Journal of Official Statistics, Vol. 21, No. 4, 2005, pp. 591–604

On the Bias in Gross Labour Flow Estimates Due to Nonresponse and Misclassification

Li-Chun Zhang[1]

I evaluate and compare the bias due to nonresponse and misclassification in the sample gross labour flow estimates from the Norwegian Labour Force Survey (LFS). These help also to explain the level and net change estimates from the same survey. The main conclusions are the following: (a) the overall labour market stability, i.e., the proportion of people without change in status, should be boosted after adjusting for both nonresponse and misclassification, (b) neither nonresponse nor misclassification affects the net change estimates, and (c) misclassification has very little effect on the level estimation of the characteristics "employed", "unemployed" and "not in the labour force".

Key words: Post-stratification; nonignorable nonresponse model; Markov latent class analysis; sparse classification errors.

1. Introduction

Gross flows are cross-classified counts according to the states of a categorical variable over time. The most frequently studied gross flows in the Labour Force Survey (LFS) are month-to-month, or quarter-to-quarter, cross-classified counts according to the labour force statuses "employed" (E), "unemployed" (U) and "not in the labour force" (N). These can be arranged in a 3×3 contingency table of turnovers EE, EU, EN, . . . , and NN, where EE stands for "E at time point 1 and E and time point 2", and EN stands for "E at time point 1 and N at time point 2", and so on. Not only are such gross labour flow estimates important for social researchers and policy makers, they also help to explain the level and net change estimates from the LFS.

The sample of the Norwegian LFS contains 8 rotation groups. The overlap between two successive quarters consists of 7 rotation groups and about 20,000 persons. Nonresponse at each point in time is about 10 per cent. Estimation of gross labour flows based on panel data must solve several problems, including period-to-period nonresponse, response error (i.e., misclassification of labour force status), inconsistency between panel-based and complete-sample-based point-in-time estimates, differences in sample weights over time, etc. In this article I evaluate and compare the bias in the sample gross flow estimates caused by nonresponse and misclassification. Previous studies on the subject have largely concentrated on one or the other of these two errors.

Both individual (e.g., Stasny 1986, 1988) and household-level nonresponse models (e.g., Clarke and Chambers 1998; Clarke and Tate 2002) have been proposed for the

[1] Statistics Norway, Kongens gt. 6, P.B. 8131 Dep., N-0033 Oslo, Norway. Email: lcz@ssb.no
Acknowledgment: I am grateful to two anonymous referees for comments and suggestions.

© Statistics Sweden

labour force data. At present the Norwegian LFS selects the sample from a population register of families, which can be understood as a particular type of household. Nonresponse occurs for families as well as individuals: about 60% of all the nonrespondents come from single-person families where there is no difference between individual- and family-level modeling; about 20% of the nonrespondents come from families with more than one person where nonresponse is recorded for the entire family. In this study I consider only individual-level models.

Thomsen and Zhang (2001) found that nonresponse leads to considerable (i.e., compared to the sampling error) over-estimation for "employed" at each point in time, whereas it has very little effect on net change estimates based on the panel data. The present investigation suggests that this happens because nonresponse does not affect the off-diagonal cells of the measured turnover table, i.e., people whose classified labour force status is changed from one quarter to the next. To explore the bias for the diagonal cells, i.e., people without change in status, I develop a nonignorable nonresponse model under which nonresponse is approximately independent of net changes, and compare it to the method of dynamic post-stratification (Thomsen and Zhang 2001).

Most studies on gross flow estimates subject to misclassification assume data in the form of contingency tables. Two types of approach can be distinguished. The first one is to estimate the classification probabilities based on reinterview data (e.g., Abowd and Zellner 1985; Poterba and Summers 1986; Chua and Fuller 1987). Singh and Rao (1995) discussed modifications of some of the methods proposed. The second approach is known as Markov latent class analysis (MLCA), introduced for labour market data by Van de Pol and de Leeuw (1986). While it does not presume reinterview data, the MLCA can incorporate such data when they are available (Van de Pol and Langeheine 1997). Biemer and Bushery (2000) evaluated the MLCA approach to the Current Population Survey (CPS). Reinterview data or not, empirical studies suggest that misclassification causes under-estimation of the diagonal cells (i.e., EE, UU and NN) of the turnover table and over-estimation of the off-diagonal cells. Van de Pol and Langeheine (1997), however, argued against the risk of over-adjustment, which they attribute to the commonly made assumption of independent classification errors (ICE). That is, classifications at different time points are conditionally independent given the latent labour force status at these times.

