Review of Income and Wealth
Series 44, Number 1, March 1998

THE ESTIMATION OF POVERTY DYNAMICS USING DIFFERENT MEASUREMENTS OF HOUSEHOLD INCOME

BY ULRICH RENDTEL

University of Frankfurt

ROLF LANGEHEINE

IPN-Institute for Science Education, Kiel

AND

ROLAND BERNTSEN

University of Frankfurt

If surveys offer two different measurements of household income, one can use them simultaneously to identify the potential effects of measurement error on the observed-income mobility of the poor. In this paper we investigate transition tables between subsequent income states. Latent Markov models are used to model incorrect classifications of income states. Misclassifications are interpreted as measurement error or spurious changes that are not consistent with a simple transition table model. The empirical results for the German Socio-Economic Panel (GSOEP) show that the observed transition tables overestimate the mobility between poverty states.

1. INTRODUCTION

In household panel surveys there are two methods to measure household income. Primarily, one may ask the head of the household to self-assess the total household income. Alternatively, one may ask all household members for their individual incomes and transfer payments. Here, the household income is obtained by adding all individual components to a computed household income. Both measurements may be taken as indicators of the true household income. Since these measurements will not coincide, in general, measurement error is a common tool for the explanation of such differences.

Poverty dynamics are frequently analyzed in the framework of transition tables between subsequent poverty states (see Bane and Ellwood, 1986). If, due to a possible measurement error, a poverty state is incorrectly indicated a sequence such as poor/non-poor/poor may appear to be simply an imperfect measurement of "always poor." Now, if the two alternative poverty measurements derived from the reported and the computed income are indicators for the same poverty state, we may used them simultaneously. In this case, a sequence like poor/non-poor/

Note: The authors thank Greg Duncan, Bruce Headey, Jim Witte and two anonymous referees for helpful comments.

poor from measurement 1 is accompanied by a second sequence from measurement 2. If the second measurement also yields poor/non-poor/poor this provides evidence for a true change between poverty states. In case of non-corresponding poverty sequences, one would conclude that measurement error may be present.

We use latent Markov chain models with one and two manifest indicators (Langeheine and Pol, 1990, 1992, 1993, 1994) in order to assess the reliability of the two measurements and the impact of measurement errors on the observed income mobility of the poor. Such models have been used for the analysis of transitions between employment states (Abowd and Zellner, 1985). These models generalize the turnover tables between subsequent poverty states. The entries of such tables are interpreted as probabilities to slip into or out of poverty over one time period. Consequently, the risk over longer periods is computed by the product of the risk between subsequent periods. This is equal to assuming that the poverty states form a Markov chain.

The latent Markov chain model assumes that the true poverty states behave like a Markov chain. The observed poverty states are linked to the true states by response matrices. These response matrices reflect the probabilities to observe the manifest poverty states for different true (or latent) poverty states. If the true poverty states behave like a Markov chain and no measurement error is present the response matrices are equal to the unit matrix.

The response matrices may be different from the unit matrix for two reasons (a) there is measurement error, or (b) the true poverty states do not behave like a Markov chain. In the second case, the response matrices measure the spurious transitions between the poverty states that are not explained by the Markov model.

Spurious changes in income have been a major point of interest in the discussion of the shape of the income distribution and measures of inequality, like the Gini coefficient (Slemrod, 1992 and Shorrocks, 1978). Within the context of income inequality it is expected that spurious changes generate a higher mobility of earnings and, as a consequence, a higher inequality. In the context of discrete poverty states the latent Markov chain approach seems to be a useful tool to treat spurious changes.

The potential pitfalls of latent Markov models, namely the assumption of independent misclassifications, have been analyzed by Skinner and Torelli (1993). The existence of two different measurements for the same true value makes it possible to overcome this remedy.

The data we use in this analysis come from the first 4 waves of the German Socio-Economic Panel (GSOEP) and cover the years 1984, 1985 and 1986.

The article is organized as follows: Section 2 gives a brief description of the GSOEP database and the definition of the sample. Section 3 specifies the self-assessed and the computed household income in the GSOEP and discusses their potential sources of measurement error. Section 4 gives the definition of the poverty states used here. Section 5 presents a comparison of the observed transition tables between the poverty states. Then we introduce the latent Markov models and report our estimation results for these models. We discuss a separate treatment of measurement error in Section 6, followed by a model that uses both measurements simultaneously but assumes independent measurement errors in Section 7.

The model of Section 8 relaxes the assumption of independent measurement errors across waves. Section 9 concludes.

2. The Database

The GSOEP is a household panel survey which started in 1984 with 6,000 West-German households. All household members older than 15 are interviewed. Every year a new panel wave was launched. The main subjects of this ongoing survey are income and labor force participation. A brief description of the GSOEP can be found in Wagner *et al.* (1993) and the references cited there.

For the analysis we used a balanced data file of households where both measurements are available for the years 1984, 1985 and 1986. In order to compile the relevant information we had to use the data from the 1987 file (= wave 4) also, since the individual incomes are recorded in a retrospective calendar.

The use of a balanced data file excluded newly founded households and households that died, moved abroad or refused to participate during the first four years of the GSOEP. The sample size of this longitudinal sample was 3,944. For the computation of the poverty lines, however, we used for each indicator and each year the cross-sectional sample of all households where the respective indicator was known. This results in sample sizes of about 5,000 households.

3. Household Incomes in the GSOEP

At the household level the questionnaire of the GSOEP asks for the current monthly household income. At the individual level the GSOEP respondents were asked to provide their "average" monthly earnings. These individual "average" monthly figures were the basis for the computation of the total annual household income which is divided by 12 in order to be comparable with the figures from the household questionaire.[1]

The use of longer reference periods has been a major tool in removing the transitorial income fluctuations, see for example Benus and Morgan (1975), Headey and Krause (1995) and Slemrod (1992). In all cases, however, the empirical findings suggest that results on inequality or transition behavior are stable with respect to different length of the reference period. Thus we may expect that these conceptual differences are of minor importance.

