Article

Dealing With Incomplete Household Panel Data in Inequality Research

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

Abstract

In trying to capture complete within-household heterogeneity, household panel surveys typically try to interview *all* adult household members. Following from this, such surveys tend to suffer from partial unit nonresponse (PUNR), that is, the nonresponse of at least one member of an otherwise participating household, most likely yielding an underestimation of aggregate household income. Using data from the German Socio-Economic Panel (SOEP), the authors evaluate four different strategies to deal with this phenomenon: (a) *ignorance*, that is, 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 to suffer PUNR and reweighting of households observed to be at risk of but not affected by PUNR; and (d) *longitudinal imputation* of the missing income components. The aim of this article is to show how the choice of technique affects substantive results in inequality research. The authors find indications of substantial bias on income inequality and poverty as well as on income mobility.

[1]DIW Berlin, Berlin, Germany and Technische Universitaet Berlin
[2]University of Bremen, Bremen, Germany

Corresponding Author:
Markus M. Grabka, DIW Berlin, Mohrenstrasse 58, 10117 Berlin, Germany
Email: mgrabka@diw.de

Motivation

One of the standard assumptions of welfare economics is that individuals living together in a ''needs unit''—usually a private household—pool and share all their available resources. This approach, which is generally accepted in the inequality research, is not just a means of adapting to data limitations: It is based on the idea that household members seek to achieve a common and equal standard of living (e.g., Canberra Group 2001; Atkinson and Bourguignon 2000; Smeeding and Weinberg 2001).[1] This requires that all incomes received in a given household are aggregated across all members and that the total sum is distributed among all of them. Typically, an equivalence scale is then applied to adjust for differences in household composition and size, thus allowing for economies of scale in larger households as well as variation in 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 U.S. Panel Study of Income Dynamics, 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. Various sorts of nonresponse behavior appear in most of these surveys, however, posing a serious threat to the implicit assumption of (representative) full coverage of all the resources and needs of individuals living together in the household.[2] In such cases, unit nonresponse (UNR, i.e., the nonresponse of an entire household) is addressed by means of proportionally weighting successfully surveyed observations, while item nonresponse (INR) is corrected for by either weighting or imputation. But there remains the problem of partial unit nonresponse (PUNR) in those households where at least one member does not participate while other members do. In fact, there appears to be a generally increasing percentage of households affected by PUNR in population surveys. Ignorance of this phenomenon may give rise to several problems: (a) misreporting income aggregates; (b) increasing bias in results on income inequality, poverty, and mobility, and (c) bias in analyses on the intrahousehold distribution.

In the literature on social welfare, various approaches have been proposed for dealing with this phenomenon in the generation of ''equivalent household income,'' the outcome variable relevant for inequality and poverty analysis: First is ignoring the fact that a household member (and his or her income information) is missing, thus assuming the nonresponding individual's income is zero. This does not rule out that the person is living entirely from household-based transfers (public and private), which can be assumed to be measured correctly and comprehensively based on the information provided by other household members. Second is adjusting the calculation of the equivalence scale to ignore the person's share in household needs to compensate for his or her missing income. This approach implicitly assumes that the incomes of other household members are independent of the income of the missing individual. Third is eliminating all households observed with PUNR from the analysis population, thus assuming that there are no systematic differences between them and completely surveyed households, that is, that the underlying missing process is completely at random. An extension of this approach would try to compensate for potential selectivity by means of weighting. This can be accomplished by proportionally increasing the share of those at risk of but not affected by PUNR. That is, for all households that are not at risk of PUNR because there is only one adult respondent, the weighting factor remains unchanged. Obviously, this weighting strategy can be more or less complex depending on the case at hand. Last is imputing the missing income components at the individual level and aggregating across all household members (including those with PUNR).

At first glance, the assumptions of the first three options are very strong given the (likely) selectivity issues involved in the missing mechanism. The third (with reweighting) and final option appear to be less selective, and, in principle, the last option has the additional advantage of maintaining the entire survey population. Having said that, there is quite some normativity involved in the actual implementation of any such imputation process, as is true in any case of imputation. For example, one may argue that the degree of misspecification is actually 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 for household and individual characteristics related to the missing mechanism. For example, the severity of misreporting is probably very different for a household where data on the 85-year-old mother of the household head is missing because of lingering illness than for a household where the main

income earner is absent because he or she is drilling for oil on an offshore platform.

The results from research on equivalent household income inequality and especially on relative poverty may crucially depend on the choice of the aforementioned option because this decision will affect the income of all individuals living in households affected by PUNR and the relative poverty line to be derived from the national mean or median income. Having said that, any bias in cross-sectional income measures will also affect mobility analyses based on those measures.

Using more than 20 waves of micro data from the German Socio-Economic Panel (SOEP), this article assesses how approaches to PUNR can potentially affect income inequality, income poverty, and mobility. The article is set up as follows: In the second section, we first describe the incidence of and trends in 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) and for individual characteristics of the missing person. The third section presents the various techniques for dealing with PUNR, focusing on the principles of our three-stage imputation strategy for income components at the individual level. In the fourth section, we provide sensitivity analyses showing the variation in the results for income inequality and poverty when choosing among the aforementioned options. Making explicit use of the panel nature of the underlying data, we also demonstrate the importance of PUNR (and how it is treated in the micro data) for poverty dynamics and income mobility. Here, our findings show that dynamics are exaggerated if PUNR is present in at least one wave. The fifth section concludes with some remarks on the potential relevance of our findings for cross-national comparability of research on income inequality, poverty, and mobility.

Incidence and Selectivity of PUNR in the SOEP

The Data

The SOEP is a representative longitudinal survey of individuals living in private households in Germany (Wagner, Frick, and Schupp 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 more than 12,000 respondents, with everyone aged 17 and older in sample households being interviewed. In recent years, several representative new subsamples have been drawn, which have

approximately doubled the initial sample size. Other additional samples were explicitly designed to sample specific subgroups of the population. In 1995, the SOEP introduced an oversampling of immigrants to cope with the misrepresentation of recent immigrants in ongoing panel surveys. In 2002, to overcome the problem of a lack of information on ''rich'' individuals in representative population surveys—an important group for welfare analyses given the high concentration of economic resources (income and wealth) at the top of the distribution—the SOEP introduced a high-income sample overrepresenting the top 3 percent of the income distribution. The sample analyzed below employs all available observation years up to survey year 2007.

One of the main problems population surveys face when asking for (specific) income and wealth information is nonresponse, and SOEP is no exception. To make effective use of the panel nature of SOEP, all cases of INR are corrected for using 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 potential biases arising from the aforementioned selectivity can be reduced (see Frick and Grabka 2005).[3]

Incidence and Selectivity of PUNR

Incidence of PUNR at the person versus household level. There are at least two ways to express the overall incidence of PUNR in a household panel survey. For example, while ''only'' 5.4 percent of all adult household members did not fill in the requested questionnaire themselves in 2005, their nonparticipation behavior affected the measurement of relevant outcomes for all other members of their respective households as well. Thus, 11.5 percent of all individuals (including children who not yet reached the respondent age of 17) lived in a household that was affected by PUNR.

