Dealing with incomplete household panel data in microsimulation models

by Joachim R. Frick[,1,2,3], Markus M. Grabka[1,2] and Olaf Groh-Samberg [1]*

Paper prepared for the
2nd General Conference of the International Microsimulation Association (IMA)
"Microsimulation: Bridging Data and Policy"
Ottawa, Canada, June 8-10, 2009

Berlin, May 2009

[1] SOEP at DIW Berlin. [2] Berlin University of Technology. [3] IZA Bonn.
* Corresponding author <jfrick@diw.de>
Mail Address: c/o DIW Berlin, Department SOEP, Mohrenstr. 58, 10117 Berlin, Germany.

Abstract:

Dynamic micro-simulation models clearly profit from the availability of longitudinal data collected in household panel surveys, where interviewing *all* adult household members allows to better aggregate household resources as well as to model household decisions. However, such household panel surveys around the world are affected by partial unit-non response (PUNR), i.e. unit-non response of at least one member of an otherwise cooperating household. In those cases, the *aggregation of income* across all household members falls just short of at least one individual's income. These processes are typically not at random, thus results of micro-simulation models, including any policy recommendation derived from those, are subject to bias.

Using data on more than twenty waves of the German Socio-Economic Panel (SOEP) we evaluate four different strategies to deal with this phenomenon: (a) *Ignorance*, i.e., assuming the missing individual's income to be Zero. (b) *Adjustment* of the equivalence scale to account for differences in household size and composition. (c) *Elimination* of all households observed with PUNR with subsequent re-weighting procedures of those observations not affected. (d) *Longitudinal Imputation* of the missing income components. In order to assess the impact the choice of the technique has on a fictitious policy change we simulate an increase in (non means-tested) child allowances and evaluate the resulting effect on poverty separately for each of those four approaches. We find indication that especially ignoring PUNR overstates child poverty as well as biases the poverty reduction effect in contrast to a more complex imputation approach. Obviously, these findings are even more important in cross-nationally comparative micro-simulation models, if national data providers deal differently with PUNR in the underlying data.

Keywords: Household Panel data, Partial Unit Non-Response, Inequality and Mobility, Poverty, Longitudinal Imputation, SOEP

JEL-Codes: D33, I32, C81

Content

1 Motivation 1

2 Incidence and selectivity of PUNR in the German Socio-Economic-Panel (SOEP) 4

2.1 The data 4

2.2 Incidence and selectivity of PUNR 5

2.2.1 Incidence of PUNR 5

2.2.2 Selectivity of PUNR 8

3 Dealing with PUNR in income-based analyses of economic well-being 10

3.1 Alternative approaches 10

3.2 The imputation strategy 12

3.2.1 Imputation of single income categories 12

3.2.2 Imputation of receipt and of amount received 13

3.2.3 Results of the imputation process 16

3.3 Intermediate conclusions and hypotheses 18

4 Empirical analyses: Inequality and Mobility effects 19

4.1 Inequality and Poverty 20

4.1.1 Hypothetical effects 20

4.1.2 How does PUNR affect income levels, inequality and poverty? 20

4.1.3 Consistency of poverty “assignment” when using different approaches to tackle PUNR 23

4.2 Poverty and Income mobility 26

5 Application: Simulating a change in child benefits using different treatments of PUNR 29

6 Conclusions 31

7 References 33

List of Figures & Tables

Figure 1: Incidence of PUNR in the German SOEP, 1984-2007 5

Figure 2: Incidence of PUNR by number of adult household members 7

Figure 3: Incidence of PUNR by panel duration 8

Figures 4.a-c: The impact of imputation: population share holding income components and average income 17

Figure 5: The impact of PUNR-treatment on the distribution of equivalent income 21

Figure 6: The impact of PUNR-treatment on income inequality 22

Figure 7: The impact of PUNR-treatment on relative income poverty 23

Figure 8a,b: Consistency of poverty measurement using alternative approaches to handle PUNR 25

Figure 9: The impact of PUNR-treatment on poverty mobility 27

Figure 10: The impact of PUNR-treatment on 2-wave income mobility 27

Figure 11: The impact of PUNR-treatment on 5-wave income mobility 28

Figure 12: Simulated income effects by PUNR-treatment 29

Figure 13: Simulated poverty effects by PUNR-treatment 30

Table 1: Probability of PUNR – Results from a pooled probit regression 9

Appendix A: Excerpt from the ECHP documentation on the PUNR-adjustment procedure 35

Appendix B: Results from a simulated increase in child benefits, 2007 36

1 Motivation

One of the standard assumptions in welfare economics is pooling and sharing of resources across individuals living together in a needs unit, mostly a private household. Not only due to data limitations this approach, which is generally accepted in inequality research, argues that all household members pool and share all available resources (i.e., income) so that everyone‘s standard of living in the household is the same (e.g. Canberra Group 2001, Atkinson & Bourgignon 2000, Smeeding & Weinberg 2001).[1] This requires that all incomes received in a given household are aggregated across all members and the total sum is reassigned to all of them. Typically, an equivalence scale is then applied in order to adjust for differences in household composition and size, thus allowing for economies of scale in larger households as well as variation of needs across age groups (see e.g. Buhmann et al 1988). This approach crucially depends on either one household representative (e.g. the household head) providing complete proxy information on behalf of all members (similar to the approach used in the US Panel Study of Income Dynamics, PSID, which was started in 1968) or, as is common in almost all household panel surveys started thereafter, all adult household members actually providing an interview themselves. However, various sorts of non-response behavior are reality in most of those surveys posing a massive threat to the implicit assumption of (representative) full coverage of all the resources and needs of individuals living together in one household[2]. In such cases, generally unit non-response (UNR, i.e., the refusal of a complete household) is taken care of by means of proportionally weighting successfully surveyed observations, while item non-response (INR) is corrected for by either weighting or by imputation. But there remains the problem of partial unit-non response (PUNR) in those households where at least one member refuses to participate while other members do cooperate. In fact, there appears to be a general trend of an increasing share of households being affected by partial unit-non response (PUNR) in population surveys and ignorance over this phenomenon may give rise to several problems: a) misreporting of income aggregates; b)

1 This approach is in line with the unitary model as proposed by Becker (1991), which assumes a set of coherent preferences across all household members. However, it should be noted that there is quite some literature arguing against the general validity of this assumption (see e.g. Vogler & Pahl 1994, Bonke & Uldall-Poulsen 2007).

2 For a comprehensive discussion of non-response behaviour in household surveys in general as well as in panel surveys in particular, see the various contributions in the readers edited by Groves & Cooper 1998, Groves, Dillman, Eltinge & Little 2002, De Leeuw et al 2008 and Lynn 2009.

increasing bias in results on income inequality, poverty and mobility, as well as c) bias for analyses on the intra-household distribution.

In the welfare oriented literature one can find various ways on how to deal with this phenomenon when it comes to generating the outcome variable relevant for inequality and poverty analysis, namely "equivalent household income": (a) Ignoring the fact that a household member (and its income information) is missing, thus assuming the non-responding individual's income is Zero. This does not rule out that the person is living only on household based transfers (public and private) which can be assumed to be measured correctly and comprehensively from other household members. (b) Adjusting the calculation of the equivalence scale by ignoring the person's contribution to household income as well as to household needs, thus in principle ignoring the person's existence. This approach implicitly assumes that the incomes of other household members are independent from the income of the missing individual. (c) Eliminating all households observed with PUNR from the analysis population, thus assuming that these households are missing completely at random. An extension of this approach would try to compensate for the obvious embedded selectivity by means of weighting, thus assuming that these households are missing at random. This can be accomplished by crossing up the share of those at risk of PUNR according to the share of households actually affected by PUNR. I.e., for all households which are not at risk of PUNR, because there is only one respondent, the weighting factor remains unchanged. Obviously, this weighting strategy can be more or less complex. (d) Imputing the missing income components at the individual level and aggregating across all household members (including those with PUNR).

At first glance, the assumptions underlying options (a) through (c) are most likely very strong given the (potential) selectivity issues involved in the missing mechanism. Options (c) [with re-weighting] and (d) appear to be less selective, and in principle option (d) has the advantage of maintaining the entire survey population. Having said that – and as is true in any case of imputation –, there is quite some normativity involved in the actual implementation of any such imputation process. For example, one may argue that the degree of misspecification is more like a general underreporting that can be corrected for by either adjusting the household income by means of a "(relative) factor" or by adding an "(absolute) flat sum". More appropriately, one may allow for more variation with respect to the contribution of various income components to the overall household income measure and by controlling household and individual characteristics related to the missing mechanism. E.g., the severity of misreporting is most likely very different for a household where the 85 year old mother of the household head is missing due to lingering illness as compared to a household where the main income earner is absent because he/she is drilling for oil on an offshore platform.

Thus, results from research on equivalent household income inequality and even more so on relative poverty may crucially depend on the choice of the aforementioned option, because this decision will affect (a) the income of all individuals living in households which happen to be concerned by PUNR, and (b) the relative poverty line to be derived from the national mean or median. Having said that, any such bias in cross-sectional income measures will (c) also affect mobility analyses based on those measures. As a consequence, any microsimulation analysis based on household panel data, e.g. on the monetary effects of welfare reforms, might be affected by the treatment of partial unit-non-response implemented in the database.

Using more than twenty waves of microdata from the German Socio-Economic Panel (SOEP) this paper assess the potential impact of dealing with PUNR on the results of microsimulation analyses focusing on economic well being. The paper is set up as follows: In Section 2 we first describe incidence and trends of PUNR in the SOEP over the period 1984 to 2007 before turning to the analysis of selectivity of PUNR. Here we control for the relevance of concurrent household characteristics (e.g. size and composition) as well as for individual characteristics of the missing person. Section 3 presents the various techniques to deal with PUNR. We especially elaborate on the principles of our three-stage imputation strategy of income components at the individual level. In Section 4 we provide sensitivity analyses showing the variation in the results for income inequality and poverty when choosing between the aforementioned options. Making explicit use of the panel nature of the underlying data, we also point to the relevance of PUNR (and the way it is treated in the microdata) for poverty dynamics and income mobility. Apparently, dynamics are exaggerated if PUNR is present in at least one wave. Section 5 is devoted to a comparison of the results obtained from standard static microsimulation models based on those four data versions. We show the impact the choice of the technique has on a simulated policy change. Following current discussions in Germany on the poverty reducing effect of child benefits we simulate an increase in (non means-tested) child allowances. We find indication that ignoring PUNR overstates the poverty reduction effect, while this bias is reduced for in the more complex imputation approach. Finally, Section 6 concludes with some remarks on the potential relevance of our findings for cross-national comparability of research on income inequality, poverty and mobility. Obviously, our findings are most important in comparative micro-simulation models (e.g. EUROMOD), especially if national data providers deal differently with PUNR in the underlying data.

2 Incidence and selectivity of PUNR in the German Socio-Economic-Panel (SOEP)

2.1 The data

The German Socio-Economic Panel (SOEP) is a representative longitudinal survey of individuals living in private households in Germany (Wagner et al. 2007). The survey was started in 1984 in West Germany and was extended to East Germany in June 1990, somewhat more than half a year after the fall of the Berlin Wall. The initial sample included over 12,000 respondents, with everyone aged 17 and over in sample households being interviewed. In recent years, various representative new sub-samples have been drawn, which approximately doubled the initial sample size. Other additional samples were explicitly designed to cope with specific subgroups of the population. In 1995, the SOEP introduced an oversampling of immigrants to cope with the phenomenon of misrepresentation of recent immigrants in ongoing panel surveys. Due to the high concentration of economic resources (income and wealth) at the top of the distribution, welfare analyses based on representative population surveys are often confronted with the lack of information on “rich” individuals. In order to overcome this problem, the SOEP introduced a high income sample in 2002, over-representing the top 3% of the income distribution—this sample is thus included in the more recent years of our time series. The sample analyzed below employs all available observation years up to survey year 2007.

One of the main problems when asking for (specific) income and wealth information in any population survey is non-response, and SOEP is no exception to this rule. Making effective use of the panel nature of SOEP, any item non-response is corrected for by applying longitudinal row-and-column imputation procedures (see Little and Su 1989) and purely cross-sectional imputation techniques if longitudinal information is lacking. Thus, at least eventual bias arising from the above-mentioned selectivity can be reduced (see Frick and Grabka 2005).

Finally, it should be noted that all of the following analyses refer to the population in private households only, i.e., we exclude individuals living in institutions such as nursing homes.