3.1. *Reported Income*

In each wave of the GSOEP the head of the household is asked: "If you add everything together, what is the total monthly net income of all household members todays? Please report the monthly net amount, i.e., the amount after tax and social insurance. Add to the amount regular payments such as housing assistance, child payments, student aid, alimony payments, etc. In the case of 'I don't know', please estimate the amount."

[1]Average is used with quotation marks because it is an estimated value which may not be the exact average over the year.

With respect to measurement error three points are noted. First, the respondent was asked to add the contributions of *all* household members. However, there may be situations where the respondent is not well informed about the income of all household members or some incomes may not be regarded as part of the household pool, for example, the income of persons who don't belong to the family of the respondent or the income of children. For instance, Schwarze (1995) obtained the result that a change in the reporting person has a systematic impact on the variance of the household income profile, even if the household composition does not change during the analysis and after control for the number of gainfully employed persons in the household.

Second, the respondent was asked to include *regular* transfers, which is a vague formulation open to the respondents interpretation. Third, they were asked for the *net* income in a certain month. If the household belongs to a self-employed person, only a guess of the net income is possible since the annual net income is only known after taxation.

3.2. *Computed Income*

The GSOEP questionnaire asks all adult members of a sampled household for their individual earnings in a retrospective monthly calendar. For each activity of the preceding year they are asked to provide "the *average* monthly gross income." The term "average" may also introduce some measurement error because it is not clear how the respondents estimate the average. One may provide the (correct but difficult to calculate) weighted mean or the unweighted mean of two or more amounts.

There are, however, further sources of a potential measurement error in the computed annual income. The GSOEP asks the respondents for their gross earnings. For the assessment of the poverty status the net household income is relevant. Since the household and family composition is known, one may then try to estimate taxes in order to assess net annual income, see Berntsen (1992) for details.[2] Although the taxation rules are well known, the estimated tax is an apparent source of measurement error.

The computed household income in the GSOEP is conceptually linked to the annual household net income. Consequently, all annual gratifications, like a 13th-month salary, are included. Finally, in cases of owned flats the estimated cost for the rent was added to the household income.[3] This was done so as to be able to compare the welfare status of households with and without ownership of their flat. For comparison with the reported income the total annual income was divided by 12.

[2]In principle, it is also possible to ask respondents for the taxes they paid. However, when this information was explicitly collected it led to high non-response: 25 percent in wave 2 to 17 percent in wave 7. Also, in about one-third of the cases the taxation in year t was not yet known at the interview in year $t+2$. This is due to the possibility that in Germany the final assessment of taxes may be deferred for a time span up to two years. Thus, panel attrition also creates a relevant proportion of missing values. In these cases the individuals have to participate in over three panel waves in order to collect the information about gross income and its taxation from the questionnaire.

[3]The hypothetical rent for an owned flat was estimated on the basis of a statistical model, see Berntsen (1992) for details.

3.3 *Equivalence Income*

The household net income is not the only determinant for the assessment of poverty. The other determinant is the size and the composition of the household, which reflects the needs of its members. Therefore, we use the equivalence income obtained by dividing the household income by a weighted average of the household composition.

The use of equivalence incomes facilitates comparisons of incomes for household with different compositions. This is necesary not only for inter-household comparisons, but also for intra-household comparisons over time, if the household composition changes over time. The equivalence income is different from a simple per capita income. The GSOEP equivalance income takes into account the needs of the household members as well as economics of scale. The needs of children vary with their age. The scale of needs used in this analysis was in accordance with German legislation for social aid. At least theoretically there are numerous ways to calculate different equivalence scales, see Buhmann *et al.* (1988). However, as Rohwer (1991) demonstrated for the GSOEP, empirical results on income mobility are quite stable with respect to different equivalent scales.

The calculation of the equivalence income does not cause a problem with monthly household figures. However, in the case of annual figures it is no longer obvious how this has to be done if the composition of the household changed during the year. In some instances the simple rule of using the household composition during the time of the interview, which takes place approximately three months after the reference period of the annual income, may produce misleading results. However, in the case of the annual income there is no easy way out of this dilemma. Therefore, we used the above simple rule for our analysis.

Finally, there also appears to be severe rounding of amounts apparent in the data, see Rohwer (1991) and Rendtel and Schwarze (1991) for empirical findings with the GSOEP.

4. Definition of the Poverty States

In poverty research two different concepts, absolute and relative poverty, are used to define the poverty status of a household: one can use a poverty line which is defined by the requirements necessary to live at a certain level of welfare. Such a poverty line is based on absolute figures. The concept of relative poverty states that a household is considered to be poor, if it has less than a given percentage of the average equivalence income at its disposal. It is an open question how to choose the percentage of average income that defines the (relative) poverty line. Often the 50 percent level is used for the dichotomization of households into poor and non-poor. Here we use a more informative three-state description of poverty: less than 40 percent, between 40 percent and 60 percent, and above 60 percent of average equivalence income.[4]

[4]In order to study the effect of different categorizations, we also used the 2-category scheme as well as a 4-category scheme that splits the medium range into two separate categories: 40 percent to 50 percent and 50 percent to 60 percent. The results differ only slightly from the conclusions for the 3-category scheme displayed in this article, see Rendtel *et al.* (1992).

The two income measurements will give in general two different means for the equivalent income and consequently two different poverty lines. We decided to define the poverty state of a household according to the respective poverty line; i.e., the poverty state according to the reported income refers to the poverty line based on the reported income while the poverty state according to the computed income refers to the poverty line based on the computed income.[5]

5. Comparison of Observed Transition Tables Between Poverty States

To begin, we compare the shape of the distribution of the equivalence income for both measurements. Figure 1 displays a kernel estimate of the 1984 income distributions.[6] It appears that the two densities differ in their lower part only by an additive transformation. Note also the unsmooth behavior of the density for the reported household income, which indicates the presence of rounding effects. Since several components of the computed household income are estimated by a statistical model, rounding effects should play a minor part. This is confirmed by the smooth shape of the corresponding density in Figure 1.