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, there has been a clear tendency toward a growing percentage of nonparticipating respondents since then.[4] This process is even more striking given the secular trend toward smaller households: The population at risk of PUNR is actually shrinking because of the increasing share of singles and lone parents in the population (with minor children up to 17 years of age, the respondent age in SOEP), that is, household types for which nonresponse of the only respondent living in that household by definition yields a complete dropout (= UNR). Of course, the
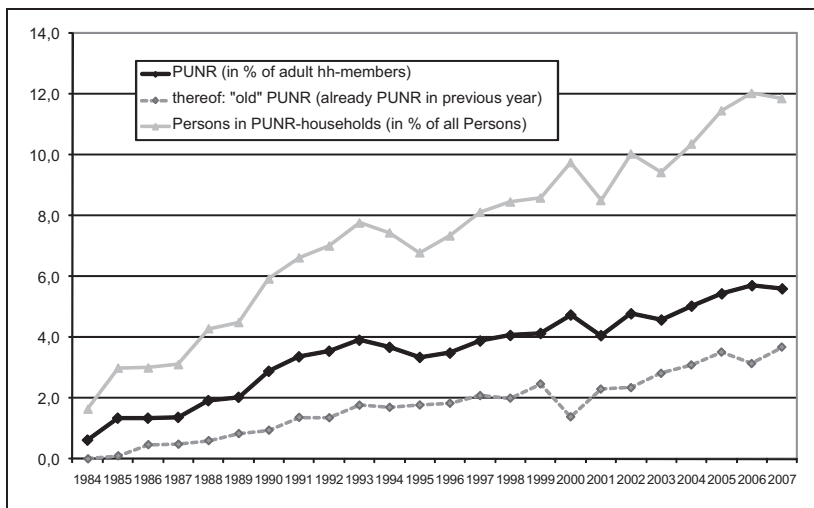
Figure 1. Incidence of PUNR in the German SOEP, 1984-2007
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

increase in the incidence of PUNR is also driven by the accumulation of ''old'' PUNRs over time, that is, persons who basically never give personal interviews while other household members continue to do so.

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 only two adults. While children younger than respondent age do not bear the risk of PUNR themselves, they may be affected by PUNR of adult household members: Any misrepresentation of the adults' resources in the household aggregate will affect measures of child poverty.

Figure 2 clearly illustrates this effect by indicating consistently increasing shares of individuals affected by PUNR, either directly (because of their own nonparticipation) or indirectly (because of nonresponding coresidents): In 2005, for example, the share of individuals affected by PUNR within the household context is around 10 percent in households with two persons of
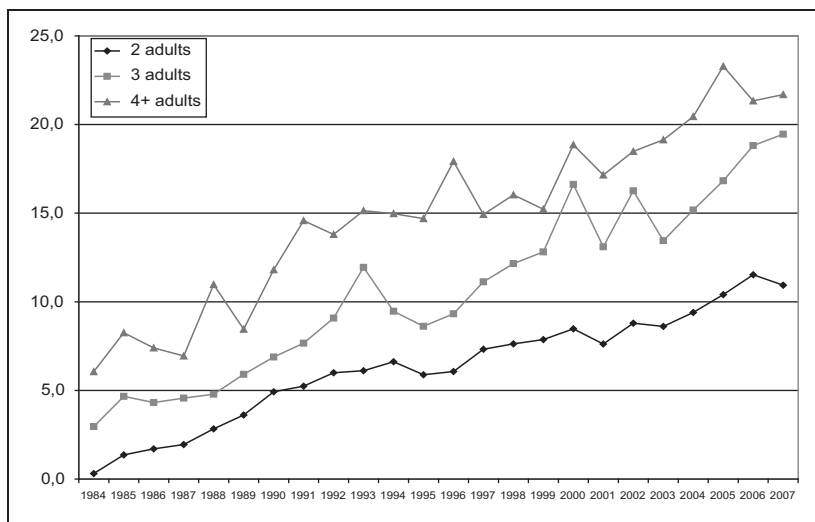
Figure 2. Incidence of PUNR by number of adult household members
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

respondent age, about 17 percent in households with three respondents, and greater than 20 percent for the rather few observations in large households of four and more adults.

Incidence of PUNR by panel experience. For a long-running household panel such as SOEP, any nonresponse behavior is crucial for maintaining the quality and representativeness of the *longitudinal* data. The research on the scope and selectivity of UNR in household panel studies provides rather robust evidence that the probability of dropping out decreases with panel experience; thus, any additional interviews reduce the probability of UNR (see, e.g., Watson and Wooden 2009 on the Australian Household, Income and Labour Dynamics in Australia [HILDA] Survey). Although there is clear empirical support for this hypothesis on UNR in the SOEP data (on the weighting scheme in SOEP, see Kroh and Spiess 2007), the probability of PUNR does not necessarily monotonically decrease for long-term respondents (see Figure 3). We find a clear reduction in this probability only over the first few years of panel experience (including the years of childhood that a person spent ''growing up'' in the survey without being a respondent himself or herself) and again after approximately 20 years.
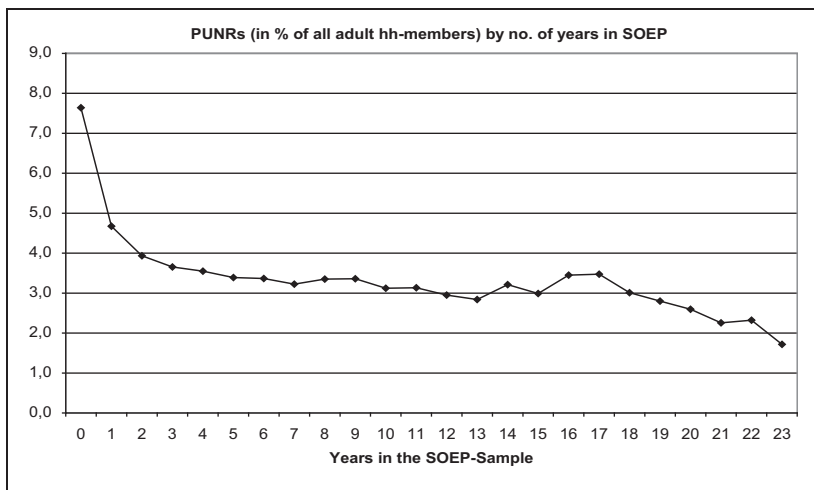
Figure 3. Incidence of PUNR by panel duration
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

It may be that a respondent in an otherwise cooperative household is simply more likely to drop out temporarily in a period of 20 years than in just 5 years. As such, a temporary dropout of a long-term respondent may be less problematic than that of a short-term respondent, that is, it may reduce the number of observations in a balanced panel sample, thus reducing efficiency. However, an alternative hypothesis that is not tested in this article could be that the established relationship between interviewer and respondent for long-term panel members makes these individuals more likely not to participate in times when ''unusual'' events occur, that is, occurrences that they are embarrassed to tell the interviewer about (e.g., a successful manager's job loss). This second hypothesis is more in line with findings by Kapteyn et al. (2006), who argue that attritors in the Health and Retirement Survey who are recruited back into the survey are very different from permanent attritors.

Selectivity of PUNR. To control for potential selectivity of PUNR, we make use of multivariate analyses. Table 1 shows results from a pooled probit regression model based on more than 45,000 individuals observed in the SOEP during 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.