2.2 Incidence and selectivity of PUNR

2.2.1 Incidence of PUNR

Incidence of PUNR at person vs. household level

There are at least two ways to express the overall incidence of PUNR in a household panel survey. E.g., while in the survey year 2005 “only” 5.4% of all adult household members did not fill in the requested questionnaire themselves, their non-participation behaviour affected the measurement of relevant outcomes for all their co-residents in their respective household as well, thus 11.5% of all individuals (including children who not yet reached the respondent age of 17) lived in a household which is affected by PUNR.

Figure 1: Incidence of PUNR in the German SOEP, 1984-2007

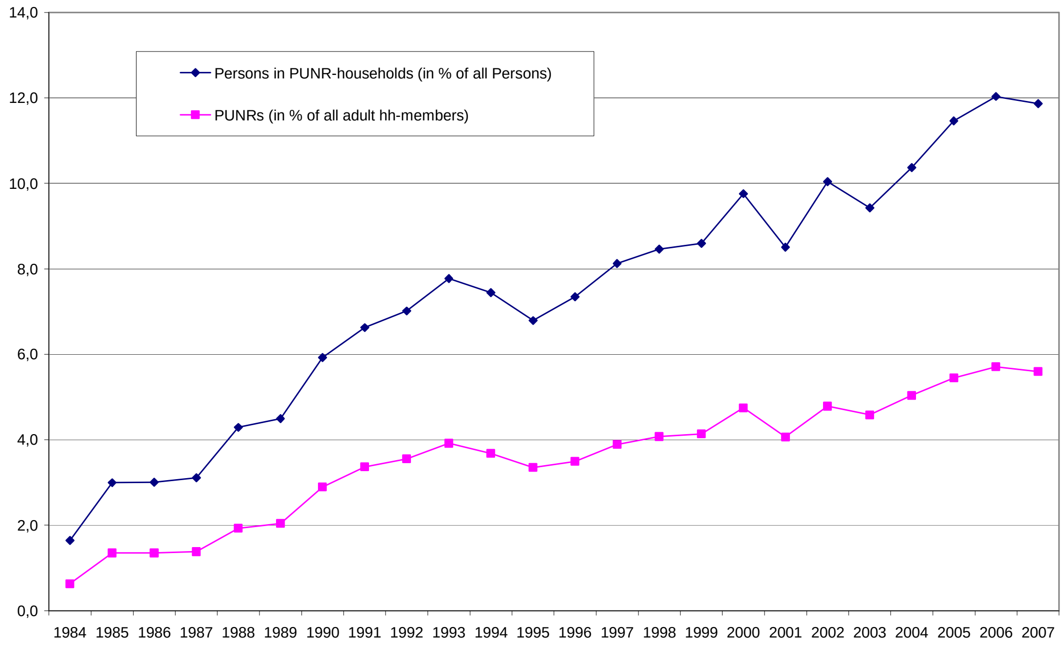


Figure 1 presents time series information on the incidence of PUNR in the SOEP data over the period 1984 through 2007. While there was almost no PUNR in the starting wave 1984 (the contract with the fieldwork agency stated explicitly that in the first wave households had to be “completely” interviewed with only a few strict exceptions allowed), there is a clear tendency towards increasing shares of non-participating respondents since then. This process is even more striking given the secular trend towards smaller households, thus the population at-risk for PUNR is shrinking due to the increasing share of singles and lone-parents (with minor children up to 16 years of age which is the respondent age in SOEP), i.e., household types for which non-response of the only respondent living in that household by definition yields a complete drop-out (= UNR, Unit Non-response). Of course, the increase in the incidence of PUNR is also driven by the accumulation of “old” PUNRs over time, i.e., persons who basically never give personal interviews while other household members continue to do.

Incidence of PUNR by household size

Following from this, a straightforward way to identify potential selectivity in PUNR comes with the number of adults (= target respondents) living in a given household. Obviously, those households with only one respondent do not bear the risk of PUNR. The risk of the household unit being affected by PUNR, however, clearly increases with the number of respondents, thus a household of six is more likely to be affected by PUNR than a household of two adults only. While children below respondent age themselves do not bear the risk of PUNR, they do however may be affected by PUNR of adult household members, and any misrepresentation of those persons resources in the household aggregate will have relevance for measures of child poverty.

Figure 2: Incidence of PUNR by number of adult household members

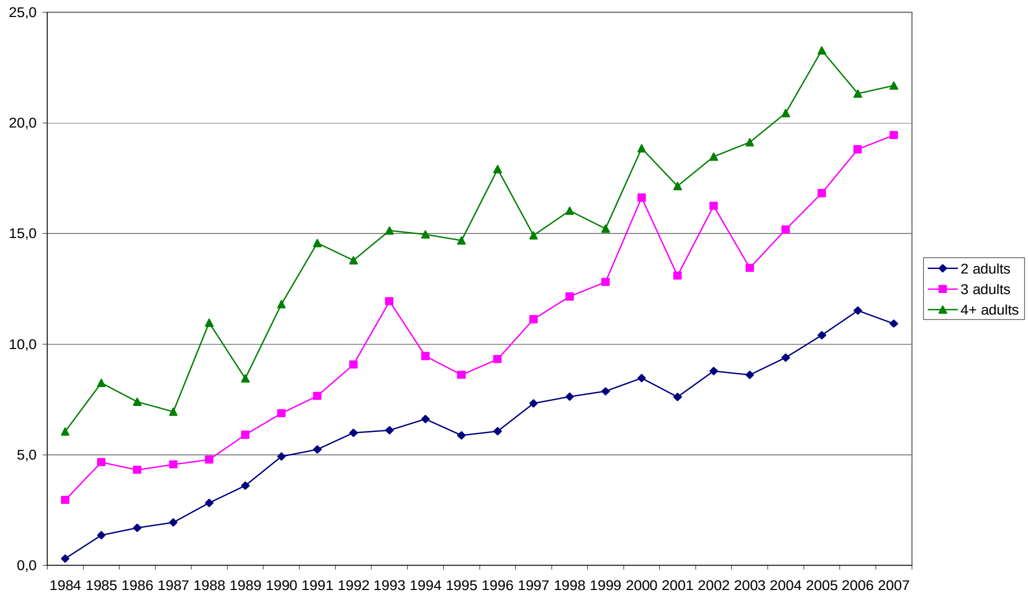
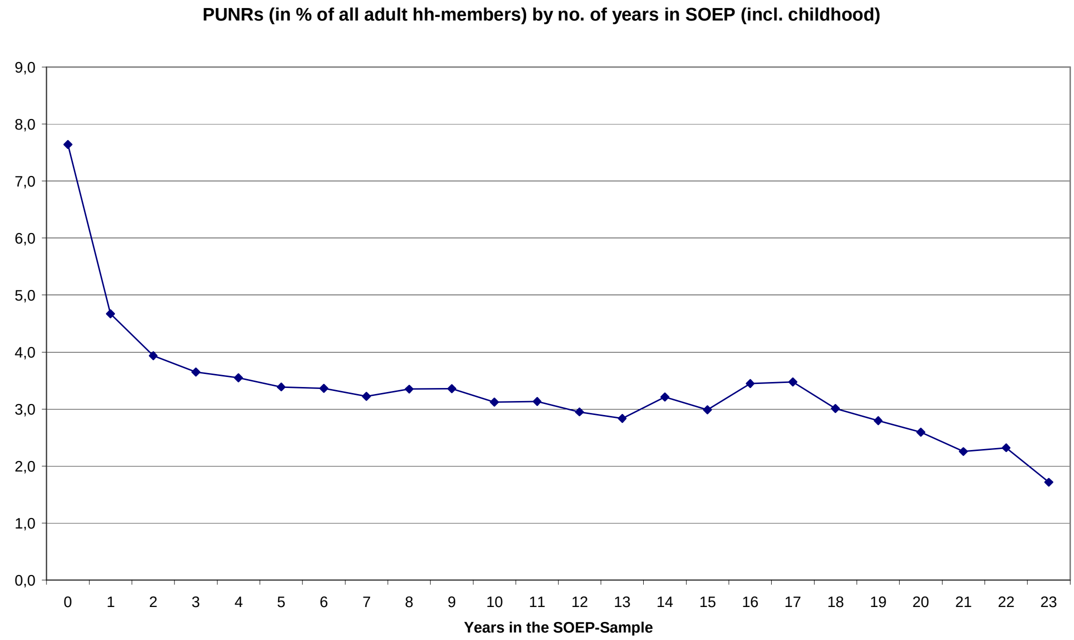


Figure 2 clearly illustrates this effect by indicating consistently increasing shares of individuals being affected by PUNR either directly (due to own non-participation) or indirectly (due to non-responding co-residents): For example in 2005, the share of individuals being affected by PUNR within the household context is around 10% in households with two persons of respondent age, about 17% for those living in households with three respondents and even above 20% for the rather few observations living in large households of four and more adults.

Incidence of PUNR by panel experience

For a long-running household panel such as SOEP, any non-response behaviour is crucial for maintaining the quality and representativeness of the *longitudinal* data. A general finding in research on scope and selectivitiy of UNR in household panel studies provides a rather robust picture for the probability of dropping-out being decreasing in panel experience, thus, any additional interview reduces the probability of UNR (see e.g. Watson & Wooden 2009 for the case of the Australian HILDA Survey). Although with respect to UNR there is clear empirical support for this hypothesis in the SOEP data (a most relevant result for designing the weighting scheme in SOEP, see Kroh & Spieß 2007), the probability of PUNR does not necessarily monotonically diminish for long term respondents (see Figure 3). We find a clear reduction in this probability only over the first years of the panel experience (including the years that a person "grew" up in the survey during childhood without being a respondent herself) and again after some 20 years.

Figure 3: Incidence of PUNR by panel duration



It may be rather the case that a temporary drop-out of a respondent in an otherwise co-operative household is just more likely to happen over a period of twenty years than over only five years. As such, a temporary drop-out of a *long-term* respondent may be less selective, i.e., it reduces the number of observations in a balanced panel sample, thus reducing efficiency, but does not cause substantive bias in the results. However, an alternative hypothesis – not tested in this paper – could be that just because of the well-established interviewer-respondent relationship of long-term panel members, these individuals may be more likely not to participate in times when "unusual" events occurred which they do not want to communicate with the interviewer (e.g. unemployment for a long-term successful manager). This second hypothesis is more in line with findings by Kapteyn et al. (2006) who argue that attriters in the Health and Retirement Survey (HRS) who are recruited back into the survey are very different from those who are permanent attritors.

2.2.2 Selectivity of PUNR

Above and beyond the finding that individuals in "younger" subsamples of SOEP are more likely to produce PUNR, further descriptive analyses (not shown here in detail) point towards more selectivity underlying the missing mechanism: the probability of PUNR increases in household size, middle age groups and men are more likely for PUNR than the elderly and women, respectively. Finally, any third household members are found more often among the PUNR than household head's partners or children. These findings may be taken as indication for underreporting of economically active household members and thus most likely for understating major contributions to overall household resources.

Table 1: Probability of PUNR – Results from a pooled probit regression

		Coeff.
sex	female	-0.443***
age groups (ref. 56-65 yrs of age)	17-24	-0.003
	25-40	-0.053*
	41-55	-0.210***
	66 and over	-0.304***
relation to household head (ref: household head)	partner	0.977***
	child	0.984***
	other	1.700***
migration background (ref: autochthonous)	native born foreigners	0.204***
	foreign born Germans	-0.264***
	foreign born foreigners	-0.347***
East Germany (ref: West Germany)	East Germany	-0.245***
change in household composition (ref: no change)	change	0.076***
	first wave	0.528***
no. of adults in household (ref: 2)	3 adults	0.052*
	4+ adults	0.125***
no. of children in household (ref: none)	1 child	-0.078***
	2 children	-0.163***
	3+ children	-0.170***
home owner (ref: tenant)	home owner	-0.051**
highest education of hh head/partner (ref: intermediate)	lower secondary	0.039
	higher secondary	-0.102***
	tertiary	-0.124***
item non-response monthly household income (ref: no)	yes	0.396
Control dummies for survey year		yes
Constant		-1.779
Pseudo-R2=		0.1568
Observations / Individuals		325414 / 45038

Robust standard errors. * significant at 10%; ** significant at 5%; *** significant at 1%

Source: SOEP 1985-2007. Authors' calculations.