No reinterviews are conducted in the Norwegian LFS. Application of the MLCA produces improbably large adjustments of the observed gross flow estimates. This has motivated us to adopt an alternative approach. The saturated model of classification probabilities which maps the nine latent turnovers to the nine measured ones contains 72 unknown parameters. Allowing only a selected few of these classification probabilities to be identified based on the data, I find that simple sparse classification error models provide adjustments which are more plausible than those provided by the MLCA, with low risks of over-correcting the bias due to misclassification. In addition, it so happens that the fitted sparse classification error mechanisms preserve the marginal totals of the flow table without any further constraining, providing empirical evidence that misclassification has little effect on the estimates of marginal totals.

In the sequels nonresponse is considered in Section 2, and misclassification in Section 3. A summary of the empirical findings is given in Section 4.

2. Nonresponse

2.1. Post-stratification

Let $R_{it} = 1$ indicate that person i responds at time t, and let $R_{it} = 0$ indicate nonresponse, where $t = 1, 2$. Let Y_{it} be the measured labour force status if $r_{it} = 1$, which takes value 1 for E, 2 for U and 3 for N. In the case of $r_{it} = 0$, let Y_{it} be the conceptual measured status had r_{it} been 1. Let $(y_{i1}, y_{i2}) = (j, k)$, for $j, k = 1, 2, 3$, and denote the turnovers EE, EU, EN, . . . , and NN. Let $x_{it} = 1$ if person i is "registered employed" at time t, and $x_{it} = 0$ otherwise. The register employment status is available from the administrative registers, which can be linked to the sample through a unique personal identification number whether or not a person responds in the survey.

Dynamic post-stratification uses register information of both time points. Let $h = 1, \ldots, 4$ denote the post-stratum corresponding to $(x_{i1}, x_{i2}) = (1, 1), (1, 0), (0, 1)$ and (0, 0). Let n_h be the sample size of the hth post-stratum, and let $n = \sum_{h=1}^{4} n_h$. Let m_h be the number of persons in the hth post-stratum who respond at both time points, and let $m = \sum_{h=1}^{4} m_h$. Let p_{hjk} be the sample proportion of flow (j, k) in post-stratum h. The proportion of (j, k) in the gross panel is then given by $p_{jk} = \sum_h (n_h/n) p_{hjk}$. Let $\hat{p}_{hjk}$ be the observed proportion of flow (j, k) in post-stratum h. The observed and post-stratified estimates for p_{jk} are, respectively,

$$\hat{p}_{jk} = \sum_{h=1}^{4} (m_h/m)\hat{p}_{hjk} \quad \text{and} \quad \tilde{p}_{jk} = \sum_{h=1}^{4} (n_h/n)\hat{p}_{hjk} \tag{1}$$

Both sets of estimates have been plotted in Figure 1 for the 12 quarter-to-quarter panels between the first quarter of 1999 and the first quarter of 2002.

Firstly, I notice that the estimated proportions of persons who change status are almost identical by both methods for all the panels I have looked at. It follows that the estimates of measured net changes must be the same by both methods, which corroborates the findings of Thomsen and Zhang (2001) based on the data from 1995 to 1997. Under the assumption of dynamic post-stratification, i.e., nonresponse is independent of labour force status within each post-stratum, the post-stratified net change estimates determine the marginal distribution of the net changes, whereas the observed sample net changes provide the distribution of the net changes conditional on $(r_{i1}, r_{i2}) = (1, 1)$. The fact that they turn out to be almost the same suggests empirically that measured net changes may be independent of nonresponse.