From Figure 1 one might conclude that the reported income is always less than the computed income. Such a view would support the assumption that the

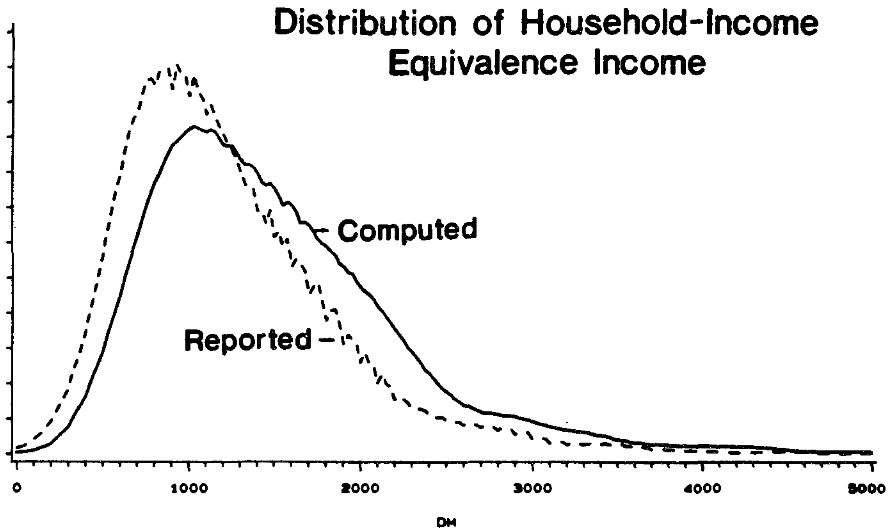

Figure 1. Comparison of the Distribution of Household Equivalence Income for Computed and Reported Income (1984) in the GSOEP

[5]The means of the equivalence income in 1984, 1985 and 1986 are: 1,523 DM, 1,586 DM and 1,671 DM for the computed household income and 1,274 DM, 1,307 DM and 1,373 DM for the reported household income.

[6]The kernel function was a normal density with a standard deviation of 100 DM. This value corresponds to about 0.13 of the standard deviation of the income distribution. The optimal smoothing factor is $1.06/n^{1/5} = 0.20$, if the density to be estimated is normal, cf., Silverman (1986). So there has been a slight undersmoothing in order to preserve a potential multimodality of the income distribution.

respondents underreport their household income. However, Figure 2, which presents the differences between the computed and the reported equivalence income, shows that there is no systematic ordering of the two income measurements. Figure 2 demonstrates that for all points in time the ordering of the two meaurements is reversed for a substantial part of the sample. The proportion of observations where this is true does not change over time.

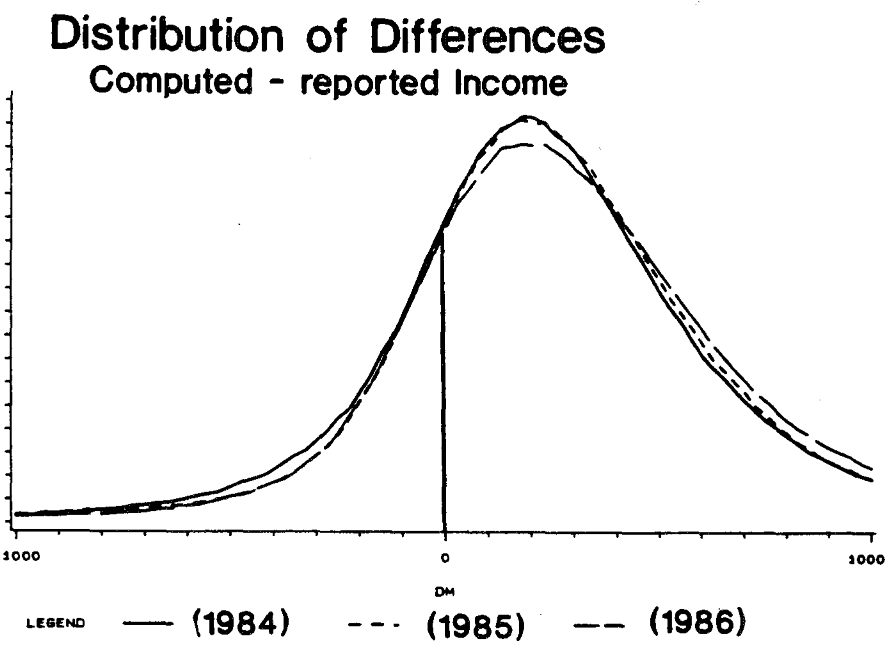

Figure 2. The Distribution of the Differences Between Computed and Reported Equivalence Income in the GSOEP, 1984 to 1986

We will now compare the transition tables $T_{2|1}$ between the poverty states in wave 1 and wave 2 and $T_{3|2}$, the corresponding transition table between wave 2 and wave 3. Table 1 compares the starting distribution and the estimated transition probabilities between subsequent poverty states for the computed and the reported income.

A comparison of the starting distributions reveals that the differences with respect to the level of poverty may be ignored. This result is to be expected from Figure 1 where the two income distributions seem to differ in their lower range only by an additive transformation.

Furthermore, with respect to mobility both measurements result in roughly the same transition tables between subsequent poverty states.[7] Both measurements predict a high probability of leaving the ≤ 40 percent state, which amounts to 0.52 and 0.48 for the computed income and 0.55 and 0.60 for the reported income.[8]

[7]Similar results may be shown in the case of two and four category schemes, see Rendtel *et al.* (1992).

[8]Due to rounding of results not all rows of the transition matrix sum to one.

