994). The ALLBUS is conducted as face-to-face interviews. Until 1992 the underlying sampling design of the ALLBUS theoretically selected German households with equal probabilities and one respondent within a chosen household proportional to the inverse of the number of the target people in that household. 1994 and 1996 the design was changed as follows: After stratification of the communities due to regional aspects, and selecting communities as primary units proportional to the number of the inhabitants belonging to the target population, the respondents were selected from local registers. This procedure implies equal inclusion probabilities for all persons of the target population. If all selected people in the gross sample were attainable and would answer the questions in the face-to-face interviews, the researcher could estimate the population total of a variable of interest by the sample total or by the ratio estimator each multiplied by the sampling fraction. Unfortunately, only a subset of the chosen people were willing or able to answer the questions. About 54%, i.e. 3518 persons were recorded in the official ALLBUS file distributed by the Zentralarchiv in Cologne.

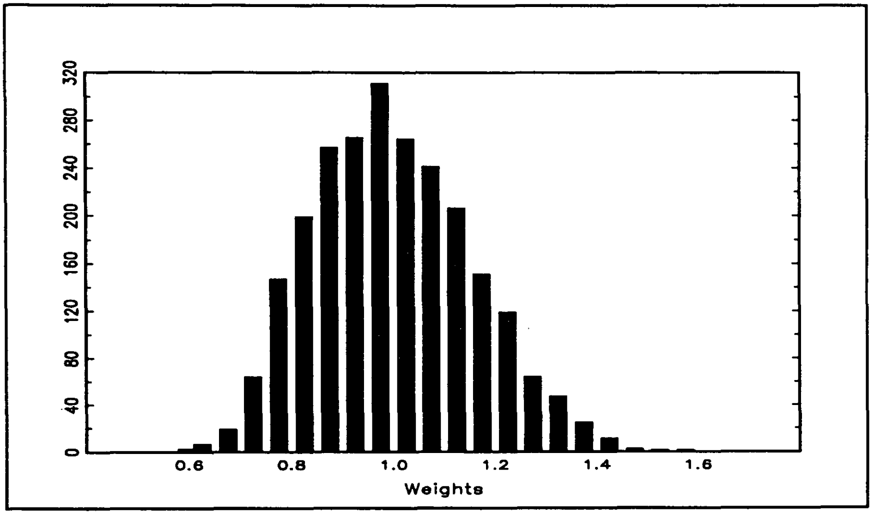
One of the advantages of this ALLBUS design is the knowledge of additional information for (nearly) all people of the gross sample. More specifically we obtained after some corrections 6488 selected persons from the local registers, 4430 in West Germany and 2058 in East Germany. Since West and East Germany are treated in the design separately we should apply our formulae also for both parts of Germany separately. We will present the results exemplary for West Germany with 2402 respondents.

As auxiliary variables we use age[1], gender, nationality, regional characteristics and the BIK variable which classifies the communities. Other variables - if available - could of course be used. For the gross sample the values of these variables are distributed similar to that in the population. For example, the portion of males in the gross sample is 48%, in contrast to 50% in the ALLBUS net sample. The columns of the matrix Q consist of the values of the variables which were categorized except age. Q has 4430 rows and 23 columns, 1 for age, 1 for gender, 1 for nationality, 10 for the federal states and 10 for the BIK variable. As R we used $Q(Q'Q)^{-1}$ and for a the vector consisting of ones. After computing b^c in the conditional minimax estimator we normed the weights in order to have sum equals 2402 which is the size of the ALLBUS sample in West Germany. The

[1] In Berlin the age was available only in classified form. We used the class mean as age. Sometimes the age value of the respondents differs from the age values of the local register. The reason could be that the interviewer asked the wrong person in the household.

following Figure 2 shows the distribution of the weights.

Figure 2: Frequencies of the optimal weights



The minimum, maximum, mean and standard deviation of the weights are 0.5889, 1.5543, 1.0000, 0.1562.