Next, I notice that the post-stratified estimate of EE is considerably lower than the observed estimate of EE, is much higher for NN, and is only slightly higher for UU. The essential effect of dynamic post-stratification is to move some of the flow EE to NN. Should we conclude that nonresponse has very little effect on UU? Or has post-stratification failed to provide sufficient adjustment for likely over-representation of UU among the nonrespondents? The latter may be the case because, while the register information distinguishes between "employed" and "unemployed", it does not distinguish between "unemployed" and "not in the labour force".

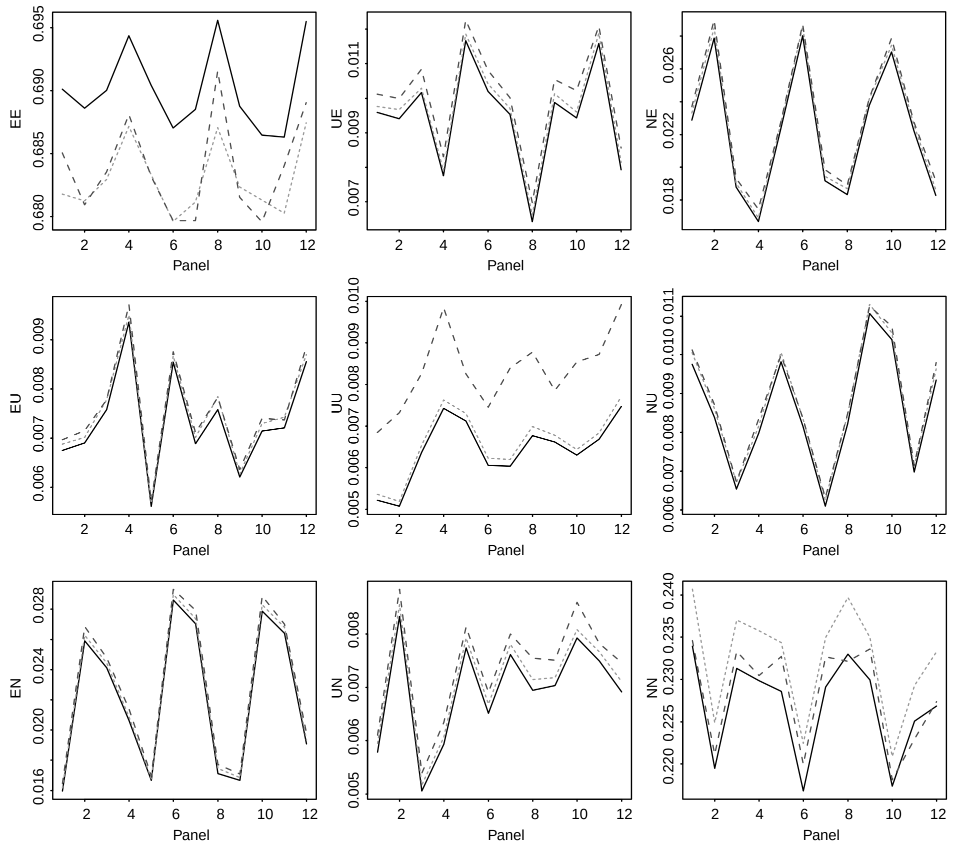

Fig. 1. Quarter-to-quarter labour flow estimates in the Norwegian LFS sample from 1999 to 2002: observed (solid), post-stratified (dotted), and nonignorable model based (dashed)

2.2. Nonignorable nonresponse model

The post-stratification is based on a missing-at-random (Little and Rubin 1987) nonresponse model, where nonresponse is independent of the variable of interest (i.e., gross flow) given the auxiliary variable (i.e., employment register). To explore the possible bias due to this assumption I consider next nonignorable nonresponse modelling, where nonresponse depends on the variable of interest even after conditioning on the auxiliary variables available.