Given that some of these characteristics are well correlated with each other we try to identify correlates (and eventually determinants) of PUNR by means of multivariate analyses simultaneously controlling for various potential factors. Table 1 shows results from a pooled probit regression model based on more than 45,000 individuals being observed in the SOEP over the period 1985 to 2007. This totals more than 325,000 person-year observations, thus we use robust standard errors obtained from clustering at the level of individuals. The descriptive findings mentioned above are by and large confirmed in our estimation: women are less likely for PUNR, which is also true for the middle aged (25-40) and the elderly (66 years and over). Other characteristics associated with a reduced probability for PUNR include home-ownership, increasing number of dependent children as well as increasing educational level. On the other hand, the risk for individual non-cooperation is as expected increasing in the number of adults to be interviewed. Last but not least, we find item-non-response (INR) on the question targeting at the current monthly household income to be significantly related to partial unit-non-response of at least one household member, which is in line with findings based on data from various panel surveys showing a positive impact of such INR on income questions in wave *t* on the probability of UNR (i.e., attrition) in wave *t*+1 (see e.g., Loosveldt et al. 2002, Frick & Grabka 2009).

3 Dealing with PUNR in income-based analyses of economic well-being

Keeping those selection issues in mind, the following section briefly introduces the various approaches to deal with PUNR before applying some prototypical empirical applications from a welfare economics point of view based on an aggregated measure of equivalent household income.

3.1 Alternative approaches

There exists a variety of ways to deal with PUNR in empirical analyses, out of which four will be used in this paper: *Ignorance*, i.e., assuming the missing individual's income to be Zero. (b) *Adjustment* of the equivalence scale to account for differences in household size and composition. (c) *Elimination* of all households observed with PUNR with subsequent re-weighting procedures. (d) *Longitudinal Imputation* of the missing income components.

- Ignorance: Assume the individual affected by PUNR has no own income to add to the household's overall resources, but she does have needs which ought to be considered when constructing the household's equivalence scale. This means effectively ignoring

PUNR in the measure of household *income*, while maintaining her *needs*, i.e., *Y(PUNR)=0 & Needs(PUNR)>0*.

- Adjustment: Assume the individual has no own income as well as no needs, thus completely ignore the existence of the individual with PUNR. This effectively means deleting non-responding individuals from PUNR-households by "downward adjusting" the respective equivalent scale, i.e., *Y(PUNR)=0 & Needs(PUNR)=0*, which implies that income and needs of the missing individual are identical to those of the observed household members.
- Elimination: Delete *all* individuals living in households affected by PUNR (i.e., also the successfully interviewed persons) and re-scale the population weights for those households who bear a risk of PUNR, but who did take part in the survey completely. This assumes that the income and needs of households with PUNR are mirrored by successfully completed households with two and more respondents.[3]
- Imputation: Impute any income measure missing due to PUNR, thus considering all households with their respective completed income as well as needs, by assigning incomes to PUNRs on the basis of comprehensive (cross-sectional and longitudinal) imputation procedures (details are given in the next section below).
- Finally, another approach – which will not be considered in the remainder of this paper – is the "flat correction factor" which has been applied in the European Community Household Panel (ECHP) (see Eurostat 2000). In this case, each household is assigned a specific "within household non-response inflation factor" (see Appendix A for details). The basic assumption underlying this approach is that all income components in a given household are affected by PUNR in the very same way – thus, even if this was considered a pseudo-correction of the misreported income level, the income portfolio of the household most likely remains subject to bias.

[3] Without re-scaling of weights, this approach is also known as "listwise deletion". It is well-known from the literature that such simple deletion of incomplete cases without any subsequent correction most likely biases the results and increases the inefficiency of the statistics by reducing the sample size (see e.g., Barceló 2008 for the Spanish Survey of Household Finances, Frick & Grabka 2005 for the SOEP, Frick & Grabka 2009 for the UK BHPS, the SOEP and the Australian HILDA Survey, as well as Starick & Watson 2007 also using HILDA data).

3.2 The imputation strategy

The imputation procedure correcting for income missing due to PUNR is based on the following principles: impute as detailed income information as possible (i.e., use different components of income), make use of household context data and, most important from the panel perspective, employ longitudinal information if at all available.

3.2.1 Imputation of single income categories

As is standard in the welfare economics literature our target income measure is annual post-government income of the previous year, which is given by the sum of all market incomes (from labor, capital and private transfers) *plus* pensions and public and private transfers received *minus* taxes and social security contributions paid – aggregated across all household members.[4] However, instead of imputing just a "lump sum" of missing income we target at imputing six individual gross income components which are directly compatible to the more detailed information collected in the standard SOEP questionnaire every year. This allows aggregation at the level of the household as well as the tax unit[5] to match the very same income aggregates for all individual observations, be they PUNR or successfully interviewed. This supports the final simulation of direct taxes and social security contributions which explicitly needs to consider the interdependence of income receipt and tax calculations for joint tax filers. By this, we can derive a consistent measure of "Household Post-Government Income" as the major source for inequality analyses. Finally, this procedure is more likely to exert less bias in portfolio analyses than a "lump sum or flat factor" approach would do.

For each PUNR we impute the following six income components which are collected at the individual level in the SOEP (all other income components such as means tested public transfers or capital income are surveyed at the household level and thus, by definition, are already included in the income measures derived from the successfully interviewed household members):

---

[4] In line with the Canberra Group (2001) recommendations we also add a measure of net imputed rent, which captures the implicit income advantage of owner-occupied housing as well as any non-monetary income advantage of subsidized renters (see Frick & Grabka 2003).

[5] According to German tax laws married couples in principle file their taxes jointly, while all other individual are single filers.

1. labor income (this is the sum of all incomes from dependent employment, self-employment, secondary jobs, including extra payments such as Christmas bonus, gratification, etc.),

2. pensions (this is the sum of all pensions received from the statutory social security pension system (GRV) as well as from the tax-financed pension system for civil servants, in both systems including eventual survivor benefits),

3. unemployment compensation (this is the sum of assistance received via the unemployment insurance scheme, unemployment assistance and subsistence allowance from the labor office),

4. (public) student grants,

5. (public) maternity leave transfers, and

6. private transfers.

3.2.2 Imputation of receipt and of amount received

For each of the components mentioned above, we employ a two-step procedure. First, we need to define a "filter" indicating whether a given person received the respective component, and conditional on predicted receipt, we need to impute a positive value for that income.[6] On both occasions, we make use of longitudinal information, if at all available as suggested by ample evidence available in the literature on imputation in panel studies (see e.g. Spiess & Goebel 2003 using ECHP, Starick & Watson 2007 using HILDA, Frick & Grabka 2009 using SOEP, HILDA and BHPS data). We separate all observations with PUNR into four groups, depending on availability of information from either the previous or the following year, both years or none of them. In fact, for any PUNR with missing information in year *t* and a successful interview in *t-1*, we can derive most valuable and highly predictive information for the target information in *t* from previous year's income receipt at the time of the interview when she was asked about her current income and employment status.

As such, receipt of a given income Y_k>0 (*k* = six income components) is predicted on the basis of a multivariate regression model (probit) estimating the probability of income receipt in the observed population. For observations without any longitudinal information we employ only contemporary control variables including individual information on sex, age,

6 This two-step procedure is necessary in order to avoid imputing too low incomes for too many persons ("regression to the mean"), in particular in cases of rarely received income components like maternity leave benefits or student grants.

relationship to the household head[7], and a range of household context information[8]. For observations with such longitudinal data, we include also income and employment status from the adjacent waves. In any case, if the predicted probability for receipt of a given income exceeds a randomly chosen threshold (drawn form a normal distribution with mean 0.5 and standard deviation 0.2), we assign a value of one to the respective filter indicator. In order to adequately control for the interdependency of income receipt, we include predicted filter dummies in subsequent regressions in the following order: We first estimate the probability for receipt of pensions, then include the predicted pension dummy in the estimation for the receipt of unemployment assistance. Both of these are then considered when running the probit regressions for maternity leave benefits and student grants. All four filter dummies are then included in the regression of private transfers receipt, and finally, we use all five filter variables to predict receipt of labor income. It should be noted that the predictive power of these models is very good, especially when introducing longitudinal data: e.g., the Pseudo-R^2 for estimating the probability of receiving labor income ranges up to .7.

In a second step, we predict the amount of income conditional on predicted receipt as indicated in the filter variables. Here, we make use of longitudinal data in the imputation process by applying the row-and-column imputation as described in Little & Su (1989). Longitudinal information is used over a 7-year (or wave) shifting window around the point in time with missing data – up to three years prior and post the occurrence of PUNR. Assuming that information obtained from observations more distant from the missing data point to be less strongly correlated to the missing information, we give decreasing weights to further apart information.[9]

[7] The quality of correction by means of imputation and weighting may crucially depend on the available data collected on the missing persons by means of proxy information (see e.g. BHPS). A lack of such proxy data may not always be blamed to the survey designers but may simply result from ethical and legal restrictions about collecting proxy data on individuals who are unwilling or unable to participate in a survey. In fact, in the German context and thus in the SOEP, collecting proxy information on the income situation of non-cooperative household members is by and large not feasible.

[8] **Demographic variables** *(D=dummy variable)*: age-groups, family and household type, number of children, at least one household member is in need of care (D), relationship to head of household, region, community size, SOEP-sample identifier, change in household composition (D); **Social structure**: home owner (D), (highest) education of head of household or spouse/partner, education of children, migration background (D); **Income variables**: screener income (net household incomes at month of interview), public transfers (housing subsidies, social assistance, etc.), aggregated observed individual incomes as a share of household screener income; **Filter(s)** considering receipt of other income components (D).

[9] In more detail, the income information is weighted by 2^(3-dt), with dt denoting the distance in years, ranging from 1 to 3.

The row-and-column procedure proposed by Little & Su[10] is set up as follows: The column effect c_t is defined by the relative cross-sectional annual income for each of the seven (k) observations, thus capturing simple period effects: $c_t = Y_t / \frac{1}{k}\sum_{i=1}^{k} Y_i$ (with Y_t denoting total incomes at time *t* and *k* being 7 years). Secondly, the row effect r_i is based on the longitudinal income data collected from a given individual *i*, thus it gives the within-person average income position $r_i = \frac{1}{k}\sum_{j=1}^{k}\frac{Y_{ij}}{Y_j}$ (with *k* denoting the number of valid income information for individual *i* and Y_j denoting the average income at time *j*).[11] Multiplying the column effect c_t with the row effect r_i yields the expected income position of individual *i* at time *t*: $E_{it} = r_i * c_t$. Finally, a stochastic component is introduced by considering the deviation of the expected and the observed value of the nearest neighbor. Combining those three effects yields an estimate for the missing income $Z_{it} = E_{it} * [Y_{nt} / E_{nt}]$ (with Y_{nt} and E_{nt} denoting the observed and the expected income of the individual *n*, defined as the nearest neighbor to individual *i*).

This procedure provides imputed values for all PUNRs with at least one valid interview within the seven year-window under investigation, however, it fails to do so if no longitudinal data is available. This applies to less than one third of all PUNRs and thus requires the application of a purely cross-sectional imputation strategy for the various income components, some of which are closely linked to the life course such as student grants, maternity benefits, and pensions. Conditional on the predicted receipt we impute the money metric value for the six income components, partly also separately by gender. Again, in order to control for eventual interdependence in receipt of the various income sources, the list of regressors includes the full set of dummies for receipt of the respective other (five) components, as well as the same set of variables used to predict the filter information. In order to preserve variation when predicting the respective income value for PUNRs, we again introduce a stochastic component drawn randomly from the residuals of the regression sample.

[10] L&S is also the standard procedure for imputation of item-non-response (INR) on income questions in the SOEP (see Frick & Grabka, 2005).

[11] Note that in the procedure implemented here, r is based on weighted income information as described above.

3.2.3 Results of the imputation process

A straightforward assessment of the overall impact of these imputation procedures is given in Figures 4a-c presenting time series information on a comparison of observed and fully imputed values for the various income components. We show (a) the population share holding a given component, (b) mean values for each component (in nominal Euro) conditional on receipt, and (c) the resulting mean values for the entire population.

According to Figure 4a and in line with the regression results presented above, non-participating individuals are being assigned labor incomes clearly more often than is true among the observed population. In fact, our imputation procedures impute labor incomes for roughly 75% of all individuals with PUNR while only 60% to 70% of the observed individuals report receipt of labor income throughout the previous year. Accordingly, the share of individuals for whom we imputed receipt of any other income components (pensions, unemployment benefits, maternity benefits, student grants, private transfers) is clearly below the level among the successfully interviewed population.