Table 1. Probability of PUNR—Results From a Pooled Probit Regression

		Coeff.		Std. err.
Sex	Female	−0.443	***	0.022
Age groups (ref: 56–65 years of age)	17–24	−0.003		0.024
	25–40	−0.053	*	0.024
	41–55	−0.210	***	0.033
	66 and older	−0.304	***	0.036
Relation to household head (ref: household head)	Partner	0.977	***	0.025
	Child	0.984	***	0.036
	Other	1.700	***	0.049
Migration background (ref: autochthonous)	Native-born foreigners	0.204	***	0.045
	Foreign-born Germans	−0.264	***	0.045
	Foreign-born foreigners	−0.347	***	0.037
Region (ref: West Germany)	East Germany	−0.245	***	0.025
Change in household composition (ref: no change)	Change	0.076	***	0.016
	First household interview	0.528	***	0.015
No. of adults in household (ref: 2)	3 adults	0.052	*	0.023
	4 + adults	0.125	***	0.028
No. of children in household (ref: none)	1 child	−0.078	***	0.023
	2 children	−0.163	***	0.026
	3 + children	−0.170	***	0.032
Home owner (ref: tenant)	Home owner	−0.051	**	0.019
Highest education of hh head/partner (ref: intermediate)	Lower secondary	0.039		0.034
	Higher secondary	−0.102	***	0.023
	Tertiary	−0.124	***	0.026
Item nonresponse monthly household income (ref: no)	Yes	0.396	***	0.025
Control dummies for survey year	Yes			
Constant		−1.779		0.034
Pseudo-R2		.1568		
Observations		325,414		
Individuals		45,038		

Source: German Socio-Economic Panel (SOEP, 1985–2007)v24, authors' calculations.
Robust standard errors.
* Significant at 10 percent. **Significant at 5 percent. ***Significant at 1 percent.

Women tend to have a lower probability of PUNR, which is also true for the middle aged (25–40) and elderly (66 years and older). Similarly, home ownership, an increasing number of dependent children, and increasing levels of education are negatively related to PUNR. Compared to the head of household, we find spouses or partners, children, and any

other household members to be more likely to show PUNR. The probability of PUNR is particularly high at the time of a given household's first interview and increases when changes in household composition occur. As expected, the risk of individual nonparticipation increases with the number of adults to be interviewed. We also find significant evidence that INR on the question dealing with current monthly household income is related to PUNR of at least one household member. This is also in line with findings based on data from various panel surveys showing a positive impact of INR on income questions in wave t on the probability of UNR (i.e., attrition) in wave $t+1$ (see, e.g., Loosveldt, Pickery, and Billiet 2002; Frick and Grabka 2010). These findings may be taken as indications that economically active household members are more common among PUNR, and thus probably major contributions to overall household resources are understated.

Summing up, there is mixed evidence on the mechanisms driving PUNR. On one hand, long-term participation in panel surveys increases the risk of (temporary) PUNR, which could be completely random (CMAR) or at least random with respect to the income information (MAR). On the other hand, it is also likely that temporary nonresponse is caused by life events that affect individual income, for example, unemployment or sickness (NMAR). Thus, any approach of dealing with PUNR in income-based analyses should be capable of addressing the whole range of nonresponse mechanisms.

Dealing With PUNR in Income-Based Analyses of Economic Well-Being

Keeping the aforementioned selection issues in mind, the following section starts by briefly introducing the various approaches to dealing with PUNR before demonstrating prototypical empirical applications from a welfare economics perspective using an aggregated measure of equivalent household income.

Alternative Approaches

There exist a variety of ways to deal with PUNR in empirical analyses, four of which are applied in this article: (a) *ignorance*, that is, 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 using subsequent reweighting procedures; and (d) *longitudinal imputation* of the missing income components.

- *Ignorance:* Assume the individual affected by PUNR has no income of his or her own to add to the household's overall resources, but he or she does have needs that ought to be considered when constructing the household's equivalence scale. This means effectively ignoring PUNR in the measure of household *income*, while continuing to consider his or her *needs*, that is, $Y(PUNR) = 0$ & $Needs(PUNR) > 0$.
- *Adjustment:* Assume the individual has no income of his or her own as well as no needs and thus completely ignore the existence of the individual with PUNR. This effectively means deleting nonresponding individuals from PUNR households by adjusting the respective equivalent scale downward, that is, $Y(PUNR) = 0$ & $Needs(PUNR) = 0$, 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 rescale the population weights for those households that bear a risk of PUNR but did fully complete the survey. This assumes that the income and needs of households with PUNR are mirrored by successfully completed households with two and more respondents.[5]
- *Imputation:* Impute any income measure missing because of PUNR, thus considering all households with completed information on 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, there is another approach that is not considered in the remainder of this article: the flat ''correction factor,'' which has been applied in the European Community Household Panel (ECHP; see Eurostat 2000).[6] In this case, each household is assigned a specific ''within-household nonresponse inflation factor''. 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 were considered a pseudo-correction of the misreported income *level*, the income *portfolio* of the household would most likely remain subject to bias.

The incidence of PUNR and the treatment of this measurement problem in major household panel surveys are presented in Table 2.

Table 2. Incidence and Treatment of Partial Unit Nonresponse (PUNR) in Major Household Panel Surveys

	Share of PUNR (%)[a]	Method used
British Household Survey (BHPS)	10%-20%[b]	Ignorance[c]
Canadian Survey of Labour and Income Dynamics (SLID) 2008	7.2%[e]	Imputation
European Community household Panel (ECHP)		
- Greece	0.5%	Correction factor
- . . .	. . .	Correction factor
- Netherlands	10.3%	Correction factor
German Socio-economic Panel (SOEP) 2007	5.6%	Imputation
Household, Income and Labour Dynamics in Australia Survey (HILDA) 2007	5.9%	Imputation
Russia Longitudinal Monitoring Study (RLMS-HSE) 2009	4.0%	Ignorance
Swiss Household Panel (SHP) 2009	6.5%	Imputation
UK Household Longitudinal Study (UKHLS) 2009[f]		
- General Population Sample (first 12 months of first wave only)	18%	Imputation[d]
- Ethnic minority boost sample (first 12 months of first wave only)	29%	Imputation[d]
European Statistics on Income and Living Conditions (EU-SILC) 2007		
- Belgium	0.7%	Imputation
- Estonia	0.8%	Imputation
- France	0.6%	Imputation
- Cyprus	0.2%	Imputation
- Lithuania	0,3%	Imputation
- Austria	0.4%	Imputation
- Germany	0.4%	Correction factor
- Greece	0.6%	Correction factor
- Spain	0.7%	Correction factor
- Latvia	1.8%	Correction factor
- Portugal	0.4%	Correction factor
- Slovakia	1.5%	Correction factor

(continued)

Table 2. (continued)

	Share of PUNR (%)[a]	Method used
- Poland	6.0%	Ignorance
- Czech Republic	–	Elimination
- Ireland	–	Elimination
- Italy	–	Elimination
- Luxembourg	–	Elimination
- Hungary	–	Elimination
- United Kingdom	–	Elimination
- Denmark	–	Elimination
- Netherlands	–	Elimination
- Slovenia	–	Elimination
- Finland	–	Elimination
- Sweden	–	Elimination
- Iceland	–	Elimination
- Norway	–	Elimination
For information only: Survey of Health and Retirement in Europe (SHARE)		
- Denmark	7.0%	Imputation
. . .	. . .	Imputation
- Spain	26.2%	Imputation

Source: Authors illustration based on Bastien et al. (2010), European Communities (2003), Hayes and Watson (2009), Jenkins (2010) and Eurostat (2010: 31).
a. As a share of all eligible adult respondents.
b. Measured at the household level
c. PUNR is not considered in the construction of net income variables (Jenkins 2010).
d. Only aggregated household level information will be imputed.
e. Share of respondents of labour questions only.
f. The relative high share of PUNR is the result of a strict regulated fieldwork period.
– = not available.

The Imputation Strategy

The imputation procedure correcting for income missing because of PUNR is based on the following principles: First, we impute the most detailed income information possible (i.e., different components of income) to support augmented income analyses. Second, we employ longitudinal information, if available, to control for otherwise unobserved heterogeneity and—most important from the panel perspective—to support longitudinal analysis of income and poverty dynamics. Third, we make use of household context data and in particular household-level income data to control for potential nonrandom mechanisms of PUNR.