TABLE 1

COMPARISON OF TRANSITION PROBABILITIES BETWEEN POVERTY STATES FOR COMPUTED AND REPORTED INCOME IN THE GSOEP

Standard Errors in Parenthesis Under the Estimate

Poverty State	Starting Distrib.	Transition 1 to 2 $T_{2\|1}$ ≤40%	40–60%	≥60%	Transition 2 to 3 $T_{3\|2}$ ≤40%	40–60%	≥60%
		Computed Income					
≤40%	0.05	0.48	0.34	0.18	0.52	0.32	0.16
	(0.003)	(0.036)	(0.034)	(0.028)	(0.028)	(0.036)	(0.028)
40–60%	0.16	0.10	0.59	0.31	0.08	0.61	0.31
	(0.006)	(0.012)	(0.019)	(0.018)	(0.011)	(0.019)	(0.018)
≥60%	0.79	0.01	0.06	0.93	0.01	0.05	0.94
	(0.006)	(0.001)	(0.004)	(0.004)	(0.001)	(0.004)	(0.004)
		Reported Income					
≤40%	0.06	0.46	0.35	0.20	0.41	0.37	0.23
	(0.004)	(0.033)	(0.031)	(0.026)	(0.033)	(0.033)	(0.028)
40–60%	0.17	0.10	0.52	0.38	0.10	0.53	0.37
	(0.006)	(0.012)	(0.019)	(0.019)	(0.012)	(0.019)	(0.019)
≥60%	0.77	0.01	0.08	0.91	0.01	0.06	0.92
	(0.007)	(0.002)	(0.005)	(0.005)	(0.002)	(0.004)	(0.005)

If we compare corresponding elements on the diagonal we will recognize that the computed income appears to be slightly more stable than the reported income. This might be expected from the longer reference period of the computed income. However, the higher stability may result from other causes, for example, the use of the mean value imputation for unknown household components.

Although both measurements exhibit roughly the same observed mobility behavior, we have assumed that there is no measurement error present in the assessment of the poverty status. This unrealistic assumption will be relaxed in the following sections.

6. A SEPARATE TREATMENT OF MEASUREMENT ERROR BY A ONE-INDICATOR MODEL

The inclusion of measurement error into the framework of Markov chains dates back to Wiggins (1955, 1973). The problem of estimating such latent Markov models was solved somewhat later, cf. Poulsen (1982), Bye and Schechter (1986) and Pol and Leeuw (1986) (see also Langeheine and Pol 1990). The latent Markov model may be presented by the path diagram of Figure 3.

The observed poverty status in each wave is represented by variables P_t ($t = 1, 2, 3$). The true but unobserved poverty status is represented by the latent variables Π_t ($t = 1, 2, 3$). The transition matrices $T_{2|1}$ and $T_{3|2}$ between the true poverty states therefore operate on the latent level. The connection with the observed levels is established by a response matrix R. The elements $r_{l,i}$ of R are the conditional

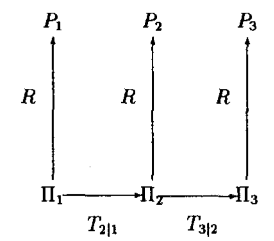

Figure 3. Path Diagram That Illustrates a Single Indicator Latent Markov Model for Three Waves

probabilities that a household is observed at poverty state i if its true poverty state is 1.

Thus, under the one-indicator latent Markov model we have:

$$\text{(1)} \qquad \Pr(P_1=i, P_2=j, P_3=k)=\sum_l \sum_m \sum_n \pi_l r_{l,i} T_{2|1}(l,m) r_{m,j} T_{3|2}(m,n) r_{n,k}$$

where $\Pr(\Pi_1=l)=\pi_l\ (l=1,2,3)$ is the starting distribution of the latent Markov chain and $T_{2|1}(l,m)$ and $T_{3|2}(m,n)$ are the elements (l,m) and (m,n) of the transition matrices $T_{2|1}$ and $T_{3|2}$. Note, that the manifest Markov model of the previous section is a special case of this latent Markov model with R being the identity matrix.

Note also, that equation (1) implies the independence of misclassifications from previous measurement errors. This is an unpleasant feature of this latent Markov model. Apparently any intertemporal dependence of misclassifications would reduce the models ability to explain an observed change in the poverty position as the result of a misclassification. Thus the latent Markov model used here has a tendency to explain observed poverty transitions as results of a misclassification of the true poverty state.

A direct answer to the question of the independence of measurement errors is not possible since the true poverty status is unknown. Only on rare occasions one is able to get additionally a high quality income measurement which may be taken as the true income. Such information was obtained in a validation study of the PSID questionnaire form. Here Bound *et al.* (1990) report in their Table 1 the correlation between subsequent measurement errors for the log of annual earnings. The correlation varied from 0.372 for a one year interval to 0.073 for a four year interval.[9] In the extended model of Section 8 the independence assumption will be relaxed.

Table 2 displays the estimated parameters of the one-indicator latent Markov model.[10] The fit for the simple Markov model (Table 1) is LR = 403.18 with df = 24, while the latent Markov model (Table 2) results in LR = 8.86 with df = 12.

[9]Bound *et al.*, do not report the standard errors of these figures. the correlation for a two-year interval, which is also relevant here, can be expected to be smaller than 0.372. However, the corresponding estimate is not reported by Bound *et al.*

[10]For computations the progam package PANMARK (Pol *et al.*, 1991) has been used. The estimates are ML and were obtained using the EM-algorithm.

Thus, there is no doubt that the manifest Markov model is inconsistent with the data.[11]

If we compare the results of Table 2 with those of Table 1, the estimates of the starting distribution π do not differ significantly. However, there are remarkable differences between the corresponding transition matrices: the probability to get out of the ≤ 40 percent poverty state is almost halved if we change from the observed to the latent level. Also, the risk of falling into poverty. i.e., leaving the ≥ 60 per cent state, dropped by a similar factor. Correspondingly the values on the diagonal raise substantially. For example, the probability of staying in the 40–60 percent state increases from about 0.55 to 0.80. Hence, the latent model suggests a much lower mobility on the true level than what is observed.