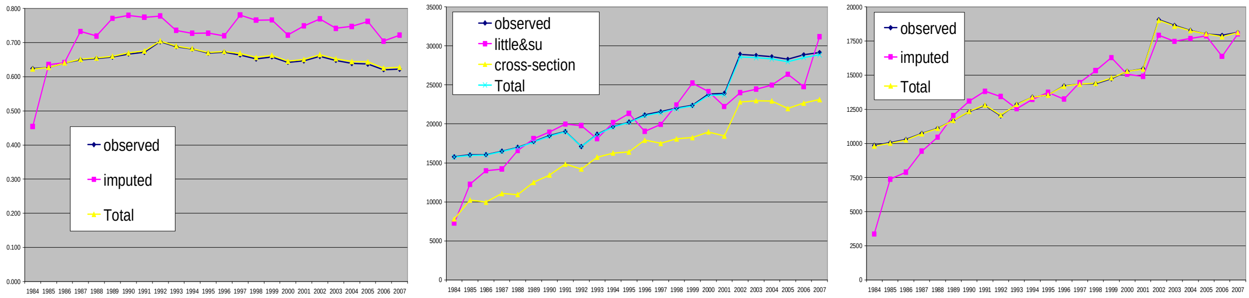
Comparing the levels of income for those who have been either observed or imputed as recipients of a given income component (see Figure 4b), we differentiate between individuals where we could apply the longitudinal imputation according to Little & Su (1989) and those where a purely cross-sectional approach was necessary due to lacking longitudinal information. By and large, for all income components both types of imputation yield similar average values – except for labor income where we can identify a consistently higher average income among the longitudinally imputed individuals. This may reflect in part a statistical artifact since for new entrants into the labor market, who typically have rather low earnings, we do not observe any previous income receipt of that type which is a prerequisite for applying the L&S imputation procedure.

Finally, Figure 4c gives the average values for the various components across the entire population prior and post full imputation. While the general result appears to be that our imputation does not alter those values very much, one should keep in mind that these figures also affect all fellow household members affected by PUNR due to the pooling and sharing assumption underlying the calculation of an equivalent household income.

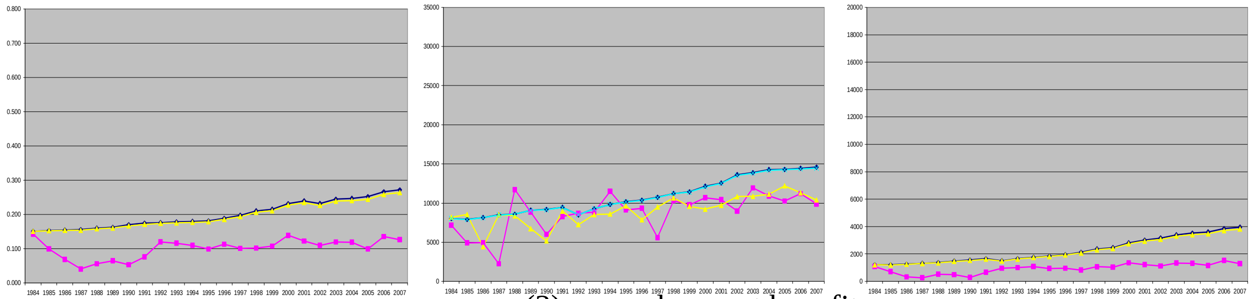
Figures 4.a-c: The impact of imputation: population share holding income components and average income

A. Share of persons with Y>0 | B. Mean values>0 | C. Mean values (incl. 0)

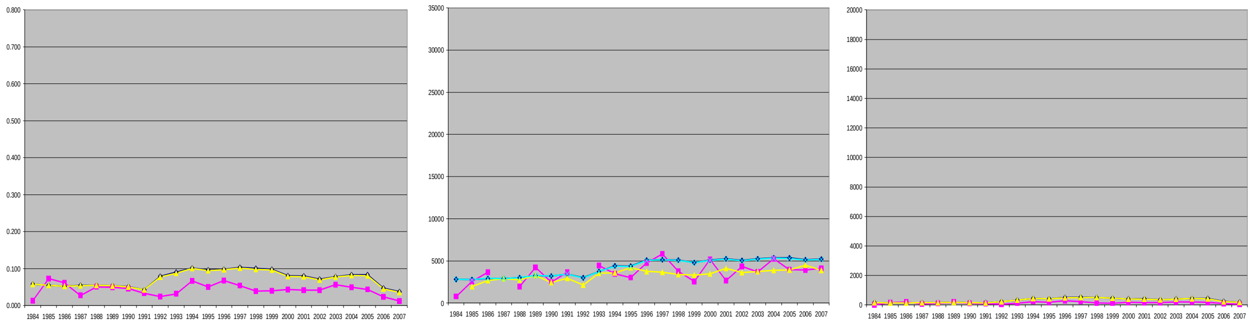
(1) labour incomes



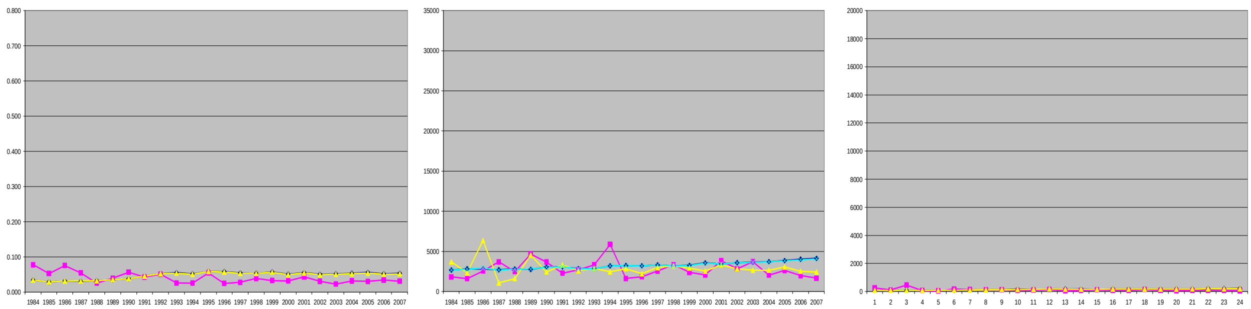
(2) pensions



(3) unemployment benefits



(4) other transfers (maternity benefits, student grants, private transfers)



Source: SOEP. Authors' calculations.

Thus, before running the welfare oriented analyses, we need to incorporate the relevant imputed gross income components into the simulation of direct taxes and social security contributions, then add any public transfers, and finally recalculate a PUNR-adjusted measure of post-government income.

3.3 Intermediate conclusions and hypotheses

What impact does the choice of one of those alternative treatments (ignorance, equivalence scale adjustment, elimination of PUNR plus re-weighting of non-PUNR, imputation) might exert on results of inequality and poverty analyses based on PUNR affected microdata? First of all, given the secular trend of increasing incidence of PUNR over time (see Figure 1 above), one should expect any eventual bias (in measures of inequality, poverty, intra-household distribution, and aggregates) arising from PUNR to be increasing over time, i.e., with duration of the panel. Secondly, the selectivity of PUNR as shown in Section 2.2.2, clearly challenges the basic assumption underlying the various approaches: Version 1 "ignoring PUNR" [thus assuming *Y(PUNR)=0*] clearly appears least plausible given that PUNR appears to be more prominent among economically active household members. In principle, this critique also applies to the basic assumption underlying version 2 "adjusting equivalence scale" where we assume the incomes and needs of the non-cooperative members to mirror those of their respective interviewed fellow household members [Y(PUNRs) ~ Y(noPUNRs) within PUNR-HH]. Finally, there may be no clear-cut answer about distributional effects arising from choosing version 3 "elimination and re-weighting" [i.e., incomes of PUNR-households are equal to those of non-PUNR-households, Y(PUNR-HH) ~ Y(noPUNR-HH)] or version 4 "imputation of PUNR" [*Y(PUNR)=f(X)+e*]. It can be supposed, however, that the more similar the variables used in the re-weighting scheme and in the imputation procedure, the more similar the treatment effects on inequality and poverty might be. In any way, while both approaches appear clearly less problematic than approaches 1 and 2 due to (adequately) controlling for selectivity, the imputation approach may be preferred due to retaining the complete panel population for mobility research.

4 Empirical analyses: Inequality and Mobility effects

The following empirical analyses are based on annual post-government income received in the calendar year preceding the survey year (including a measure of net imputed rent, see Frick & Grabka 2003). For cross-temporal comparability we express all incomes in prices of 2000, also correcting for purchasing power differences between East and West Germany till the mid 1990s. To adjust for different income needs across households due to differences in size and age composition we apply the modified OECD equivalence scale assuming adult household members (above age 14) to have 50% of the income needs of a single person and children up to age 14 to have 30% of those needs. In the following, we will compare results obtained from the four approaches to deal with PUNR on measures of income inequality, poverty and mobility. With respect to poverty measures, we will also try to identify the degree to which results do not only coincide for the entire population, but we will also look at consistency of those alternatively derived measures for each individual. Thus, it will not only be relevant to find out whether two approaches yield e.g. a similar share of individuals being at risk of relative income poverty, but also whether these results are identical for the very same persons.

In order to provide a more robust picture we apply a range of established indicators as used in the literature. We measure income inequality by means of the Gini coefficient, the mean log deviation (MLD) being more sensitive for changes at the lower end of the income distribution and the top-sensitive half squared coefficient of variation (HSCV). Relative income poverty will be measured based on a poverty threshold given by 60% of the national median. In order to also identify eventual effects *within* the population identified as poor, we make use of the family of poverty measures developed by Foster, Greer and Thorbecke (1984) allowing the poverty aversion parameter alpha to take on the values of 0 (poverty risk rate), 1 (normalized poverty gap) and 2 (giving higher weight to those further below the poverty threshold). Finally, income mobility is assessed on the basis of the measures introduced by Fields & Ok (1999) and by Shorrocks (1978). [12]

[12] All empirical results presented in this paper are based on calculations using Stata (version 9.2). We gratefully acknowledge the Stata add-ons INEQDECO, INEQUAL7, IMOBFOK, FOKMOB, SHORMOB authored by Stephen P. Jenkins and Philippe van Kerm, respectively.

4.1 Inequality and Poverty

4.1.1 Hypothetical effects

What are the hypothetical effects of treating PUNR, in whatsoever way, on relative poverty and inequality? First of all, any explicit accounting for PUNR should yield *higher* incomes among those affected by PUNR and thus, average incomes (mean and median) of the entire population should also be subject to increase. Following from that, we should expect an increase in the relative poverty threshold and consequentially, the relative poverty risk among households not affected by PUNR (and thus without a change in their PUNR-adjusted income measure) should be higher than without PUNR-correction, other things being equal. Version 2 (adjusting equivalence scale) and Version 4 (imputation) explicitly yield higher equivalent income among households affected by PUNR. Thus, for this group one should expect, *ceteris paribus*, decreasing poverty risk rates as long as their increase in household income exceeds the increase in the national poverty threshold.