Imputation of single income categories. As is standard in the welfare economics literature, our target income measure is annual postgovernment 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, aggregated across all household members.[7] However, instead of imputing just a ''lump sum'' of missing income, we aim at imputing six individual gross income components that are directly compatible with the more detailed information collected in the standard SOEP questionnaire every year.[8] This allows aggregation at the household as well as the tax unit level to match the very same income aggregates for all individual observations, whether PUNR or successfully interviewed.[9] This aids in the final simulation of direct taxes and social security contributions, which explicitly requires considering the interdependence of income and tax calculations for joint tax filers. In so doing, we can derive a consistent measure of ''household postgovernment income'' as the major source for inequality analyses. Finally, this procedure is likely to exert less bias in portfolio analyses than a ''lump sum'' or ''flat factor'' approach would.

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 and 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):

1. Labor income (the sum of all incomes from dependent employment, self-employment, secondary jobs, including extra pay such as Christmas or discretionary bonuses, etc.)
2. Pensions (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, including any survivor benefits in both systems)
3. Unemployment compensation (the sum of assistance received through the unemployment insurance scheme, unemployment assistance, and subsistence allowance from the labor office)
4. (Public) student aid
5. (Public) maternity leave benefits
6. Private transfers (including alimony)

Imputation of income received and amount received. For each of the components mentioned above, we employ a two-step imputation procedure. First, we need to impute 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.[10] In both cases, we make use of longitudinal information, if at all available, which has been shown to clearly improve the quality of imputation (see, e.g., Spiess and Goebel 2003 using ECHP, Starick and Watson 2007 using HILDA, Frick and Grabka 2010 using SOEP, HILDA, and British Household Survey data). We separate all observations with PUNR into four groups, depending on the availability of information from the previous or following year, from both years, or from neither. In fact, for any PUNR with missing information in year *t* and a successful interview in *t-1*, we can derive valuable and highly predictive information for the target information in *t* from previous year's income receipt at the time of the interview when he or she was asked about his or 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 probit model 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, relationship to the household head,[11] and a range of household context information.[12] For observations with such longitudinal data, we also include income and employment status from the adjacent waves. While longitudinal information, if available, allows us to account for otherwise unobserved heterogeneity, we are also able to make use of the net monthly household income ("income screener") provided at the household level. In conjunction with the reported individual income information from successfully interviewed household members, this information allows us to infer the magnitude of missing personal income. Hence, by employing this and other relevant household-level information in the regressions, we are able to account for nonrandom, that is, endogenous nonresponse behavior.

If the predicted probability for receipt of a given income exceeds a randomly chosen threshold (drawn from a normal distribution with mean 0.5 and standard deviation 0.2 to consider uncertainty of the underlying imputation process), we assign a value of one to the respective filter indicator.[13] 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 received private transfers, and finally we use all five filter variables to predict receipt of labor income.[14] It should be noted that the predictive power of these models is very high, especially when introducing longitudinal data: For example, 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 again make use of longitudinal data in the imputation process by applying the row-and-column imputation as described in Little and Su (1989; hereafter L&S).[15] The L&S technique takes advantage of cross-sectional as well as longitudinal information—using income data from a seven-year (or wave) shifting window around the point in time with the missing data—up to three years before and after the occurrence of PUNR. Assuming that information obtained from observations more distant from the missing data point is less strongly correlated to the missing information, we assign decreasing weights to more widely separated information.[16]

The row-and-column procedure proposed by L&S is carried out as follows: The column effect cj is defined by the relative cross-sectional annual income for each of the seven (k) waves of our shifting window, thus capturing simple period effects: $c_j = Y_j / \frac{1}{k} \sum_{i=1}^{k} Y_i$ (with Yj denoting total incomes at wave j and k being seven years). Second, the row effect ri is based on the longitudinal income data collected from a given individual i; thus, it gives the within-person average income position $r_i = \frac{1}{k} \sum_{j=1}^{k} \frac{Y_{ij}}{c_j}$ (with k denoting the number of valid income values for individual i and Y_j denoting the average income at wave j and c_j the simple wave correction from the column effect).[17] Multiplying the column effect cj with the row effect ri yields the expected income position of individual i at wave j: $E_{ij} = ri * cj$. Finally, all individuals are ordered by ri and a stochastic component is introduced by considering the deviation of the expected and the observed value from the nearest neighbor with valid information in the respective wave. Combining those three effects yields an estimate for the missing income $Z_{ij} = E_{ij} * [Y_{nj} / E_{nj}]$ (with Y_{nj} and E_{nj} denoting the observed and the expected income of the individual n, defined as the nearest neighbor of 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 are 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 aid, 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, 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 other (five) components, as well as the same set of variables used to predict the filter information. 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 to avoid a regression-to-the-mean effect when making use of standard ordinary least squares regression and to retain the original variance of the income data (Copas 1997).

Results of the imputation process. A straightforward assessment of the overall impact of these imputation procedures is given in Figures 4a–4c 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 euros) conditional on receipt, and (c) the resulting mean values for the entire population.

According to Figure 4a and in line with the regression results on the selectivity of PUNR presented above (Table 1), nonparticipating 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 percent of all individuals with PUNR, while only 60 percent to 70 percent 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 and Su (1989) and those where a purely cross-sectional approach was necessary because of 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
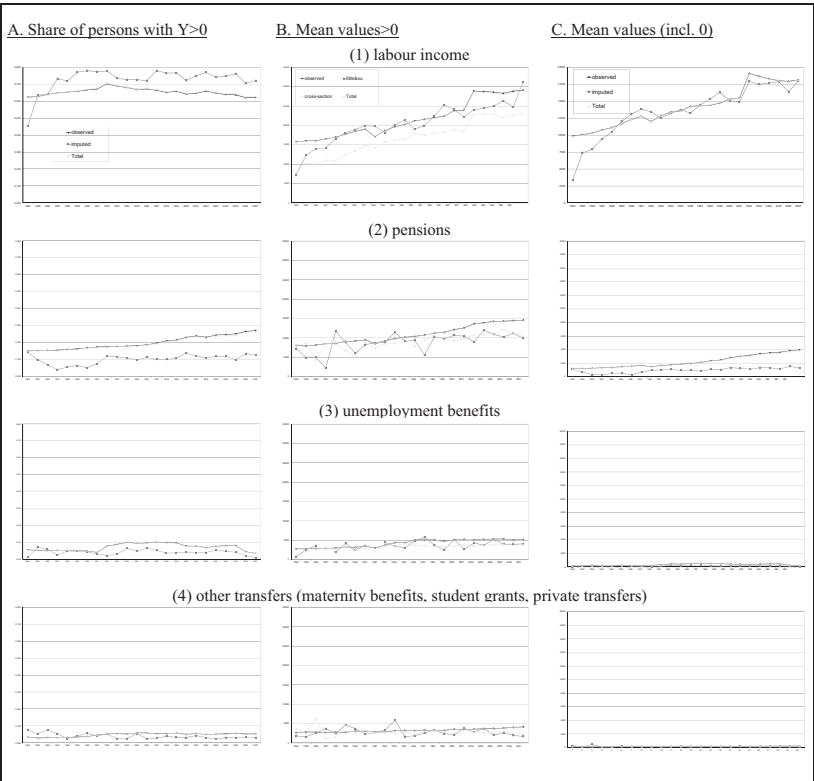
Figures 4. The impact of imputation: Population share holding income components and average income: (a) share of persons with Y > 0; (b) mean values > 0; (c) mean values (including 0)
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

imputed individuals. This may partly reflect 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 a type that requires applying the L&S imputation procedure.

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

Thus, before carrying out the welfare 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 postgovernment income.