If we compare the transition matrices of the reported and the computed income in the latent Markov model, there are no major differences with one exception: transitions from the ≤ 40 percent poverty position. Here we find differences of about 7–9 percentage points. However the standard deviation of these parameter estimates are of the same size, so we may conclude that on the latent level both incomes yield quite the same results. Indeed, restricting the latent parameters to be equal across the two income measures yields a likelihood difference of 7.19 (df = 14).

On the other hand, if we restrict the response matrices to be equal across the income measures the likelihood ratio statistic increases by 23.1 (df = 6). These different consequences of restrictions imposed on the true poverty states and their observations suggest that the differences between the two income measures occur mainly because of differences with respect to the response matrices.[12]

7. A Joint Treatment of Measurement Error Using Two-Indicator Model

From the preceding section we may conclude that the transitions between the true poverty states as they are indicated by both measurement methods exhibit more or less the same behavior. Thus, we may assume that both indicators represent the same theoretical construct, the true poverty state. Hence, we may use both indicators jointly in a two-indicator model.

Latent Markov models with multiple indicators may be reduced to single indicator models with a special transition structure, see Langeheine and Pol (1992, 1993, 1994). The idea is demonstrated by the path diagram in Figure 4.

[11]Here, the likelihood ratio refers to a two group analysis, where the first group is defined by all observations based on the computed income and the second group is defined by the observations based on the reported income. For this reason the above difference of degrees of freedom is $12 = 2 \times 6 =$ (number of groups) $\times$ (number of free parameters for each response matrix).

However, one must not ignore the fact that the two corresponding estimates are not independent since they stem from the same households. For that reason test results concerning the equality of model parameters across the two measurements may be invalid. A separate analysis for each measurement method yields a likelihood ratio statistic of approximately half the size with half the degrees of freedom.

[12]The same analysis that is reported here for three income states was repeated for two and four income states. In these cases we got similar results. The latent Markov model achieves a good fit with the data, and differences on the latent level may be ignored, while differnces with respect to **R** are substantial, see Rendtel *et al.* (1992).

TABLE 2

ESTIMATES OF THE SINGLE INDICATOR LATENT MARKOV MODEL FOR COMPUTED AND REPORTED INCOME IN THE GSOEP

Standard Errors in Parenthesis Under the Estimate

Poverty State	Starting Distrib. π	Latent States: Transition 1 to 2 $T_{2\|1}$ ≤40%	40–60%	≥60%	Transition 2 to 3 $T_{3\|2}$ ≤40%	40–60%	≥60%
			Computed Income				
≤40%	0.06	0.69	0.19	0.12	0.77	0.16	0.08
	(0.008)	(0.061)	(0.059)	(0.033)	(0.079)	(0.082)	(0.034)
40–60%	0.16	0.07	0.79	0.14	0.05	0.81	0.14
	(0.011)	(0.027)	(0.038)	(0.029)	(0.024)	(0.037)	(0.030)
≥60%	0.78	0.00	0.03	0.97	0.00	0.01	0.99
	(0.009)	(0.001)	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)
			Latent States		Response Matrix R Observed States		
				≤40%	0.75	0.22	0.03
					(0.061)	0.057)	(0.017)
				40–60%	0.02	0.85	0.13
					(0.017)	(0.025)	(0.021)
				≥60%	0.00	0.02	0.98
					(0.001)	(0.004)	(0.004)
			Reported Income				
≤40%	0.05	0.76	0.22	0.02	0.68	0.28	0.04
	(0.009)	(0.090)	(0.085)	(0.050)	(0.079)	(0.089)	(0.041)
40–60%	0.17	0.05	0.81	0.15	0.05	0.82	0.13
	(0.014)	(0.030)	(0.045)	(0.036)	(0.028)	(0.046)	(0.035)
≥60%	0.77	0.00	0.03	0.97	0.00	0.01	0.99
	(0.011)	(0.003)	(0.008)	(0.008)	(0.003)	(0.009)	(0.008)
			Latent States		Response Matrix R Observed States		
				≤40%	0.78	0.19	0.03
					(0.067)	(0.059)	(0.028)
				40–60%	0.06	0.76	0.18
					(0.021)	(0.029)	(0.025)
				≥60%	0.01	0.03	0.96
					(0.002)	(0.005)	(0.005)

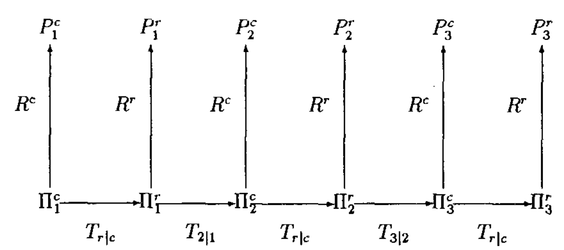

Figure 4. Path Diagram That Illustrates a Two-indicator Latent Markov Model for Three Waves

Here the observed indicators are P_t^c (computed income) and P_t^r (reported income) ($t=1, 2, 3$). The corresponding latent variables for the true state are Π_t^c and Π_t^r ($t=1, 2, 3$). If the transition matrix between Π_t^c and Π_t^r, denoted by $T_{r|c}$, is the unit matrix P_t^c and P_t^r are indicators of the same true poverty position. Temporal changes of the true poverty position are again represented by the transition matrices $T_{2|1}$ (between $t=1$ and $t=2$) and $T_{3|2}$ (between $t=2$ and $t=3$). The connection between the true poverty states and the observed poverty states is established by the response matrices R^c (between Π_t^c and P_t^c) and R^r (between Π_t^r and P_t^r).

Thus, with the two-indicator latent Markov model we have for $T_{r|c}=I$:

$$\text{(2)} \qquad \begin{aligned} &\Pr(P_1^c=i_c, P_1^r=i_r, P_2^c=j_c, P_2^r=j_r, P_3^c=k_c, P_3^r=k_r) \\ &= \sum_l \sum_m \sum_n \pi_l r_{l,i_c}^c r_{l,i_r}^r T_{2|1}(l, m) r_{m,j_c}^c r_{m,j_r}^r T_{3|2}(m, n) r_{n,k_c}^c r_{n,k_r}^r \end{aligned}$$

where $r_{l,i}^c$ are the elements (l, i) of the response matrix R^c and $r_{l,i}^r$ are the elements (l, i) of the response matrix R^r.