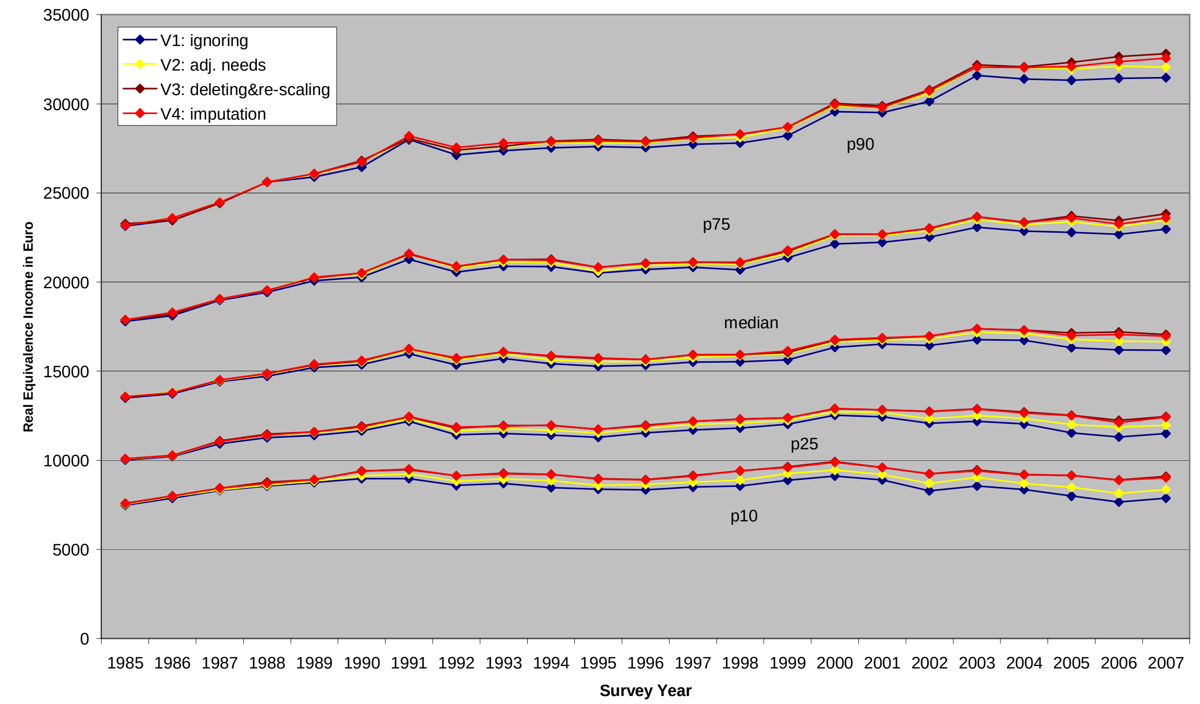
In light of those two contradicting effects the overall (net) effect of PUNR treatment on relative poverty risks at a *given point in time* remains unclear. However, due to the secular increase in the incidence of PUNR over time, *poverty trends* might be affected as well. Most likely however, there will be effects on the *socio-demographic structure* of poor households. Almost by definition, the increase in the poverty threshold will cause an increase in poverty among those households not at risk of PUNR (i.e., single adults and lone parent families with only one household member of respondent age). Similar effects may be expected for households who bear the risk of PUNR but who happened to have been completely interviewed.

4.1.2 How does PUNR affect income levels, inequality and poverty?

Putting numbers to those considerations we now turn to a comparison of results from distribution, inequality and poverty analyses based on the four approaches (labeled *Version 1*: “Ignoring”, *Version 2*: “Adjusting needs”, *Version 3*: “Deleting & re-scaling” and *Version 4*: “Imputation”).

Figure 5 gives time series information on various percentiles (P10, P25, P50=Median, P75, and P90) of the respective annual equivalent post-government income (in €, in prices of 2000). Apparently, in the early waves of the panel, when PUNR was a rather rare event, there is almost no variation across our four measures. However, in line with the increasing incidence of PUNR over time (see Figure 1), we observe a clear and consistent differentiation: Across the entire distribution, the results for Version 1 (Ignoring) are, as speculated in Section 4.1.1., lower than in the other three treatments. These differences clearly pick up over the course of time. Adjusting the equivalence scale (Version 2), which implicitly argues that incomes and needs within PUNR-households are correctly specified by the observed individuals, also controls only insufficiently for the underlying selectivity whereas Version 3 (deleting & re-scaling of weights) and Version 4 (imputation) yield the highest income levels – while being very similar in scope.

Figure 5: The impact of PUNR-treatment on the distribution of equivalent income

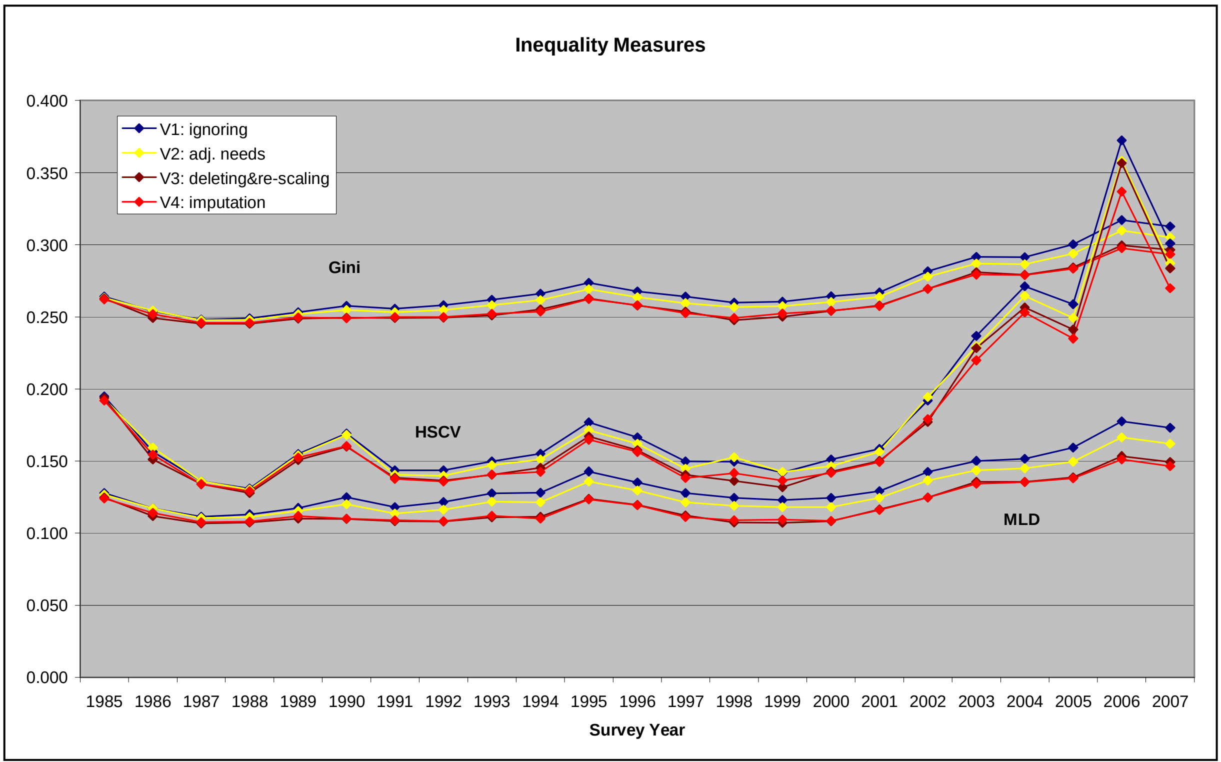


Source: SOEP 1984-2007 (until 1991 West Germany, only). In prices of 2000.

As can be expected from the accretive deviation between top and bottom income levels in Figure 5, there is a secular trend for increasing income inequality in Germany – very much accelerated since the turn of the millennium (see Frick & Grabka 2008).

Figure 6 confirms this finding using various inequality indicators.[13] More importantly for our argument, however, we observe again an increasing gap between the four treatments with Version 1 showing the highest degree in inequality (no matter which indicator is chosen), Version 2 yielding a somewhat lower degree of inequality, and finally Versions 3 and 4 producing – once again in similar fashion – the lowest level of dispersion.

Figure 6: The impact of PUNR-treatment on income inequality

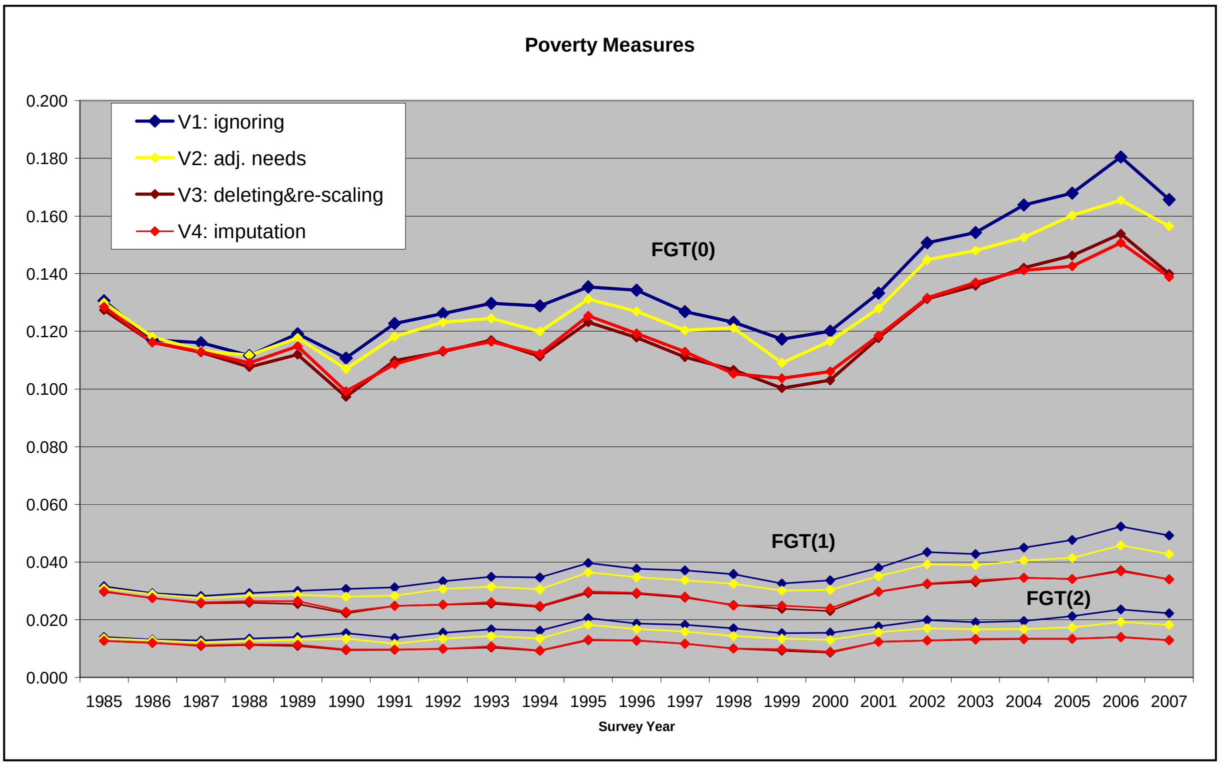


Finally, we present results on relative income poverty using the parametric family of FGT measures. Confirming the hypotheses laid out in Section 4.1.1 the results for the poverty head count ratio (FGT0) in Figure 7 are highest for the "ignorance" treatment (Version 1), while Version 2 takes on a middle position, and "reweighting" as well as "imputation" yield the lowest level of poverty risk. Again the deviation is growing with duration of the panel, thus also reflecting the increase in PUNR incidence. The difference in poverty risk rates in the most recent years is up to two percentage points!

13 Note that the very strong increase in the HSCV since 2003 is predominantly caused by the incorporation of SOEP's high income sub-sample G.

All those results are very stable using higher poverty aversion parameters in FGT1 ("normalized income gap ratio") and FGT2 (which gives more weight to the poorest poor).

Figure 7: The impact of PUNR-treatment on relative income poverty



4.1.3 Consistency of poverty "assignment" when using different approaches to tackle PUNR

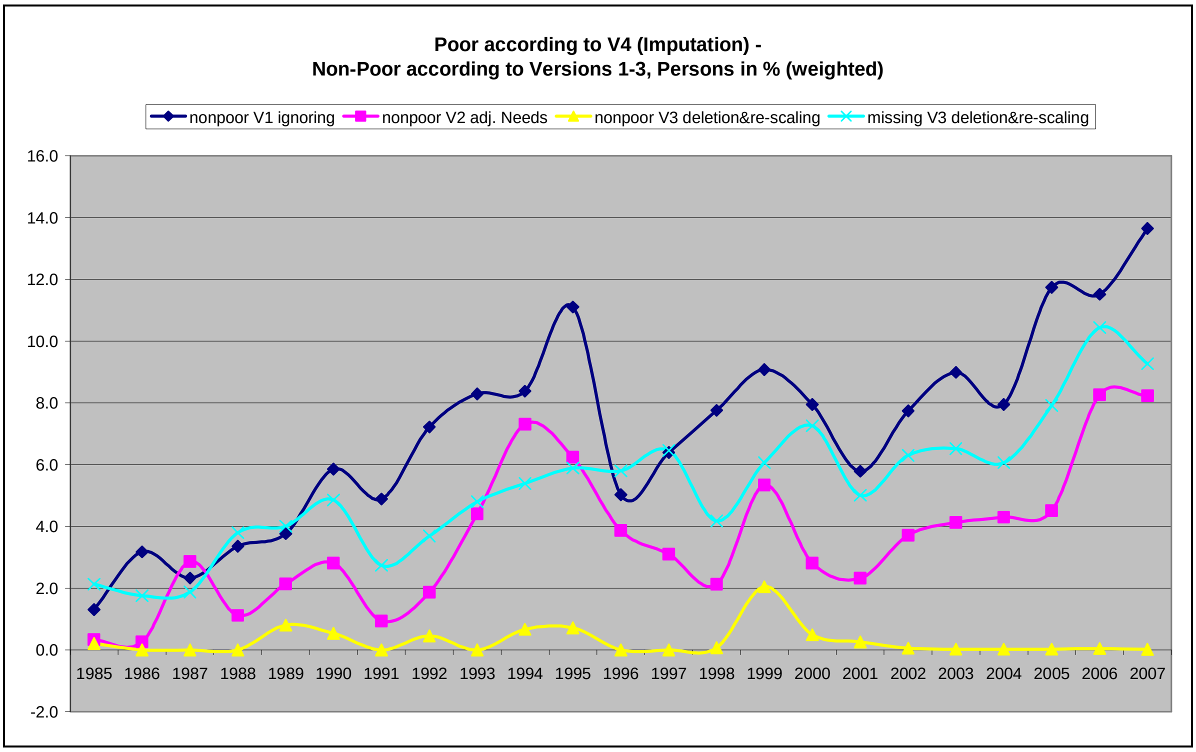
Obviously, there is clear variation in the overall level of relative poverty at the aggregate or national level across the various techniques. However, even if those results were more alike, it would not necessarily require that the various approaches identified the very same individuals as being poor. The results presented in Figure 8a and Figure 8b challenge the consistency of the poverty measurement on the basis of the four approaches, thus answering the question "who is poor according to approach x, but not poor according to approach y" and vice versa.