Intermediate Conclusions and Hypotheses

What impact might the choice of one of the alternative treatments (*ignorance, adjustment, elimination, imputation*) exert on results of inequality and poverty analyses based on micro data affected by PUNR? First of all, given the secular trend toward an increasing incidence of PUNR over time (see Figure 1 above), one should expect any bias (in measures of inequality, poverty, intrahousehold distribution, and aggregates) arising from PUNR to be increasing over time, that is, with the duration of the panel. Second, the selectivity of PUNR as shown above clearly challenges the basic assumption underlying the various approaches: Version 1 ''ignoring PUNR'' (thus assuming $Y[PUNR] = 0$) appears the least plausible given that PUNR is 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 nonparticipating members to mirror those of their fellow household members who were interviewed ($Y[PUNRs] \sim Y[noPUNRs]$ within PUNR-HH). Finally, there may be no clear-cut answer about distributional effects arising from choosing Version 3 ''elimination and reweighting'' (i.e., incomes of PUNR households are equal to those of non-PUNR households, $Y[PUNR-HH] \sim Y[noPUNR-HH]$) or Version 4 ''imputation of PUNR'' ($Y[PUNR] = f[X] + e$). It can be assumed, however, that the more similar the variables used in the reweighting scheme and in the imputation procedure are, the more similar the treatment effects on inequality and poverty will be. In any case, while both approaches appear clearly less normative than Approaches 1 and 2 because of (adequately) controlling for selectivity, the imputation approach should be chosen for four reasons: First, imputation retains the complete panel population, thus making the data easier for less experienced data users to handle. Second, by preserving the complete population, imputation should also lead to lower standard errors than is the case with reweighting approaches, which usually reduce efficiency. Third, only the imputation procedure is capable of accounting for random as well as nonrandom mechanisms of nonresponse. Fourth, and most important for a panel survey, the imputation procedure facilitates mobility research.

Empirical Analyses: Inequality and Mobility Effects

The following empirical analyses are based on annual postgovernment income received in the calendar year preceding the survey year (including a measure of net imputed rent; see Frick and Grabka 2003). For cross-temporal comparability, we express all incomes in 2000 prices, also correcting for regional purchasing power differences between East and West Germany up to the mid-1990s. To adjust for different income needs across households because of differences in size and age composition, we apply the modified Organisation for Economic Co-operation and Development equivalence scale, assuming adult household members (older than 14) to have 50 percent of an adult's income needs and children up to age 14 to have 30 percent of those needs. In the following, we 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 also not only try to identify the degree to which results coincide for the entire population but also look at the consistency of those alternatively derived measures for each individual. Here, it will be important to find out not only whether the two approaches yield, for example, a similar share of individuals at risk of relative income poverty but also whether these results are identical for the very same persons.

To provide a more robust picture, we apply a range of established indicators used in the literature. We measure income inequality by means of the Gini coefficient, the mean log deviation, which is more sensitive to changes at the lower end of the income distribution, and the top-sensitive half-squared coefficient of variation. Relative income poverty is measured based on a poverty threshold of 60 percent of the median. To also identify eventual effects *within* the population defined 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, and particularly important for addressing the relevance of these alternative treatments for panel research, income mobility is measured on the basis of the measures introduced by Fields and Ok (1999) and by Shorrocks (1978).[18]

Inequality and Poverty

Hypothetical effects. What are the hypothetical effects of treating PUNR, in whatever way, on income levels, relative poverty, and inequality? First of

all, any explicit accounting for PUNR should yield *higher* incomes among those affected by PUNR; 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 (*adjustment*) 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 these two contradictory effects, the overall (net) effect of PUNR treatment on relative poverty risks at a *given point in time* remains unclear. However, because of 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 *sociodemographic 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 that bear the risk of PUNR but that were nevertheless completely interviewed.[19]

How does PUNR affect income levels, inequality, and poverty? Putting numbers to those considerations, we now turn to a comparison of results from inequality analyses based on the four approaches (labeled Version 1: ignoring, Version 2: adjusting needs, Version 3: deleting and rescaling, 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 postgovernment income. Apparently, in the early waves of the panel, when PUNR was a rather rare event, there was 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 above, 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 means 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 and rescaling weights) and Version 4 (imputation) yield the highest income levels—while being very similar in scope. As can be expected
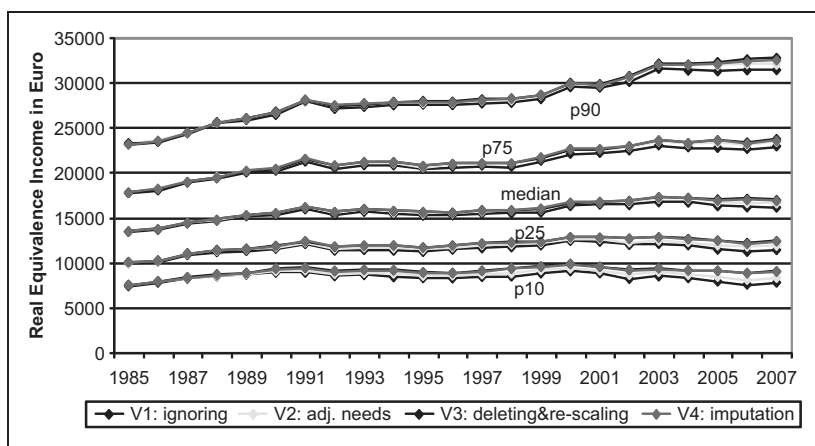
Figure 5. The impact of PUNR treatment on the distribution of equivalent income
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

from the increasing deviation between top and bottom income levels in Figure 5, there is a secular trend toward increasing income inequality in Germany, which has accelerated substantially since the turn of the millennium (see Grabka and Frick 2008).

Figure 6 confirms this finding using various inequality indicators. More important for our argument, however, we observe again an increasing gap among the four treatments, with Version 1 showing the highest degree of 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.

Finally, we present results on relative income poverty using the parametric family of Foster-Greer-Thorbecke (1984) measures. Confirming the hypotheses laid out in the Hypothetical Effects section, the results for the poverty head count ratio (FGT0) in Figure 7 are highest for the Version 1 (ignoring), 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!

All these 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).
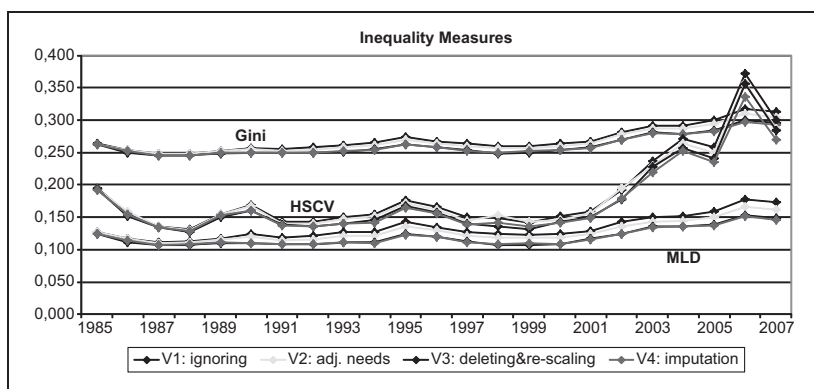
Figure 6. The impact of PUNR treatment on income inequality
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

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 the results were more similar, 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 of ''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 know more about the socioeconomic structure of the population 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 micro data used. If that were the case, for example, we would expect child poverty to look different if PUNR were mostly a problem among house-holds with dependent children, thus misreporting their income and making them appear poorer than they actually are.