However, the single indicator model is not nested within the two-indicator model. Consequently, we cannot test the parameters obtained in the previous section against the estimation results from the two-indicator model. Under the assumption that both indicators are valid for the same latent construct, the use of only one indicator appears to be an inefficient use of the observed information. In order to justify the basic condition $T_{r|c}=I$, we also estimated a model without this restriction. It turned out that this restriction was not significant.[13]

Note also, that the two-indicator model makes assumptions about the occurrence of misclassifications. With respect to the intertemporal independence of measurement errors the same objections apply as in the case of the single indicator model. Furthermore, the model assumes that there is no direct dependence between misclassifications from the two different measurements.

If we compare the measurement errors related to the two income measurements there are some facts that weaken a possible correlation between the measurement errors since they:

(a) are based on different parts of the questionnaire,

(b) were recorded at different points in time, and

(c) were provided by different individuals, if there are different sources of incomes within a household.

The estimation of the net income is also independent from the reported income.

However, in one earner households one may suspect a positive correlation between the two measurement errors. Here it seems plausible that an earner, who underestimates his gross income, will also underestimate his net income which is then the (reported) household income.

[13]In order to stabilize the estimation we used also the restriction $T_{2|1}=T_{3|2}$. This restriction did not appear to be relevant for the single indicator model. Also the results of Table 3, below, suggest that the differences between $T_{2|1}$ and $T_{3|2}$ may be ignored for the two-indicator model.

TABLE 3

ESTIMATES OF THE TWO-INDICATOR LATENT MARKOV MODEL FOR COMPUTED AND REPORTED INCOME IN THE GSOEP

Standard Errors in Parenthesis Under the Estimate

		Latent States					
		Transition 1 to 2			Transition 2 to 3		
	Starting Distrib.	$T_{2\|1}$			$T_{3\|2}$		
Poverty State	π	$\leq 40\%$	40–60%	$\geq 60\%$	$\leq 40\%$	40–60%	$\geq 60\%$
$\leq 40\%$	0.06	0.95	0.05	0.00	0.85	0.12	0.03
	(0.006)	(0.053)	(0.052)	(0.027)	(0.055)	(0.060)	(0.019)
40–60%	0.20	0.01	0.92	0.07	0.01	0.88	0.11
	(0.009)	(0.001)	(0.021)	(0.021)	(0.009)	(0.022)	(0.020)
$\geq 60\%$	0.74	0.00	0.02	0.98	0.00	0.01	0.99
	(0.009)	(0.001)	(0.005)	(0.005)	(0.001)	(0.003)	(0.003)

		Observed States					
		Computed Income			Reported Income		
	Latent States	Response Matrix R^c			Response Matrix R^r		
	$\leq 40\%$	0.62	0.37	0.02	0.61	0.35	0.05
		(0.035)	(0.034)	(0.008)	(0.031)	(0.029)	(0.012)
	40–60%	0.05	0.62	0.33	0.09	0.59	0.32
		(0.008)	(0.016)	(0.017)	(0.009)	(0.013)	(0.014)
	$\geq 60\%$	0.00	0.02	0.98	0.01	0.04	0.95
		(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)

Table 3 displays estimates of the two-indicator model. It reveals a decrease in reliability which uniformly affects both income measures, lowering the reliability of the 40 percent and the 40–60 percent state from about 0.75 to approximately 0.60. As a consequence the probabilities of not moving from poverty rise from a level of about 0.70 to approximately 0.90. Correspondingly, the probability of leaving the ≤ 40 percent state drops from about 0.28 in Table 2 to approximately 0.10. The findings indicate that there are many cases where the poverty profile with respect to measurement 1 does not agree with the profile with respect to measurement 2.

Note however, that the level of poverty, i.e., the starting distribution π, is in close accordance with Table 2. Hence, the level of poverty is quite stable under different models of measurement error.

8. A TWO-INDICATOR MODEL WITH CORRELATED MEASUREMENT ERRORS

In the one-indicator model it is hard to relax the assumption of independent measurement errors. Without imposing restrictions on the model parameters a dependence of misclassifications from previous misclassifications will be not identified. It is obvious that there is a direct trade-off between poverty patterns at the latent level and corresponding patterns of errors at the observed level.

The two-indicator model offers us the possibility of treating serially correlated measurement errors within each indicator. The identification of the model parameters is based here on the independence of the measurement errors between the

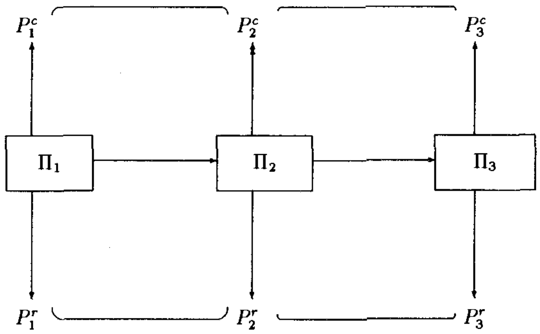

Figure 5. Path Diagram That Illustrates a Two-indicator Latent Markov Model with Correlations Between Measurement Errors

two indicators. The model equation is as follows:

(3) $$\Pr(P_1^c=i_c, P_1^r=i_r, P_2^c=j_c, P_2^r=j_r P_3^c=k_c, P_3^r=k_r) = \sum_l \sum_m \sum_n \pi_l r_{l,i_c}^c r_{l,i_r}^r T_{2|1}(l,m) r_{m,i_c,j_r}^c r_{m,i_r,j_r}^r T_{3|2}(m,n) r_{n,j_c,k_c}^c r_{n,j_r,k_r}^r$$

Contrary to the independence model it is assumed here that the response probabilities in waves 2 and 3 depend on the observed outcome of the indicator of the preceding wave. This model can be represented by the path of Figure 5.