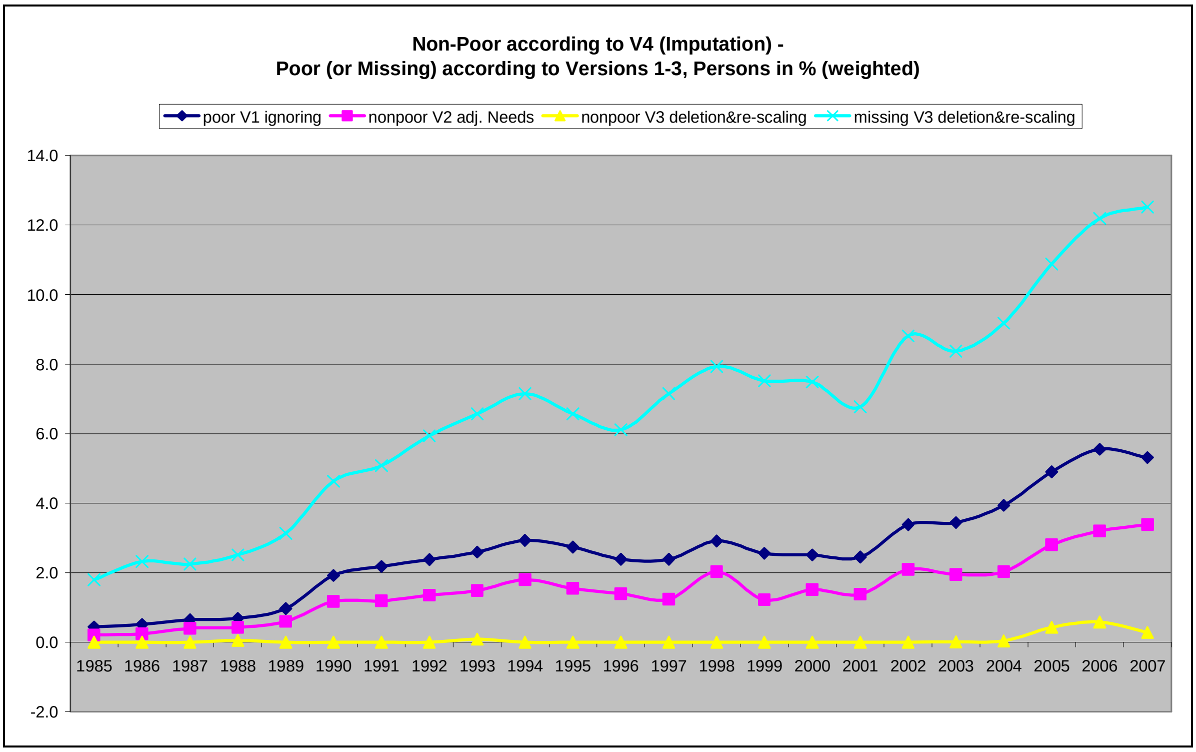
In other words, above and beyond the sheer interest in the overall share of people living below the poverty line, it is of utmost importance for the design, application and evaluation of social policy programs to better know about the socio-economic structure of the population being affected by relative income poverty. The effects of any such reform should certainly not just mirror the assumptions underlying the approach to deal with PUNR in the microdata at use. If that was the case, however, we may expect child poverty to look different if PUNR was mostly a problem among households with dependent children thus misreporting their income measure and making them appear poorer than they actually are.

In Figure 8a we restrict our sample to those who happened to live below the poverty line according to Version 4 ("Imputation"). The time series graphs show the share of those who have been identified as non-poor according to the three other approaches. In line with the analyses results so far, there is a high degree of concordance with the identification of poverty in Version 3 ("elimination and re-weighting" as given by the yellow graph), although the population used in Version 3 is more selective as can be seen in the share of those missing from the analysis (light blue graph) which accrues up to about 10% of the baseline population. Clearly less comparable are the results based on Version 1 (dark blue graph) and Version 2 (pink graph) which both show an increasing deviation from the results obtained from the imputed data.

In contrast, Figure 8b is based on the population of non-poor in Version 4 ("Imputation") and the various graphs show time series of the share of those persons being identified as poor in Versions 1 to 3. Again, we observe a high degree of similarity with the results of Version 3, while the share of those eliminated in Version 3 is as high as 12% in the most recent waves. The results obtained from Versions 1 and 2 are significantly less comparable, and differences again are growing over time.

Figure 8a,b: Consistency of poverty measurement using alternative approaches to handle PUNR





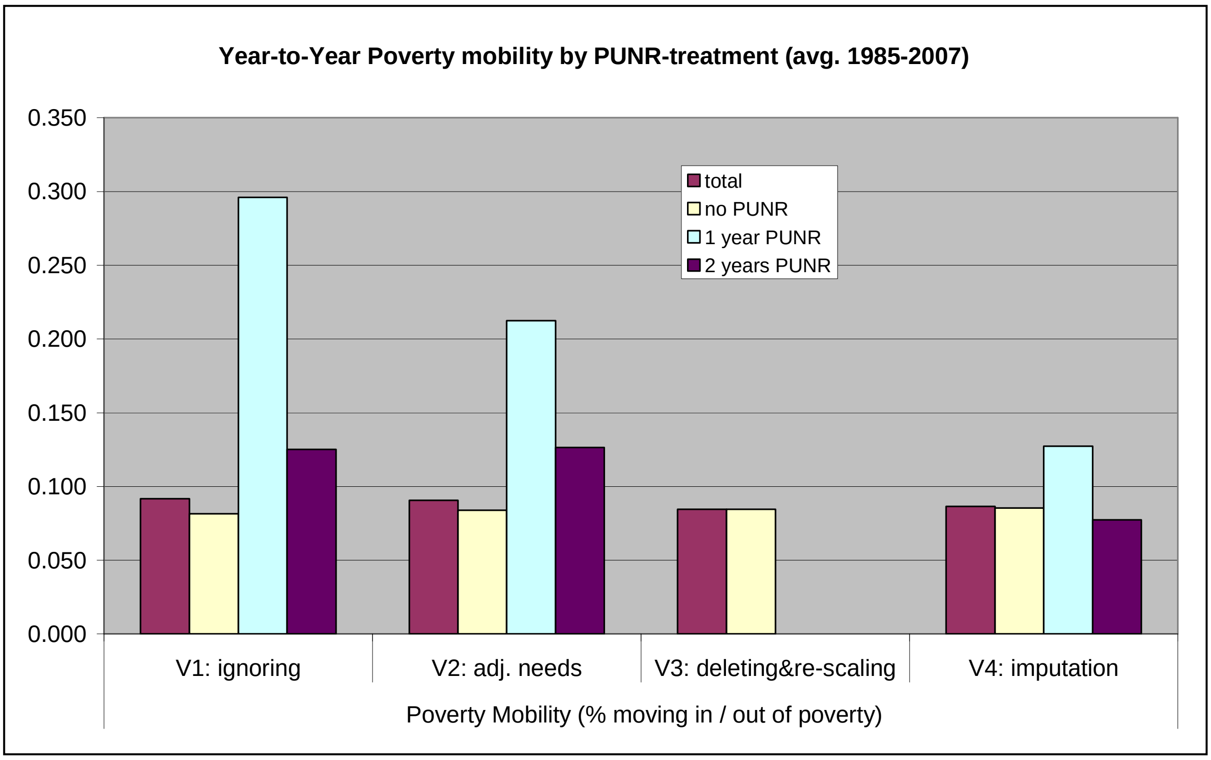
4.2 Poverty and Income mobility

The analyses so far indicated a significant impact of the methodological decision on how to cope with PUNR on cross-sectional results (inequality and poverty). In the following we address the question to what degree this is true for longitudinal analyses as well by comparing results derived from the four methods with respect to poverty and income mobility.

Figure 9 presents results from simple wave-to-wave poverty mobility analyses, for robustness purposes averaging all pooled two-year balanced panels over the period 1985-2007. For each of the four approaches, we show the share of individuals moving in or out of poverty within a given two-year interval. In order to better assess the impact of PUNR on poverty mobility we separate the population into three groups: those who were continuously living in households not affected by PUNR in both waves, those who experienced PUNR in only one wave, and finally the group of people being affected by PUNR in both years. Comparing the four techniques the selectivity of the population in Version 3 (elimination and re-scaling) becomes apparent. Indeed, one may argue that maintaining the entire survey population, i.e., including households with PUNR, may be more important for mobility analyses than for purely cross-sectional analyses, a strong argument in favor of imputation (Version 4) over weighting (Version 3).

Although the degree of poverty mobility in the aggregate does not differ much across the four versions, PUNR-households show clearly higher mobility rates, simply because PUNR in at least one wave increases the probability of being poor due to understated incomes as shown above in Section 4.1. Results based on imputed data (Version 4) still exhibit above average mobility, in particular if PUNR was present only in one wave. One the one hand, this reflects in part the uncertainty of the underlying imputation procedure (i.e., we inject variation by adding residuals to the predicted incomes), one the other hand, we cannot rule out that the missingness simply reflects “true” mobility, if the PUNR, for example, had been caused by a change in labor market status of the respondent which interfered with survey participation. If the latter was the case, then indeed the mobility results in version 3 (elimination and re-scaling) would be downward biased.

Figure 9: The impact of PUNR-treatment on poverty mobility



Using the very same data, Figure 10 gives very consistent results with respect to income mobility over two years when applying the Fields & Ok (1999) index. It should be noted that those results presented here are insensitive to the choice of the mobility measure – applying for example the Shorrocks (1978) index (using Gini coefficient) yields more or less identical results (available from the authors upon request).

Extending the observation period to five-year intervals, Figure 11 again shows the lowest degree of income mobility among PUNR households when using imputed data instead of applying Versions 1 or 2. Confirming the results from Figure 10, the elimination of PUNR-households from Version 3 consequentially yields the lowest degree of mobility for the entire population.