In Figure 8a, we restrict our sample to those who live below the poverty line according to Version 4 (imputation). The time series graphs show the share of those identified as *non*poor according to the three other approaches. In line with the analysis results so far, there is a high degree of concordance
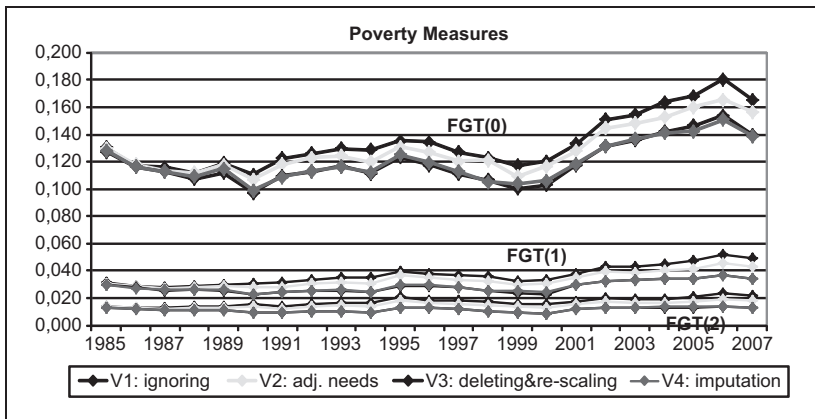
Figure 7. The impact of PUNR treatment on relative income poverty *Source*: SOEP. Authors' calculations.
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

with the identification of poverty in Version 3 (elimination and reweighting), although the population used in Version 3 is more selective, as illustrated by the share of those missing from the analysis which accrues up to about 10 percent of the baseline population. Clearly less comparable are the results based on Version 1 and Version 2 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 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 percent in the most recent waves. The results obtained from Versions 1 and 2 are significantly less comparable, and differences again are growing over time.

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 of the degree to which this is true for longitudinal analyses as well by comparing results derived from the four methods with respect to poverty and income mobility.
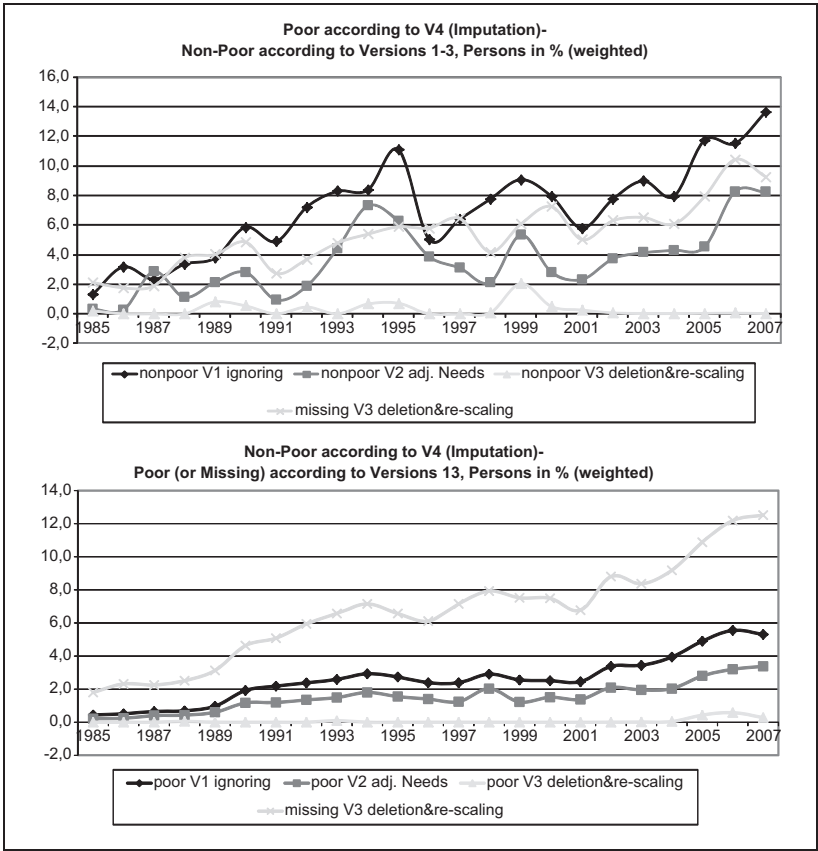
Figure 8. Consistency of poverty measurement using alternative approaches to handle PUNR
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

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 into or out of poverty within a given two-year interval. 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 affected
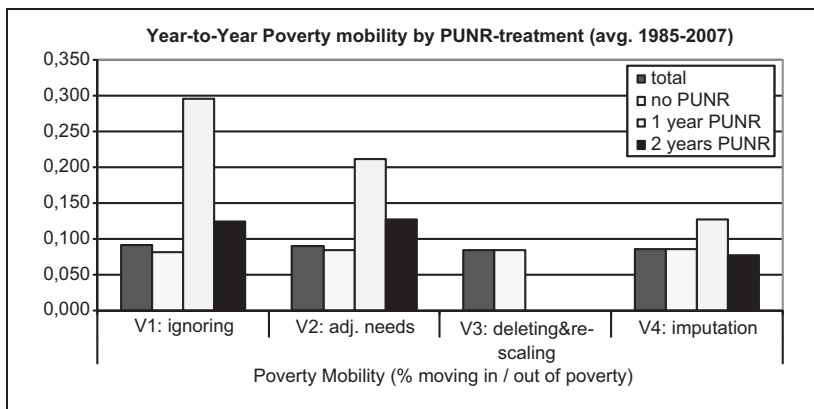
Figure 9. The impact of PUNR treatment on poverty mobility
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

by PUNR in both years. Comparing the four techniques, the selectivity of the population in Version 3 (elimination and rescaling) becomes apparent. Indeed, one may argue that maintaining the entire survey population, that is, 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 much higher mobility rates, simply because PUNR in at least one wave increases the probability of being poor because of understated incomes as shown above in the Inequality and Poverty section. Results based on imputed data (Version 4) still exhibit above-average mobility, in particular if PUNR was present only in one wave. On one hand, this partly reflects the uncertainty of the underlying imputation procedure (i.e., we inject variation by adding residuals to the predicted incomes); on 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 that interfered with survey participation. If the latter were the case, then indeed the mobility results in Version 3 (elimination and rescaling) would be downward biased.

Using the same data, Figure 10 gives very consistent results with respect to income mobility over two years when applying the Fields and Ok (1999) index. It should be noted that the results presented here are insensitive to the
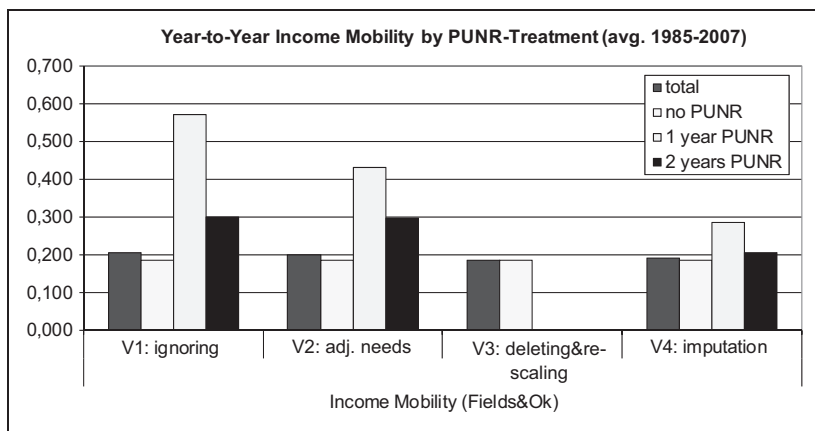
Figure 10. The impact of PUNR treatment on two-wave income mobility
Source: German Socio-Economic Panel (SOEP)v24, authors' calculations.

choice of the mobility measure; applying, for example, the Shorrocks (1978) index (using the Gini coefficient) yields more or less identical results (available from the authors on request).[20]