In light of the results above, we would like the nonignorable nonresponse model to satisfy the assumed independence between nonresponse and net changes. Consider the model

$$P[R_{i1} = 1|(y_{i1}, y_{i2}) = (j, k)] = \begin{cases} \alpha_j & \text{if } j = k = 1, 2, 3 \\ \alpha_0 & \text{if } j \neq k \end{cases} \tag{2}$$
$$P[R_{i2} = 1|(y_{i1}, y_{i2}, r_{i1} = r)] = P[R_{i2} = 1|r_{i1} = r] = \beta_r \quad \text{for} \quad r = 1, 0$$

At the first wave, the diagonal flows are allowed for distinct response probabilities, while all the off-diagonal flows share the same response probability. There are not enough degrees of freedom left in the data to allow for R_{i2} to depend on both r_{i1} and $(y_{i1},\ y_{i2})$. In our experience with nonresponse models for panel data, r_{i1} is usually the strongest explanatory variable for R_{i2}. We assume by (2) that R_{i2} is independent of $(y_{i1},\ y_{i2})$ given r_{i1}. Under the model above, all the off-diagonal flows have the same probability of being fully observed. In this way, I expect that the model would lead to similar adjustments of all the observed off-diagonal flows, and that the adjusted net changes would stay close to the observed ones, in which case the model would satisfy the assumed independence between nonresponse and net changes. Finally, for parameter estimation I assume in addition the saturated model with probability p_{jk} for $(Y_{i1},\ Y_{i2})$. Altogether there are 14 parameters on 15 degrees of freedom in the data.

The imputed sample gross flow estimates, i.e., the expected complete panel conditional on the observed data, have been plotted in Figure 1. The estimated off-diagonal flows are very close to the observed and post-stratified ones, such that model (2) approximately captures the assumed independence between nonresponse and net changes. When it comes to the diagonal flows, (i) the estimates for UU are markedly higher than the observed and post-stratified ones, (ii) the adjustments for EE are more or less on the same level as by post-stratification, but appear to be somewhat more volatile over time, and (iii) the adjustments for NN have the same direction as by post-stratification except on one occasion. Despite the indications, it is difficult to conclude at this stage that the post-stratification does under-adjust for UU. We expect further evidence on the size of this adjustment to become available once the ongoing work at Statistics Norway, which is designed to combine the unemployment register with the current employment register, is finished.

3. Misclassification

3.1. Spurious flows due to misclassification

Let Z_{it} be the latent true labour force status of person i at time t. Let $(z_{i1}, z_{i2}) = (a, b)$ denote the latent turnovers, for $a, b = 1, 2, 3$. Let π be the vector obtained by stacking the columns of the latent 3×3 flow table of $(Z_{i1},\ Z_{i2})$. Let ξ be the expectation of the vector similarly derived from the measured flow table of $(Y_{i1},\ Y_{i2})$. Let $b_{jk,ab}$, be the probability of classifying the latent flow $(a,\ b)$ as $(j,\ k)$. Let the 9×9 matrix $\mathbf{B}$ have elements $b_{jk,ab}$, where the rows correspond to the measured flows and the columns correspond to the latent flows, such that

$$\xi = \mathbf{B}\pi \tag{3}$$

In particular, $b_{jk,ab}\pi_{ab}$ is the expected part of the latent flow $(a,\ b)$ which is classified as $(j,\ k)$. It is *genuine* if $(j, k) = (a, b)$, but *spurious* otherwise.

The diagonal cells of the gross labour flow table typically contain 90 to 95 per cent of all the persons in the panel. Previous studies indicate that spurious flows due to misclassification cause more movements away from the diagonal than the reverse, even after taking account of possible higher misclassification rates of the off-diagonal latent flows. It follows that we may be able to account for a large part of the mapping (3) using only the *net* spurious flows away from the diagonal, and ignoring altogether the spurious flows arising from the off-diagonal cells. Moreover, the ICE assumption implies that the probability of misclassification decreases geometrically as the number of errors increases. Studies based on reinterview data identify the probability of correctly classifying labour status E or N as being close to unity. The probability is lower for labour status U, but may still range between 85 and 90 per cent (e.g., Chua and Fuller 1987; Poterba and Summers 1995; Singh and Rao 1995). Thus, among the spurious flows arising from the diagonal cells, we only need to take into account those which involve a single classification error. Both the assumption of net spurious flows and the assumption of single classification errors allow us to reduce the number of parameters in the classification matrix **B**. We shall refer to them as the assumptions of *sparse classification errors* (SCE).