5 Results

There exist only minor differences in the distributions of the auxiliary variables age, gender, nationality, regional characteristics and the BIK-variable between the gross sample s and the net sample r. This is an indication for the relative high quality (unbiasedness) of the net sample with regard to these variables. Therefore, the adjustment results in only little changes - and mostly in the expected direction.

For several survey variables we compared the portions of the unweighted and the weighted sample. This way we show the effect of the weighting procedure. In general, the differences between the weighted and the unweighted variables are small, in most cases less than one percent (see Table 1) and therefore negligible.

Table 1: Unweighted and weighted distributions for selected ALLBUS variables

ALLBUS Variable	Unweighted	Weighted	Difference col.2-col.3
V10	Women should care for children		
Miss.	3.16	3.18	-0.02
Agree strongly	23.19	24.35	-1.16
Agree	26.23	26.55	-0.32
Disagree	27.52	26.69	0.83
Disagree strongly	19.90	19.23	0.67
V141	Sex		
Male	50.12	48.31	1.81
Female	49.88	51.69	-1.81
V142	Education		
Miss.	0.21	0.23	-0.02
No formal education certificate	2.33	2.33	0.00
Lower secondary school certificate	47.84	48.97	-1.14
Intermediate secondary school certificate	24.77	24.34	0.43
Specialized Abitur	5.75	5.63	0.12
Abitur	17.86	17.39	0.47
Other certificate	0.42	0.43	-0.01
Pupil	0.83	0.69	0.14
V155	Occupational status		
Full-time employed	48.79	46.02	2.78
Half-time employed	6.70	6.74	-0.04
Part-time employed (not main job)	4.75	4.78	-0.04
Not employed	39.76	42.46	-2.70
V183	Marital status		
Married, live together	63.16	63.63	-0.48
Married, live apart	1.71	1.68	0.02
Widowed	7.66	9.15	-1.49
Divorced	4.95	5.18	-0.23
Single	22.52	20.35	2.18
V263	Number of persons living in the household		
1 Person	16.78	17.94	-1.16
2 Persons	35.05	36.79	-1.74
3 Persons	21.23	20.33	0.91
4 Persons	17.99	16.74	1.25
5 Persons	6.00	5.58	0.42
6 Persons	1.87	1.67	0.20

Table 1 (continued):

Variable	Unweighted	Weighted	Difference col.2-col.3
7 Persons	0.79	0.70	0.09
8 Persons	0.17	0.16	0.01
9 Persons	0.08	0.07	0.02
18 Persons	0.04	0.04	0.00
V354	Number of inhabitants in community		
To 1.999 inh.	4.91	4.51	0.40
2.000 - 4.999 inh.	9.20	8.30	0.91
5.000 - 19.999 inh.	24.65	23.62	1.03
20.000 - 49.999 inh.	16.61	16.60	0.01
50.000 - 99.999 inh.	8.24	8.65	-0.41
100.000 - 499.999 inh.	21.11	21.87	-0.76
500.000 inh. and more	15.28	16.45	-1.17
V355	BIK - Type of community		
Type 1 (small urban areas)	4.91	4.51	0.40
Type 2	8.08	7.20	0.88
Type 3	16.57	15.49	1.08
Type 4	9.16	9.14	0.02
Type 5	1.04	0.93	0.12
Type 6	3.29	3.57	-0.28
Type 7	5.70	6.32	-0.62
Type 8	9.49	9.12	0.37
Type 9	12.53	11.76	0.77
Type 10 (centers of large cities)	29.23	31.96	-2.74
	Number of persons 18 years and older living in the household		
Miss.	0.12	0.10	-0.02
1 Person	66.74	69.12	-2.38
2 Persons	16.49	15.41	1.08
3 Persons	11.74	10.82	0.92
4 Persons	3.62	3.32	0.30
5 Persons	1.17	1.10	0.07
6 Persons	0.08	0.08	0
7 Persons	0.04	0.04	0

In the following we discuss only those differences between the unweighted and weighted portions of our variables of interest where the differences are greater than one percent. First we present for some demographic variables how the weights change their distributions.