Conclusions

Using 24 waves of panel data from the German SOEP, we find an increasing incidence of PUNR together with clear indications of the selectivity of PUNR. A major consequence of this phenomenon is a systematic downward bias in the level and development of income inequality and relative poverty, whereas income mobility will be overstated because of people moving into and out of PUNR. Our strategy of imputing single income components as well as including them in the estimation of taxes and social security contributions appears to be an appropriate means to cope with both of these problems, thus guaranteeing a less biased measure of postgovernment income as the empirical basis for analyses of income inequality and poverty.[21] The imputation of various components instead of just adjusting the ''annual postgovernment 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 that exclude those affected by PUNR and reweight those at risk of PUNR). Finally, from

a theoretical perspective, the imputation strategy outperforms all other approaches because it makes it possible to account for nonrandom nonresponse behavior. From a practical perspective, it enhances the usability of the data, a matter of utmost importance to data providers.[22]

Future research will have to address the following questions:

- Is there a correlation between PUNR und subsequent attrition (UNR) that would be relevant for mobility analyses?
- Does the choice of PUNR treatment affect comparability in cross-country comparisons (see Frick and Grabka 2010 for the need to harmonize the procedures used for the imputation of INR)?
- While our analysis dealt with the missing contribution of individuals to their respective household's resources, PUNR may also yield a similar bias in other research areas, such as labor economics, where the interaction among household members is of crucial importance—for example, when modeling labor supply decisions of couples. The missing mechanism for PUNR may not be random at all if the individual is unable to participate in the survey because, for example, he or she is earning a great deal of money drilling for oil on an offshore platform, or simply because he or she does not want to answer questions while unemployed or severely ill.
- Finally, what do our results imply for designing incentives targeted at increased participation in household panel surveys (see Laurie and Lynn 2009; De Leeuw, Hox, and Huisman 2003; Hill and Willis 2001)? Instead of correcting the micro data *after* data collection, one should instead try to prevent missing data from occurring. Thus, it might be preferential from a data collection point of view to consider ideas for early interventions, such as, first, collecting proxy information on individuals with restricted interview capability, for example, because of severe sickness, dementia, Alzheimer's, and so on or, second, increasing incentives to participate, for example, through monetary incentives. While this additional participation may exert positive spillover effects on other household members, it may also yield some habituation effect from a panel perspective: Interviewees who were paid extra money once may want to keep their ''bonus'' in future waves as well, making this approach a rather expensive one. One also can consider providing an additional incentive at the household level only if all adult respondents participate (this is done in the HILDA survey; see Watson and Wooden 2009). Third, an alternative might arise with a short drop-off questionnaire to be filled in by

PUNR respondents 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'' of the regular survey to reduce the response burden.

Above and beyond the arguments brought forward in this article, there may also be other reasons why PUNR is increasing in prevalence over the past few years than simply measurement issues. One argument arises from the increasing number of individuals who have *multiple residences*, for example, because of long-distance commuting between ''home'' and ''work'' or choosing ''modern'' lifestyles such *living apart together* (Asendorpf 2009*)*, that is, couples who do not share a common address but rather live in two separate places. By making it difficult to determine which household a given person belongs to and whether his or her resources and needs should be assigned to just one or (partially) to several households, these recent social developments make it ever more difficult to precisely define the concept of ''private households.''

Declaration of Conflicting Interests

Funding

Notes

1. This approach is in line with the unitary model proposed by Becker (1991), which assumes a set of coherent preferences across all household members. However, it should be noted that there is a substantial body of literature questioning the general validity of this assumption (see, e.g., Vogler and Pahl 1994; Bonke and Uldall-Poulsen 2007).
2. For a comprehensive discussion of nonresponse behavior in household surveys in general as well as in panel surveys in particular, see the various contributions in the readers edited by Groves and Cooper (1998), Groves et al. (2002), De Leeuw, Hox, and Dillmann (2008), and Lynn (2009).
3. It should be noted that all of the following analyses refer to the population in private households only, that is, we exclude individuals living in institutions such as nursing homes.

4. 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.

5. Without rescaling 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 and Grabka 2005 for the German Socio-Economic Panel [SOEP], Frick and Grabka 2010 for the U.K. British Household Survey [BHPS], the SOEP, and the Australian Household, Income and Labour Dynamics in Australia [HILDA] Survey, as well as Starick and Watson 2007, also using HILDA data).

6. See Nicoletti and Peracchi (2006) for a discussion of methodological problems arising from the flat-rate correction in the European Community Household Panel.

7. 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 nonmonetary income advantage of subsidized renters (see Frick and Grabka 2003).

8. Obviously, imputing missing income components is more complex than imputing aggregate individual income. However, more detailed income components not only offer greater research potential but also are needed to simulate tax and social contributions accurately given the complexity of the German tax system, with its various allowances and variable taxation of different income components.

9. According to German tax law, married couples file their taxes jointly, while all other individuals are single filers.

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

11. 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 is not always the fault of the survey designers but may simply result from ethical and legal restrictions on the collection of 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 nonparticipating household members is by and large not feasible. But even if the collection of proxy information were allowed, one would expect a relatively strong misspecification bias arising from the restricted knowledge base of the reporting individual with respect to the missing income information of the respective partial unit nonresponse (PUNR).

12. **Demographic variables** (D = *dummy variable*) include 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-subsample identifier, change in household composition (D). **Social structure** variables include home owner (D), (highest) education of head of household or spouse/partner, education of children, migration background (D). **Income variables** include net household income at month of interview, public transfers (housing subsidies, social assistance, etc.), aggregated observed individual incomes as a share of net household income at month of interview. **Filter(s)** include considering receipt of other income components (D).

13. Using a static threshold of 0.5 for assigning the filter to either zero or one would essentially ignore the nature of probabilities and result in biased predicted shares of individuals with or without the income component. Thus, we use random thresholds. We fixed the mean of the randomly drawn thresholds at 0.5, the default in probability predictions, and decided on a mean variation of 0.2 around that mean.

14. The rationale for a specific order when imputing the six income components follows their relative incidence. However, one exception is labor income. Robustness checks altering the order of the imputation process lead to an overestimation of labor income, which suppresses the relative incidence of all other income components compared to the incidence of the observed cases.

15. In an evaluation of various imputation methods, Starick (2005) argues in favor of the L&S methods rather than the standard (single) imputation techniques. Furthermore, the L&S methods perform better in maintaining cross-wave relationships and income mobility. This finding is also confirmed by Frick and Grabka (2005) for the SOEP. L&S is also the standard procedure for imputation of item nonresponse (INR) on income questions in the SOEP (see Frick and Grabka 2005) and HILDA (Starick and Watson 2007).

16. In more detail, the income information is weighted by $2^{(3-dt)}$, with dt denoting the distance in years, ranging from 1 to 3.

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

18. All empirical results presented in this article are based on calculations using Stata (Version 9.2). We gratefully acknowledge Stephen P. Jenkins and Philippe van Kerm for their Stata add-ons INEQDECO, INEQUAL7, IMOBFOK, FOKMOB, and SHORMOB.

19. To give an example, the head-count poverty rate (based on 60 percent of national median income, 2007) for persons living in households with children varies from 21 percent for ''ignoring'' and 19 percent for ''adjusting needs'' to about 15 percent for ''deleting and rescaling'' and ''imputing.''

20. These findings for a two-year interval are confirmed when analyzing five-year intervals. The lowest degree of income mobility among PUNR households can be found when using imputed data instead of applying Versions 1 or 2, while

''elimination'' of PUNR households from Version 3 yields the lowest degree of mobility for the entire population.