Figure 10: The impact of PUNR-treatment on 2-wave income mobility

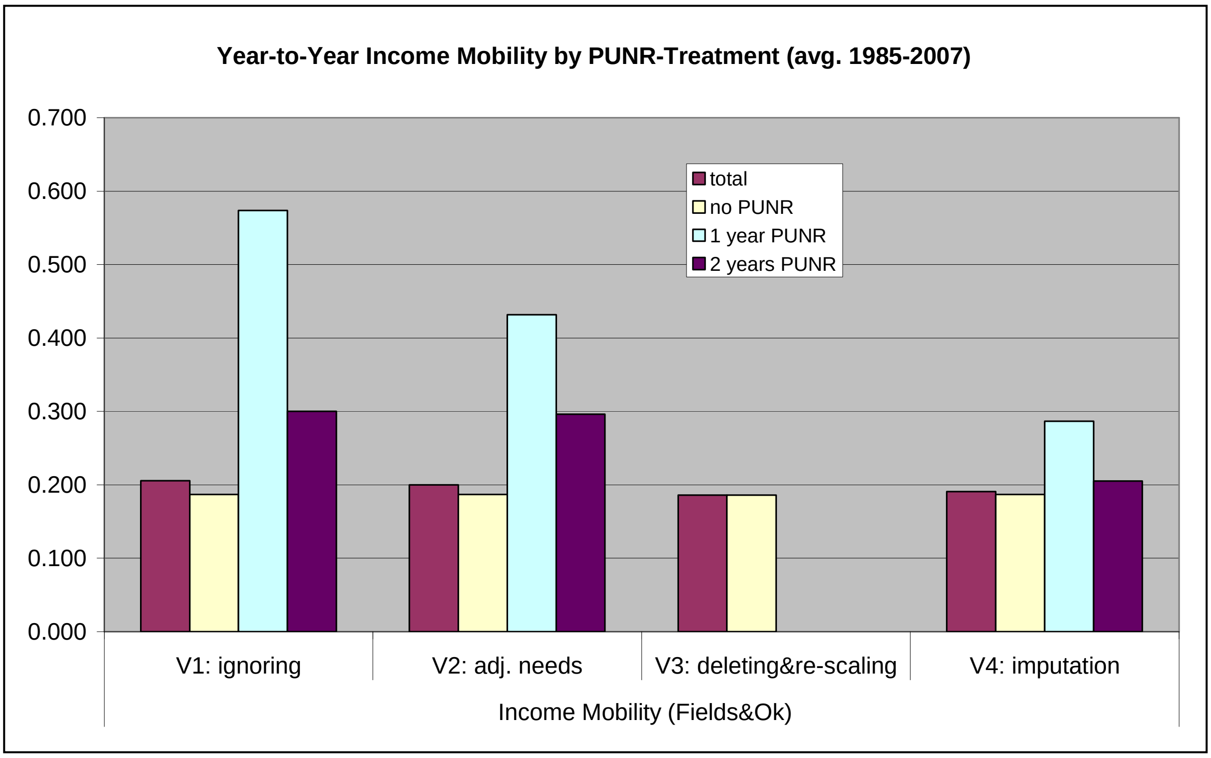
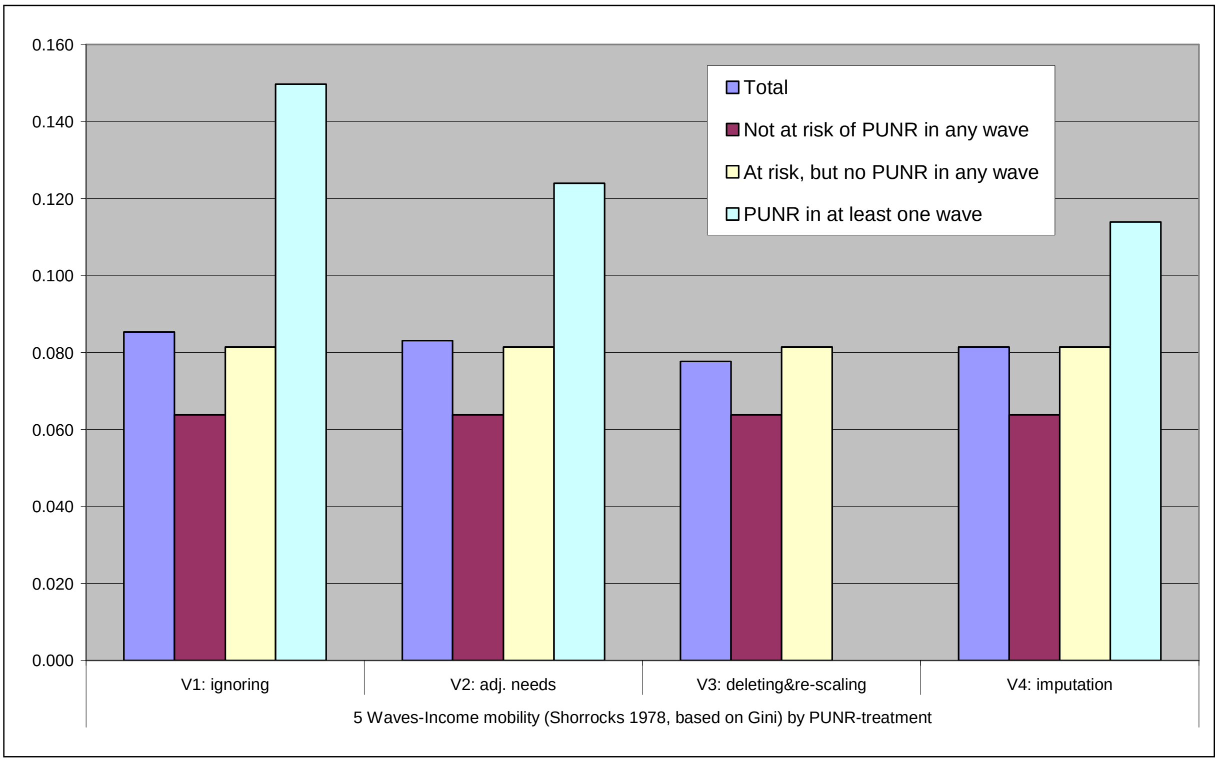


Figure 11: The impact of PUNR-treatment on 5-wave income mobility

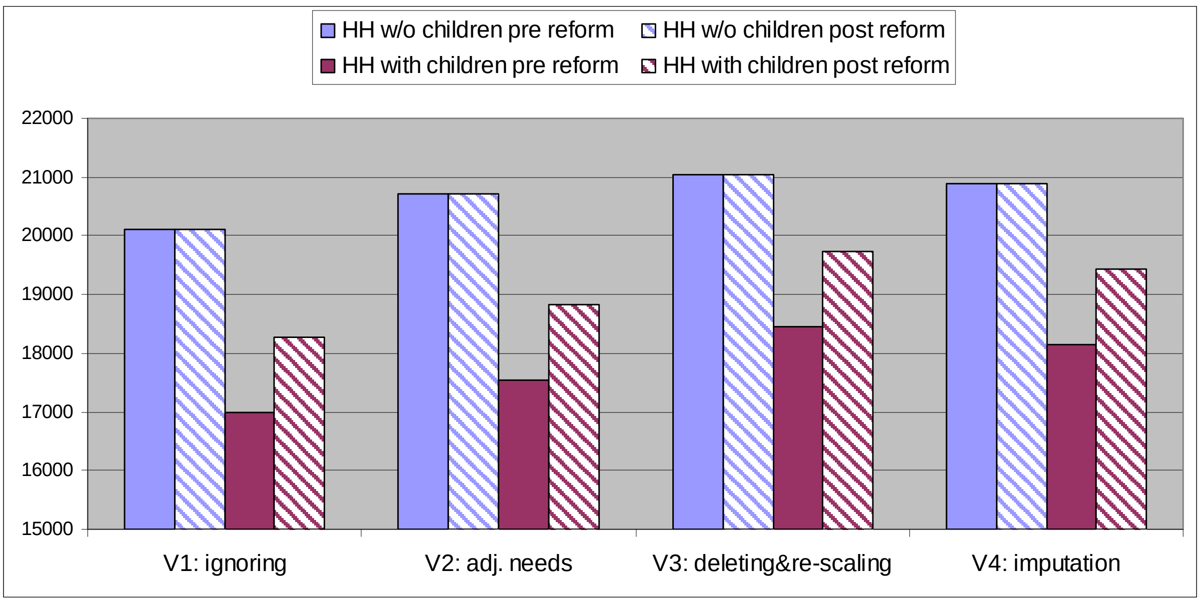


5 Application: Simulating a change in child benefits using different treatments of PUNR

In order to assess the impact the choice of the technique has on a fictitious policy change we simulate an increase in (non means-tested) child allowances and evaluate the resulting effect on poverty separately for each of the four approaches. The microsimulation is carried out only for the survey year 2007 and assigns 150 Euro additional child allowance per child to all households with children up to 16 years of age.[14] Following, we compare pre- and post-simulation results with respect to mean, inequality and poverty risk ratios separately for the population in households with and without eligible children (details are given in Appendix B). Adding the transfer yields an increase in average income by about 2.5%. Note that the poverty thresholds are dynamically adjusted to this increase.

While the simulation leaves the income of childless households unchanged, Figure 12 shows the expected increase in income by about 7 to 7.5% among households with children, depending on the PUNR-treatment. More important and irrespective of our simulation, there are considerable differences in income levels by household type depending on the technology used to counter PUNR. Apparently, any approach to actively deal with PUNR (i.e., versions 2 to 4) yields an improved income position of children when compared to Version 1 (“ignoring PUNR”).

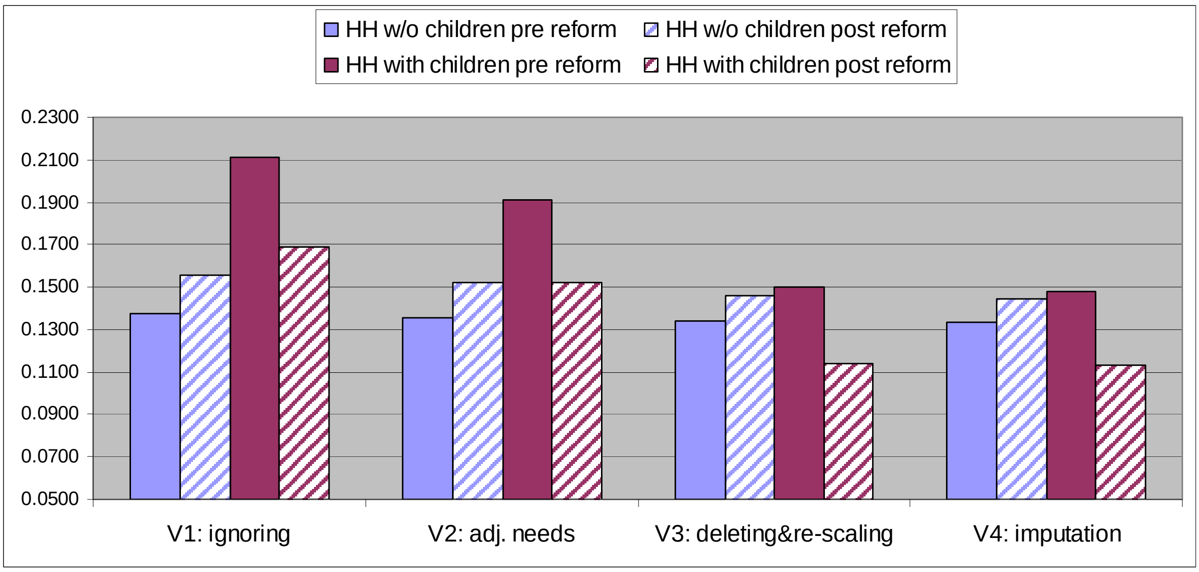
Figure 12: Simulated income effects by PUNR-treatment



Consistent with their below average income position, children in PUNR-households are also more exposed to poverty risks in Versions 1 and 2 (see Figure 13). Already prior to a simulated reform of child benefits, the gap in poverty risks for households with children versus households without children is much larger according to Version 1 (ignoring PUNR) [more than 7%-points] than in any other approach, especially in Version 4 [barely 1.5%-points]. This is clear indication for downward biased income measures among households with children when ignoring PUNR.

The poverty effects arising from our microsimulation (see bars with striped pattern in Figure 13) can be valued rather differently. In all four versions we find an increase in poverty rates among the childless population due to the dynamic upward adjustment of the poverty threshold (by 8% to 13% depending on the technique). At the same time we observe a significant reduction in child poverty rates (by approximately 20% to 24%), inducing however marked differences with respect to the poverty structure across household types. In version 1, even after increasing benefits, the population in households with children faces above average poverty risks. Adjusting needs in Version 2 already sees children's poverty risk on par with that of childless households. Finally, Versions 3 and 4 yield – once again – a rather similar picture. Here, child poverty rates are about 3%-points below those of childless households indicating a "more successful" poverty reduction policy induced by the additional benefits. .