We may explore the SCE assumptions empirically. Take any estimated ICE mechanism, denoted by $\mathbf{A}_{3\times 3}$. Form the classification probability matrix $\mathbf{B} = \mathbf{A} \otimes \mathbf{A}$, where $\otimes$ denotes the direct product. Take any observed flow table to be an estimated ξ. Obtain $\pi = \mathbf{B}^{-1}\xi$ and then the decomposition $b_{jk,ab}\pi_{ab}$ for all (j, k) and (a, b). We find, for each measured flow (j, k), the part which corresponds to (I) the spurious flows away from the diagonal cells that involves only one misclassification, and (II) the genuine flow. Using the Norwegian LFS data, we have examined the ICE mechanisms reported by Chua and Fuller (1987), Poterba and Summers (1995), Singh and Rao (1995), and Biemer and Bushery (2000). Very seldom is the sum of (I) and (II) below 95% of any measured flow. Thus, while they obviously cannot fully explain the classification mechanism, the SCE assumptions may nevertheless be able to account for the main part of the mapping (3) for practical purposes.

3.2. *Marginal constraints*

Denote by θ_{ab} the proportion of latent flow (a, b) where $\sum_{a,b}\theta_{ab} = 1$. Without any restrictions on the θ_{ab}'s, there will be no degree of freedom left in the data for the estimation of any classification errors. While misclassification can cause large relative bias in gross flow estimates, its effect on the estimates for marginal totals at a particular time (i.e., stocks) is often considered to be negligible. In the literature this is referred to as the assumption of unbiased stocks (Singh and Rao 1995), which has both practical and theoretical motivations. An idea is therefore to obtain degrees of freedom for identification of the classification errors by introducing marginal constraints. However, having adjusted for nonresponse and obtained the imputed gross panel, we cannot simply fix the sample marginal proportions as such, because $\hat{\theta}_{ab} = p_{jk}$ for $(a, b) = (j, k)$ where p_{jk} is the proportion of (j, k) in the imputed gross panel, and $\hat{\mathbf{B}} = \mathbf{I}_{9\times 9}$, i.e., the identity matrix, would then always fit the data perfectly.

Instead, I use the fact that the samples do overlap in time, and set up the following marginal constraints akin to the assumption of unbiased stocks. For the panel between time $(t - 1, t)$, I require that (i) the estimates for the latent marginals of time t should agree with the corresponding imputed marginal proportions based on the same panel, and (ii) the estimates for the latent marginals of time $t - 1$ should agree with the corresponding estimates based on the previous panel, i.e., the one between time $(t - 2, t - 1)$. The first constraint is an assumption of unbiased stocks whereas the second one helps us to avoid the trivial situation in which $\mathbf{B} = \mathbf{I}_{9\times 9}$ "out-fits" other genuine misclassification mechanisms. Together, this makes way for 4 degrees of freedom in the data for any model of classification errors.

3.3. Two SCE models

Of the three diagonal cells, UU is of a size comparable to that of the off-diagonal cells. In the first SCE model, denoted by SCE1, I assume perfect classification for all the off-diagonal flows as well as for UU. I assume that the probability of classifying EE as EU is the same as the probability of classifying EE as UE, as both involves a single misclassification of *E* as *U*, and likewise for the classification of EE as EN or NE. Notice that this special kind of symmetry is also implied by the ICE assumption. I set the probability of classifying EE as UU, NN, UN, or NU to be zero, since these all involve misclassification on both occasions. I set up the classification probabilities of the latent NN similarly. In terms of $b_{jk,ab}$, I need to identify

$$(b_{11,11}, b_{12,11}, b_{21,11}, b_{13,11}, b_{31,11}) \quad \text{and} \quad (b_{33,33}, b_{13,33}, b_{31,33}, b_{23,33}, b_{32,33})$$

where $b_{12,11} = b_{21,11}$, and $b_{13,11} = b_{31,11}$, and $b_{11,11} + 2b_{12,11} + 2b_{13,11} = 1$, and $b_{13,33} = b_{31,33}$, and $b_{23,33} = b_{32,33}$, and $b_{33,33} + 2b_{13,33} + 2b_{23,33} = 1$. The rest are 1 if $(j, k) = (a, b)$, and 0 otherwise.