21. Future extensions may consider *multiple* imputation (Rubin 1987), which would also take into account the uncertainty embedded in the single two-stage imputation procedure applied here.

22. This raises the question of whether such corrections should be made by data providers or by the users themselves. The main advantage of in-house correction is that data providers usually have more experience in handling ''their'' information and are also aware of the data's shortcomings and potential restrictions. Another aspect is the availability of expertise in the field of (longitudinal) imputation. Finally, data generation is time-consuming, which could discourage users from working with these kinds of micro data. On the other hand, since imputation is associated with normative decisions, transparent documentation of the methods used and decisions made is crucial to ensure that users are properly informed.

References

Asendorpf, J. B. 2009. ''Living Apart Together: Alters- und Kohortenabhängigkeit einer heterogenen Lebensform'' (Living Apart Together: Age- and Cohort-dependence of a Heterogeneous Life Style). *Kölner Zeitschrift für Soziologie und Sozialpsychologie Cologne Journal for Sociology and Social Psychology* 60:749-64.

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

Barceló, C. 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 No. 0829, Banco De Espana, Madrid.

Bastien, J.-F., C. Cauchon, W. Jocelyn, A. Lim, and M. J. Yang. 2010. ''Data Quality for the 2008 Survey of Labour and Income Dynamics (SLID).'' Statistics Canada, Ottawa.

Becker, G. S. 1991. *A Treatise on the Family*. Enlarged ed. Cambridge, UK: Cambridge University Press.

Bonke, J. and H. Uldall-Poulsen. 2007. ''Why Do Families Actually Pool Their Income? Evidence from Denmark.'' *Review of Economics of the Household* 5:113-28.

Buhmann, B., L. Rainwater, G. Schmaus, and T. M. Smeeding. 1988. ''Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (LIS) Database.'' *Review of Income and Wealth* 34:115-42.

Canberra Group. 2001. *Final Report and Recommendations*. Retrieved from http://www.lisproject.org/links/canberra/finalreport.pdf

Copas, J. B. 1997. ''Using Regression Models for Prediction: Shrinkage and Regression to the Mean.'' *Statistical Methods in Medical Research* 6:167-83.

De Leeuw, E. D., J. J. Hox, and D. A. Dillmann, eds. 2008. *International Handbook of Survey Methodology*. New York: Taylor & Francis.

De Leeuw, E. D., J. J. Hox, and M. Huisman. 2003. ''Prevention and Treatment of Item Nonresponse.'' *Journal of Official Statistics* 19:153-76.

De Luca, G., and F. Peracchi. 2005. ''Survey Response.'' in *Health, Ageing and Retirement in Europe—First Results from the Survey of Health, Ageing and Retirement in Europe*, edited by A. Boersch-Supan, A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist, and G. Weber. Mannheim, Germany: MEA:355.

European Communities. 2003. ''The EC Household Panel Newsletter (01/02).'' Office for Official Publications of the European Communities, Luxembourg.

Eurostat. 2000. ''Imputation of Income in the ECHP.'' DocPan 164/00, Luxembourg.

Eurostat. 2010. ''An Assessment of Survey Errors in EU-SILC.'' 2010 ed. Methodologies and working papers, Luxembourg.

Fields, G. S. and E. Ok. 1999. ''Measuring Movement of Incomes.'' *Economica* 66:455-71.

Foster, J., J. Greer, and E. Thorbecke. 1984. ''A Class of Decomposable Poverty Measures.'' *Econometrica* 52:761-66.

Frick, J. R. and M. M. Grabka. 2003. ''Imputed Rent and Income Inequality: A Decomposition Analysis for the U.K., West Germany, and the USA.'' *Review of Income and Wealth* 49:513-37.

Frick, J. R. and M. M. Grabka. 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, J. R. and M. M. Grabka. 2010. ''Item Non-Response and Imputation of Annual Labor Income in Panel Surveys from a Cross-National Perspective.'' Pp. 355-74 in *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*, edited by J. A. Harkness, B. Edwards, M. Braun, T. P. Johnson, L. E. Lyberg, P. P. Mohler, B.-E. Pennell, and T. Smith. New York: John Wiley.

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

Groves, R. M. and M. P. Couper. 1998. *Non-Response in Household Interview Surveys*. New York: John Wiley.

Groves, R. M., D. A. Dillman, J. L. Eltinge, and R. J. A. Little, eds. 2002. *Survey Nonresponse*. New York: John Wiley.

Hayes, C. and N. Watson. 2009. ''HILDA Imputation Methods.'' HILDA Project Technical Papers Series No. 2/09, December 2009 (Revised January 2010).

Hill, D. and R. J. Willis. 2001. ''Reducing Panel Attrition: A Search for Effective Policy Instruments.'' *Journal of Human Resources* 36:416-38.

Jenkins, S. P. 2010. ''The British Household Panel Survey and its income data.'' ISER Working Paper Series No. 2010-33, University of Essex, Colchester.

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, DIW Berlin.

Laurie, H. and P. Lynn. 2009. ''The Use of Respondent Incentives on Longitudinal Surveys.'' Pp. 205-34 in *Methodology of Longitudinal Surveys*, edited by P. Lynn. London: John Wiley.

Little, R. J. A. and H. L. Su. 1989. ''Item Non-Response in Panel Surveys.'' Pp. 400-425 in *Panel Surveys*, edited by D. Kasprzyk, G. Duncan, and M. P. Singh. New York: John Wiley.

Loosveldt, G., J. Pickery, and J. Billiet. 2002. ''Item Non-Response as a Predictor of Unit Non-Response in a Panel Survey.'' *Journal of Official Statistics* 18:545-57.

Lynn, P., ed. 2009. *Methodology of Longitudinal Surveys*. London: John Wiley.

Nicoletti, C. and F. Peracchi. 2006. ''The Effects of Income Imputation on Microanalyses: Evidence for the European Community Household Panel.'' *Journal of the Royal Statistical Society Series A* 169:625-46.

Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley.

Shorrocks, A. F. 1978. ''Income Inequality and Income Mobility.'' *Journal of Economic Theory* 19:376-93.

Smeeding, T. M. and D. H. Weinberg. 2001. ''Toward a Uniform Definition of Household Income.'' *Review of Income and Wealth* 47:1-24.

Spiess, M. and J. Goebel. 2003. ''Evaluation of the ECHP Imputation Rules.'' CHINTEX Working Paper No. 17.

Starick R (2005). ''Imputation in Longitudinal Surveys: The Case of HILDA.'' Research Paper of the Australian Bureau of Statistics. ABS Catalogue no. 1352.0.55.075, Canberra.

Starick, R. and N. Watson. 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* 42:236-88.

Wagner, G. G., J. R. Frick, and J. Schupp. 2007. ''The German Socio-Economic Panel Study (SOEP)—Evolution, Scope and Enhancements.'' *Schmoller's Jahrbuch—Journal of Applied Social Science Studies* 127:139-69.

Watson, N. and M. Wooden. 2009. ''Identifying Factors Affecting Longitudinal Survey Response.'' Pp. 157-82 in *Methodology of Longitudinal Surveys*, edited by P. Lynn. London: John Wiley.

Bios

Joachim R. Frick†was deputy director of the Socio-Economic Panel (SOEP) and head of the SOEP Research Data Center (SOEP-RDC) at the DIW Berlin Germany and acting professor at the Technische Universitaet Berlin.

Markus M. Grabka is senior researcher at the department of the Socio-economic Panel (SOEP) at the DIW Berlin Germany and the Technische Universitaet Berlin.

Olaf Groh-Samberg is Junior-Professor for Sociology at the University of Bremen and DIW Research Professor. He currently works at the Bremen International Graduate School of Social Sciences at the University of Bremen.