Figure 13: Simulated poverty effects by PUNR-treatment



14 This transfer is added to household income prior to equivalizing.

6 Conclusions

Using 24 waves of panel data from the German SOEP we find an increasing incidence of PUNR together with clear indication for selectivity of PUNR. A major consequence of this phenomenon is a systematic downward bias in level and development of income inequality and relative poverty whereas income mobility will be overstated due to people moving in and out of PUNR. Our strategy of imputing single income components as well as including those into the estimation of taxes and social security contribution appears to be an appropriate means to cope with both of these problems, thus guaranteeing a less-biased measure of post-government income as empirical basis for analyses of income inequality and poverty.[15] The imputation of various components instead of only adjusting the "annual post-government income measure" (e.g. by means of a "flat correction factor") may also be considered advantageous because it supports decomposition analyses (by income source) and portfolio analyses while maintaining the entire survey population (in contrast to alternative strategies which exclude those affected by PUNR and re-weight those at risk of PUNR).

Future research will have to invest in the following questions:

- Is there a correlation between PUNR und subsequent attrition (UNR) which would be most relevant for mobility analyses?
- Does the choice of PUNR-treatment affect comparability in cross-country comparisons (see Frick & Grabka 2009 for the need to harmonize of imputation item-non-response)?
- While our analysis was concerning the missing contribution of individuals to their respective household's resources, PUNR may also yield similar bias in other research areas such as labor economics, where the interaction between household members is of crucial importance, for example when modelling labor supply decisions of couples. The missing mechanism for PUNR may not at all be at random, if the missing partner can not participate in the survey because s/he is making lots of money while drilling for oil on an off-shore platform, or because s/he is not willing to participate while being unemployed or because of being severely sick.

15 Future extensions may consider *multiple* imputation which would also take into account the uncertainty embedded in the single imputation applied here.

- Finally, what do our results imply for designing incentives in household panel surveys (see Laurie & Lynn, 2009)? Instead of correcting the microdata *after* data collection one may rather try to prevent missingness from happening. Thus, it might be preferential from a data collection point of view to consider some ideas for early interventions such as (a) collect proxy information on individuals with restricted interview capability e.g. due to severe sickness, dementia, Alzheimer disease etc.; (b) increase incentives to participate e.g. by monetary incentives. While this additional participation may exert positive spill-over effects on other household members, it may also yield some habituation effect in the panel perspective: interviewees who were paid this extra money once, want to keep that "bonus" in future waves as well, making this approach a rather expensive one. One may also consider an additional incentive at the household level being paid only if all adult respondents do participate (this is done in the HILDA survey, see Watson & Wooden 2009). (c) An alternative might come with a short drop-off questionnaire to be filled in by PUNR respondents, as to improve the basis for the imputation or weighting procedure. However, as is the case for proxy-interviews such an approach might also be used by respondents to sneak out from the regular survey in order to reduce response burden.

Above and beyond the arguments brought forward in this paper there may as well be other reasons for why PUNR is of increasing relevance over more recent years than just simply measurement issues. One argument arises from the increasing number of individuals who are affiliated with *multiple residences,* e.g. due to either long distance commuting between "home" and "work" as well as due to choosing "modern" life styles such *living apart together*, thus couples who do not share a common address but rather live in two separate places at the same time. Those developments obviously complicate a clear distinction of "private households" because it is not all that clear which household a given person belongs to or whether her resources as well as her needs should be assigned to just one or (partially) to several households. A dataset which has been "fully" imputed in case of PUNR may also well overstate the true welfare position if one does not carefully deduct the full costs of commuting.

7 References

Atkinson, A. B. and F. Bourguignon (eds) (2000): Handbook of Income Distribution, North-Holland.

Barceló, Cristina (2008): The impact of alternative imputation methods on the measurement of income and wealth: evidence from the Spanish survey of household finances, Documentos de Trabajo N.º 0829, Banco De Espana, Madrid.

Becker, I. (2000): Einkommensverteilung in Deutschland - Strukturanalyse der Ungleichheit nach Einkommenskomponenten. Jahrbücher für Nationalökonomie und Statistik, (220/4), p. 400-418.

Bonke, J. and H. Uldall-Poulsen (2007) Why Do Families Actually Pool Their Income? Evidence from Denmark, Review of Economics of the Household, 5(2): 113-128

Buhmann B, Rainwater L, Schmaus G, and Smeeding T. M. (1988) Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database. The Review of Income and Wealth. 34(2):115–142.

Canberra Group (Expert Group on Household Income Statistics) (2001): Final Report and Recommendations. Ottawa, http://www.lisproject.org/links/canberra/finalreport.pdf.

De Leeuw ED, De Heer W (2002). Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. In RM Groves, DA Dillman, JL Eltinge, RJA Little (eds.), Survey Nonresponse (pp. 41–54). New York: Wiley.

De Leeuw ED, Hox JJ, Huisman M (2003). Prevention and treatment of item nonresponse. Journal of Official Statistics, 19, 153-176.

De Leeuw, ED, Hox, JJ, DA Dillman (eds)(2008): International Handbook of Survey Methodology. New York, London: Erlbaum, Taylor & Francis.

Eurostat (2000): Imputation of Income in the ECHP. DocPan 164/00. Luxembourg.

Fields, G.S. and E. Ok (1999): Measuring Movement of Incomes, Economica, 66: 455-471.

Foster, J., J. Greer, and E. Thorbecke (1984): A Class of Decomposable Poverty Measures. Econometrica, 52 (3), 761-766.

Frick JR, Grabka MM (2005). Item non-response on income questions in panel surveys: incidence, imputation and the impact on inequality and mobility. Allgemeines Statistisches Archiv, 89, 49-61.

Frick JR, Grabka MM (2009): “Item Non-Response and Imputation of Annual Labor Income in Panel Surveys from a Cross-National Perspective”, In: Lyberg, L. et al (eds): xxx, Wiley & Sons (forthcoming).

Frick, J. R. and Grabka, M. M. (2003): Imputed Rent and Income Inequality: A Decomposition Analysis for the U.K., West Germany, and the USA. The Review of Income and Wealth, 49(4): 513-537.

Grabka, M. M. and J.R. Frick (2008): The Shrinking German Middle Class—Signs of Long-Term Polarization in Disposable Income?. DIW Berlin Weekly Report No. 4/2008, 21-27.

Groves RM (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. Public Opinion Quarterly, 70(5) Special Issue, 646–675.

Groves RM, Couper MP (1998). Non-response in Household Interview Surveys. New York: Wiley.

Groves RM, Dillman DA, Eltinge JL, Little RJA (eds.) (2002). Survey Nonresponse. New York: Wiley.

Heckman J (1979). Sample Selection Bias as a Specification Error. Econometrica, 47, 153-161.

Hill D, Willis RJ (2001). Reducing Panel Attrition: A Search for Effective Policy Instruments. Journal of Human Resources, 36, 416-438.

Kapteyn, A., P.-C. Michaud, J. Smith and A. van Soest (2006): Effects of Attrition and Non-Response in the Health and Retirement Study, IZA Discussion Paper No. 2246, August 2006, IZA Bonn.

Kroh, M. and M. Spieß (2007): Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984-2007). Data Documentation 39, Berlin: DIW Berlin.

Laurie H., Lynn P (2009). The use of respondent incentives on longitudinal surveys. In P Lynn (ed.), Methodology of Longitudinal Surveys. London: John Wiley and Sons, 205-234.

Little RJA, Su HL (1989). Item Non-Response in Panel Surveys. In D Kasprzyk, G Duncan, MP Singh (eds.), Panel Surveys (pp. 400-425). New York: John Wiley.

Loosveldt G, Pickery J, Billiet J (2002): Item non-response as a predictor of unit non-response in a panel survey. Journal of Official Statistics, 18, 545-557.

Lynn P (2008). The Problem of Nonresponse. In: ED De Leeuw, JJ Hox, DA Dillman (eds): International Handbook of Survey Methodology (pp. 35-55). New York, London: Erlbaum, Taylor & Francis.

Lynn P (ed.) (2009): Methodology of Longitudinal Surveys. London: John Wiley and Sons.

Nicoletti C, Peracchi F (2006). The effects of income imputation on microanalyses: evidence for the European Community Household Panel. Journal of the Royal Statistical Society Series A, 169 (3), 625-646.

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

Rubin DB (1987). Multiple imputation for nonresponse in surveys. New York: Wiley.

Shorrocks, A. F. (1978): Income Inequality and Income Mobility, Journal of Economic Theory, 19: 376-393.

Smeeding, T. M. and D. H. Weinberg, (2001): “Toward a Uniform Definition of Household Income,” The Review of Income and Wealth, 47(1), 1-24.

Spiess M, Goebel J (2003). Evaluation of the ECHP imputation rules. CHINTEX Working Paper No. 17.

Starick R, Watson N (2007): Evaluation of Alternative Income Imputation Methods for the HILDA Survey. HILDA Project Technical Paper Series No. 1/07, Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

Vogler, C. and J. Pahl (1994) Money, Marriage and Inequality within Marriage. Sociological Review, 42(2), 236-288

Wagner, G.G., Frick, J.R., Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP) – Evolution, Scope and Enhancements. Schmoller’s Jahrbuch – Journal of Applied Social Science Studies 127(1): 139ff.

Watson N, Wooden M (2009). Identifying factors affecting longitudinal survey response. In P Lynn (ed.), Methodology of Longitudinal Surveys. London: John Wiley and Sons, 157-182.

Appendix A: Excerpt from the ECHP documentation on the PUNR-adjustment procedure

Imputation of income in the ECHP — November 2000

6. WITHIN HOUSEHOLD NON-RESPONSE INFLATION FACTOR (HI010)

In order to have comparable figures at aggregated household level, it is necessary to make up for unit non-response if no questionnaire was answered by some persons in the household.

Remark: HI010 can only be calculated after the weighting procedure.

To do so, all persons in a country are grouped into 110 groups G - 11 age classes, sex and quintiles of equivalised net monthly household income (HI200 / HD005). For each group the weighted average (PINMEAN) of the provisional total personal income (PI100pr) is calculated.

PI100pr =
PI111 + PI112 + PI12 + PI121 + PI123 + PI131 + PI1321 + PI1322 + PI133 + PI134 + PI135 + PI136 +
(HI122 + HI137 + HI138) / HD002A

$$PINMEAN_G = \frac{\sum_{all\ persons\ in\ group\ G} PI100pr * PG002}{\sum_{all\ persons\ in\ group\ G} PG002}$$

For households in which not all the eligible persons answered, HI010 is calculated as

$$HI010 = \frac{\sum_{all\ persons\ with\ PENRES=11,13\ in\ HHD} PINMEAN_G}{\sum_{all\ persons\ with\ PELIG=1\ in\ HHD} PINMEAN_G}$$

For those household where the factor is 5 or more, it will be set to missing in order to avoid that the estimated total household income is based on less than 20% of real data:

If HI010 > 5 then HI010 = -9.

Appendix B: Results from a simulated increase in child benefits, 2007

		Pre reform	Post reform	diff	dev.%
			Means		
Total	V1: ignoring	18921	19409	488	2.6
	V2: adj. needs	19511	19999	488	2.5
	V3: deleting&re-scaling	20073	20552	479	2.4
	V4: imputation	19840	20328	488	2.5
HH w/o children	V1: ignoring	20103	20103	0	0.0
	V2: adj. needs	20712	20712	0	0.0
	V3: deleting&re-scaling	21036	21036	0	0.0
	V4: imputation	20882	20882	0	0.0
HH with children	V1: ignoring	16989	18275	1285	7.6
	V2: adj. needs	17547	18832	1285	7.3
	V3: deleting&re-scaling	18440	19731	1291	7.0
	V4: imputation	18136	19422	1285	7.1
			FGT(0)		
Total	V1: ignoring	0.166	0.161	-0.005	-3.0
	V2: adj. needs	0.156	0.152	-0.004	-2.8
	V3: deleting&re-scaling	0.140	0.134	-0.006	-4.1
	V4: imputation	0.139	0.133	-0.006	-4.4
HH w/o children	V1: ignoring	0.138	0.156	0.018	13.0
	V2: adj. needs	0.135	0.152	0.017	12.4
	V3: deleting&re-scaling	0.134	0.146	0.012	9.0
	V4: imputation	0.133	0.145	0.011	8.4
HH with children	V1: ignoring	0.211	0.169	-0.042	-20.0
	V2: adj. needs	0.191	0.152	-0.039	-20.3
	V3: deleting&re-scaling	0.150	0.114	-0.036	-23.9
	V4: imputation	0.148	0.114	-0.034	-23.2
			Gini		
Total	V1: ignoring	0.313	0.303	-0.010	-3.2
	V2: adj. needs	0.305	0.296	-0.010	-3.2
	V3: deleting&re-scaling	0.296	0.288	-0.008	-2.8
	V4: imputation	0.293	0.285	-0.009	-3.0