The portion of

- men decreases
- people with lower secondary school certificate increases
- full-time employed persons decreases
- not employed people increases
- widowed people increases
- singles decreases
- people in households with one person increases
- people in households with two persons increases
- people in households with four persons decreases
- people in communities with 5.000-19.999 inhabitants decreases
- people in communities with 500.000 and more inhabitants increases
- people in rural areas (BIK Type 3) decreases
- people in the centers of urban areas (BIK Type 10) increases
- people in households with only one person older than 17 years increases
- people in households with two persons older than 17 years decreases

in the weighted sample compared to the unweighted net sample.

Most of these differences can be explained by the adjustment variables gender and BIK. As we have already mentioned in the net sample the portion of men is somewhat higher than in the gross sample. The adjustment of gender affects such variables as marital status, occupational status and household size because females in Germany are more frequently widowed, not employed, and living alone in a household than men. The fact that the portion of people living in households with two persons increased and people in households with two persons older than 17 years decreased simultaneously can be explained as follows: It seems that single adults living with one child (mostly women) were underrepresented in the net sample. It may be that those persons form a typical group of nonrespondents. A not desirable result is the decreasing of the portion of singles in the weighted sample because this portion is in the population higher than in the unweighted sample[2]. The reason for this is the substantially smaller portion of female singles in the net sample compared to male singles.

The adjustment of the variable BIK has the effect that the portion of people in large cities, especially in their centers, increased. People living in these areas are another group with a relatively high portion of nonrespondents.

The differences in other portions of variables like denomination, union membership and

2 In 1995 the portion of single persons in West Germany was 23.5% for people 18 years and older (Statistisches Bundesamt, p.52). However, it has to be considered that this information of the Federal Statistical Office does not include foreigners.

German nationality between unweighted and weighted data are negligible. Therefore, we did not include the numbers in Table 1.

We also checked the differences between unweighted and weighted data for several survey variables which measure political and other attitudes. The portions differed only in one case by more than one percent: In the weighted data set the portion of people who strongly agree with the statement "It is much better for all persons concerned if the man preferred his job and the woman stayed at home and cared for children" is somewhat higher. That may be an effect of the adjustment of the variable age: In the gross sample the mean age of all respondents was 47.1 in contrast to 45.4 in the unweighted net sample. That means that after weighting the respondents are "older" and therefore more conservative. In all other variables (Left-Right-Scheme, Inglehart-Index, political interest, subjective social class) the differences are not worth mentioning.

6 Conclusions

The conditional minimax principle enables the researcher to treat nonresponse in a flexible way by using prior knowledge from the gross sample. He has several possibilities to incorporate the knowledge into his model. First of all he should summarize his knowledge into a matrix Q. The columns of Q may contain the values of a metric variable or a column of a design matrix of a categorical variable. If these values are not known for all respondents in the gross sample they should be estimated by a suitable procedure such as imputation to avoid missing values. In addition, Q must be of full rank. An important factor is the determination of the parameter space determined by a singular matrix V which describes the admissible variation of the values of a variable of interest. A class of such matrices V was given which involves a diagonal matrix G and a matrix R with R'Q as identity matrix. In order to take into consideration the probability of participation or the probability of contact for each respondent, the researcher could incorporate different diagonal elements into G. $R=AQ(Q'AQ)^{-1}$ is only one possibility for choosing R where A is an arbitrary non-singular matrix. Since the computation of b^c does not lead to nonnegative weights necessarily, the selection of A may help to avoid negative or other extreme weights. Further investigations are necessary to clarify this point.

Another important advantage of the presented idea is that it can be easily transferred to the estimation of parameters in subgroups.