We are interested here only in the effect of this model extention on the estimation of the transition matrices $T_{2|1}$ and $T_{3|2}$.[14] The results are shown in Table 4. It is seen there that there are only marginal differences to the estimation results from the independence model. Apparently this assumption does not affect the results for the transition matrices.[15]

TABLE 4

ESTIMATES OF THE TWO-INDICATOR LATENT MARKOV MODEL WITH CORRELATED ERRORS

		Latent States					
	Starting Distrib.		Transition 1 to 2 $T_{2\|1}$			Transition 2 to 3 $T_{3\|2}$	
Poverty State	π	≤40%	40–60%	≥60%	≤40%	40–60%	≥60%
≤40%	0.06	0.93	0.05	0.02	0.83	0.15	0.02
40–60%	0.20	0.00	0.96	0.04	0.00	0.87	0.13
≥60%	0.74	0.00	0.04	0.96	0.00	0.01	0.99

[14]The model was estimated by Vermunt's program package LEM (Vermunt 1995). In its present form LEM does not provide standard errors for the parameter estimates since it bases entirely on the EM algorithm and does not compute a Hessian matrix.

[15]However, the null hypothesis of independence has to be rejected by the likelihood ratio test: LR = 1083.4 (df = 702) for the independence model compared to LR = 694 (df = 694) for the correlation model.

9. Discussion

Our starting point was the question of how measurement error affects the observed transition between poverty states and to what extent are the two alternative income measurements reliable. The most striking result shown in the previous sections is the sensitivity of the estimated turnover tables with respect to different measurement models. The probability to slip out of the ≤ 40 percent state drops from about 0.50 at the observed level (Table 1) to about 0.28, if the measurements errors are defined separately (Table 2), and falls to about 0.10, if both measurements are analyzed jointly (Tables 3 and 4).

However, the effects on the total change expressed by

$$ch_{\mathrm{TOTAL}} = 1 - \sum_{l} \pi_l^t T_{t+1|t}(l, l) \tag{4}$$

are less dramatic. The reason for the lower sensitivity of the total change measure is the small proportion (about 6 percent) of persons in the ≤ 40 percernt state. The introduction of correlations between subsequent errors leads only to a moderate increase of ch_{TOTAL}.

TABLE 5

Total Percentage of Change (ch_{TOTAL} Across Models

Model	Income Measure	ch_{TOTAL}) $t=1$ to $t=2$	$t=2$ to $t=3$
Simple Markov	Computed	14.7	13.3
(Table 1)	Reported	18.4	17.4
Latent Markov	Computed	7.6	5.1
(Table 2)	Reported	7.0	5.6
2 Indicator latent Markov (Table 3)	Computed + Reported	3.4	4.0
2 Indicator latent Markov + Corr. errors (Table 4)	Computed + Reported	4.2	4.6

With respect to the reliability of the two measurements it can be said that there appears to be no substantial differences. The same conclusion holds for the estimated level and stability of (relative) poverty. Thus an analysis with the self-assessed household income, which was initially intended as a crude proxy variable, appears to be as reliable as an analysis with the carefully edited computed income.

The analysis of Section 8 has shown that the assumption of serially correlated misclassifications is not a relevant topic for the estimation of the latent turnover tables; i.e., the different results for observed and latent turnover tables cannot be attributed to the assumption of independent misclassifications. They have to be attributed to measurement error or misspecification of the Markov model. Up to now the second cause was ignored in the discussion of measurement errors, which is primarily based on a two wave analysis where the problem does not arise; see for example Skinner and Torelli (1993). However, as soon as we use more than two panel waves the second cause has to be considered.

It is important to notice that all models of Table 5 are Markov models. The choice of this model class is closely linked to the interpretation of turnover tables as marginal risks to slip into or out of poverty. Poverty profiles which are derived from multiplying the observed turnover tables will in general underestimate the observed number of profiles with "no change," see Berntsen and Rendtel (1991) for the case of the GSOEP. The measurement error models compensate this lack of fit by assuming a high stability on the latent level, which explains a higher percentage of observed profiles with "no change." As a consequence, observed profiles with changes are attributed to measurement error. These changes are interpreted as "spurious changes." Thus, a good deal of "measurement error" is due to the incapability of the simple Markov model to fit the observed percentage of "no change" profiles.

However, if we proceed to a model in observed variables that gives a better fit of the observed profiles, we would expect only moderate effects of a model extention with measurement error. For example, Berntsen and Rendtel (1991) have shown for the GSOEP that a second-order Markov chain gives a reasonable fit of observed poverty profiles. Such Markov models differ from the ones discussed here by the fact that the transition probabilities depend not only on the poverty state at time t but also on the poverty state at time $t-1$. The measurement models discussed here can be extented to the class of second order Markov chains. At least theoretically this can be achieved by the introduction of a new state space, that consists of two subsequent poverty states. The estimation of such second order Markov models could answer the question of the true role of measurement error compared to the specification error.

Nevertheless, the turnover tables at the latent level seem to be convenient measures of poverty risks that compensate for misclassification and misspecification or, saying it differently, spurious change.

There is an interesting by-product of these results. Some authors, for example Shorrocks (1978), assume a trade-off between the stability of incomes and the level of poverty. Under such a hypothesis one would expect a decrease in the percentage of households in the ≤ 40 percent poverty state, if we switch from the observed poverty states to the true but unobserved states that are more stable. However, this is not true: the percentage of households under the 40 percent poverty line remains almost identical.

A potential caveat in the interpretation of the results are the conceptual differences of the two income measurements. The discordant profiles from the two measurements that effect the further decrease of latent mobility (Table 2 vs. Table 3) may be caused by conceptual differences. However, there are some findings that vote against such an interpretation. First, the separate analysis of each measurement yields very similar results at the observed level. Second, there was no statistical evidence that the measurements are indicators of different latent constructs.