The application to data of the ALLBUS 1996 yields different results: The weights assigned to most of the elements of the net sample are in the small range from 0.8 to 1.2 . The reason for this are the relative small differences in the distributions of the auxiliary variables between the net sample and the gross sample of the ALLBUS 1996. In cases where the difference between the net sample and the gross sample is not so small the effects of weighting are stronger. In our application we stated only minor changes in the survey variables due to weighting - and in the expected directions.

On condition that the gross sample is drawn as a random sample that represents the population well, and auxilary information on the individuals of the gross sample is available, our suggestion for the estimation of population parameters seems to be promising.

References

Bickel, P.J. and Lehmann, E.L. (1981). A Minimax Property of the Sample Mean in Finite Populations. Annals of Statistics 9, pp. 1119-1122

Chaudhuri, A. and Stenger, H. (1992). Survey Sampling. New York: Marcel Dekker, Inc.

Davis, J.A., Mohler, P.Ph. and Smith, T. (1994). Nationwide General Social Surveys. In: Borg, I. and Mohler, P.Ph. (eds.): Trends and Perspectives in Empirical Social Research. Berlin: de Gruyter

Deville, J.-C. and Särndal, C.-E. (1992). Calibration Estimators in Survey Sampling. Journal of the American Statistical Association 87, pp. 376-382

Deville, J.-C., Särndal, C.-E. and Sautory, O. (1993). Generalized Raking Procedures in Survey Sampling. Journal of the American Statistical Association 88, pp. 1013-1020

Elliot, D. (1991). Weighting for Nonresponse - a Survey Researchers Guide. OPCS

Elliot, D. (1996). The Effectiveness of CALMAR in Reducing Nonresponse Bias on the UK Family Expenditure Survey. Paper presented at the 7th International Workshop on Household Nonresponse, Rome, Oktober 1996

Gabler, S. (1988). A Conditional Minimax Approach in Survey Sampling. Metrika 35, pp. 211-221

Gabler, S. (1990). Minimax Solutions in Sampling from Finite Populations. Lecture Notes in Statistics 64, New York: Springer Verlag

Gabler, S. (1991). Eine allgemeine Formel zur Anpassung an Randtabellen. ZUMA-Nachrichten 29, pp. 29-43

Gabler, S. and Häder, S. (1997). Wirkung von Gewichtungen bei Face-to-Face und Telefonstichproben. In: Gabler, S. and Hoffmeyer-Zlotnik, J.H.P. (eds.): Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag, pp. 221-245

Häder, S. and Gabler, S. (1997). Deviations from the population and optimal weights. In: Saris, W.E. and Kaase, M. (eds.): Eurobarometer – Measurement Instruments for Opinions in Europe. ZUMA-Spezial Band 2, pp. 37-52

Hájek, J. (1981). Sampling from a Finite Population. New York: Marcel Dekker, Inc.

Little, R.J.A. (1989). Survey Inference with Weights for Differential Sample Selection on

Nonresponse. ASA 1989 Proceedings of the Section on Survey Research Methods, pp. 62-69

Rao, C.R. and Mitra, S.K. (1971). Generalized Inverse of Matrices and its Applications. New York: John Wiley and Sons, Inc.

Schnell, R. (1996). Development of Nonresponse in Germany 1953-1994. Paper presented at the 7th International Workshop on Household Nonresponse, Rome, Oktober 1996

Statistisches Bundesamt (1995): Demographische Standards. Methoden-Verfahren-Entwicklungen

Stenger, H. (1979). A Minimax Approach to Randomization and Estimation in Survey Sampling. Annals of Statistics 7, pp. 395-399

Stenger, H. and Gabler, S. (1996). A Minimax Property of Lahiri-Midzuno-Sen's Sampling Scheme. Metrika 43, pp. 213-220

Wynn, H.P. (1976). Connected Finite Population Sampling Plans. Biometrika 63, pp. 208-210