The high degree of correspondence in the results for the level of poverty confirm the empirical results of Benus and Morgan (1975) who compare Gini coefficients for family incomes defined by different lengths of reference period. Table 3 of Benus and Morgan reveals that the length of the reference period has almost no impact on the resulting Gini coefficients. The results presented here extend their findings since the dynamics between the observed poverty states also appear to be quite similar for different reference periods.[16]

References

Abowd, J. M. and Zellner, A., Estimating gross labor force flows, *Journal of Business and Economics Statistics*, 3, 254–83, 1985.

Bane, M. J. and Ellwood, D. T., Slipping into and out of poverty: the dynamics of spells *Journal of Human Resources*, 21, 1–23, 1986.

Benus, J. and Morgan, J., Time Period, Unit of Analysis and Income Concept in the Analysis of Income Distribution, in J. Smith, (ed.), *The Personal Distribution of Income and Wealth*, 209–24, NBER, New York, 1975.

Berntsen, R., *Dynamik in der Einkommensverteilung privater Haushalte*, Campus, Frankfurt/New York, 1992.

Berntsen, R. and Rendtel, U., Zur Stabiblität von Einkommensarmut im Längsschnitt," in U. Rendtel and G. Wagner, (eds.), *Lebenslagen im Wandel, Zur Einkommensydnamik in Deutschland seit 1984*, 457–87, Campus, Frankfurt/New York, 1991.

Bound, J., Brown, Ch., Duncan, G. and Rodgers, W., Measurement Error in Cross-sectional and Longitudinal Labour Market Surveys: Validation Study Evidence, in J. Hartog, G. Ridder and J. Theeuwes, (eds.), *Panel Data and Labour Market Studies*, 1–19, North-Holland, Amsterdam, 1990.

Buhmann, B., Rainwater, L., Schmaus, G. and Smeeding, T. Equivalent Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates across Ten Countries Using the Luxembourg Income Study Database," *The Review of Income and Wealth*, 32, 115–42, 1988.

Bye, B. V. and Schechter, E. S. A latent Markov model approach to the estimation of response error in multiwave panel data, *Journal of the American Statistical Association*, 81, 375–80, 1986.

Headey, B. and Krause, P. Rich and Poor: Stability or Change? West German Income Mobility 1984–93, DIW-Discussion Paper No. 108, Berlin, 1995.

Langeheine, R. and van de Pol, F., A unifying framework for Markov modeling in discrete space and discrete time, *Sociological Methods and Research*, 18, 416–41, 1990.

———, Recent developments in discrete space discrete time Markov modeling, in F. Faulbaum, (ed.), *Softstat '91: Advances in Statistical Software 3*, 199–25, Fischer Verlag, Stuttgart, 1992.

———, Multiple indicator Markov models, in R. Steyer, K. Wender, and K. Widaman (eds.), *Proceedings of the 7th European Meeting of the Psychometric Society in Trier*, 248–52, Fischer Verlag, Stuttgart, 1993.

———, Discrete time mixed Markov latent class models, in A. Dale and R. Davies (eds.), *Analyzing social and political change: A casebook of Methods*, 170–97, Sage, London, 1994.

Pol, F., v.d., Langeheine, R. and de Jong, W., *PANMARK User Manual, Panel Analysis Using Markov Chains*, Netherlands Central Bureau of Statistics, Voorburg, 1991.

Pol, F., v.d., and de Leeuw, J., A latent Markov Model to Correct for Measurement Error, *Sociological Methods and Research*, 15, 118–41, 1986.

Poulsen, C. C., Latent structure analysis with choice modeling applications, Ph.D. dissertation, University of Pennsylvania, 1982.

Rendtel, U., Langeheine, R. and Berntsen, R., Mobilitätsprozesse und Einkommensarmut: Erfragtes und erechnetes Haushaltseinkommen im Vergleich, DIW-Discussion Paper No. 56, Berlin, 1992.

Rendtel, U., and Schwarze, J., Die Entwicklung individueller Arbeitseinkommen von 1984 bis 1989—Eine explorative Analyse von Paneldaten, in U. Rendtel and G. Wagner, (eds.), *Lebenslagen im Wandel, Zur Einkommensdynamik in Deutschland seit 1984*, 63–99, Campus, Frankfurt/New York, 1991.

Rohwer, G., Einkommensmobilität privater Haushalte 1984–1989, in U. Rendtel and G. Wagner, (eds.), *Lebenslagen im Wandel, Zur Einkommensdynamik in Deutschland seit 1984*, 379–408, Campus, Frankfurt/New York, 1991.

Schwarze, J., Der Einfluβ von Respondentenwechseln auf die Varianz von Haushaltseinkommen," Mimeo, Deutsches Institut für Wirtschaftsforschung, 1995.

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

[16]Headley and Krause (1995) obtained a result that also indicates a low influence of the reference period in the GSOEP. They compared the correlation of subsequent equivalent incomes based on a one-year period and on a five-year average of the reported income.

Silverman, B., *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, 1986.
Skinner, C. and Torelli, N., Measurement Error and the Estimation of Gross Flows from Longitudinal Economic Data, *Statistica*, 3, 391–405, 1993.
Slemrod, J., Taxation and Inequality: a Time-Exposure Perspective, in J. Poterba, (ed.), *Tax Policy and the Economy*, 105–27, NBER and MIT Press, Cambridge MA, 1992.
Vermunt, J., Loglinear Event History Analysis: A General Approach with Missing Data, Latent Variables and Unobserved Heterogeneity, Tilburg University Press, Tilburg, 1995.
Wagner, G., Burkhauser, R. and Behringer, F., The English Language Public Use File of the German Socio-Economic Panel, *Journal of Human Resources*, 28, 429–33, 1993.
Wiggins, L. M., Mathematical Models for the Analysis of Multi-Wave Panels, Ph.D. dissertation, Columbia University New York, 1955.
———, *Panel Analysis*, Elsevier, Amsterdam, 1973.