Misclassification under the SCE1 model affects all the measured flows except UU, which is implausible and at variance with the available research findings. Fitting SCE1 to the Norwegian data, I find that the estimated probability of classifying EE as UE or EU is close to zero (Table 2). In the second SCE model, denoted by SCE2, I set these two probabilities to be zero, and allow possible misclassification of UU as UE or EU, such that all the measured flows are affected by spurious flows. In terms of $b_{jk,ab}$, I now need to identify

$$(b_{11,11}, b_{13,11}, b_{31,11}) \text{ and } (b_{22,22}, b_{12,22}, b_{21,22}) \text{ and } (b_{33,33}, b_{13,33}, b_{31,33}, b_{23,33}, b_{32,33})$$

where $b_{13,11} = b_{31,11}$, and $b_{11,11} + 2b_{13,11} = 1$, and $b_{12,22} = b_{21,22}$, and $b_{22,22} + 2b_{12,22} = 1$, and $b_{13,33} = b_{31,33}$, and $b_{23,33} = b_{32,33}$, and $b_{33,33} + 2b_{13,33} + 2b_{23,33} = 1$. The rest of the error probabilities are 1 if $(j, k) = (a, b)$, and 0 otherwise.

Based on each imputed gross panel, obtained by the dynamic post-stratification, I estimate the parameters of the SCE models by maximizing the likelihood subjected to the marginal constraints described earlier. The average of the 11 sets of parameter estimates for both SCE models, i.e., between the second quarter of 1999 and the first quarter of 2002, are given in Table 2.

3.4. Various classification probabilities

We have applied the following MLCA model to all the observed three-wave panels between the first quarter of 1999 and the first quarter of 2002. For the latent transition we assume a nonstationary Markov model with 14 parameters. We assume that classification is independent, identically distributed across the three successive quarters with 6 parameters. The average of the 11 sets of estimated ICE mechanisms, denoted be NLFS, is given in Table 1. The classification probabilities of the latent E and N are similar to those reported in the literature whereas the probability of

Table 1. Estimated ICE mechanism NLFS

Observed labour status	Latent labour status		
	E	U	N
E	0.9894	0.1707	0.0230
U	0.0036	0.5245	0.0074
N	0.0070	0.3048	0.9696

correct classification of the latent U is just above 50%, suggesting that almost half of the cases are misclassified. This is much lower than any estimate previously reported.

In Table 2 I compare the classification probabilities of the SCE models to the corresponding ones derived from the ICE mechanism NLFS, and those reported by Singh and Rao (SR, Canadian LFS data year 1989), Chua and Fuller (CF, CPS data year 1982 month 1), Poterba and Summers (PS, CPS data year 1981) and Biemer and Bushery (BB, CPS data year 1996). The estimated probabilities for spurious flows arising from EE and NN are typically smaller under the SCE models than under the ICE mechanisms. The SCE2 probability for spurious flow from UU to UE or EU is larger than the reinterview based estimates (i.e., SR, CF and PS), but smaller than the MLCA estimates (i.e., BB and NLFS). Notice that the MLCA yields a much smaller probability of correct classification of UU than do the other methods.

3.5. Adjustments for misclassification

We applied all the classification mechanisms in Table 2 to adjust the post-stratified estimates. The mean adjustments between the first quarter of 1999 and the first quarter of 2002 are given in Table 3. Details of the SCE2 and NLFS estimates are plotted in Figures 2 and 3. All the adjustments are upwards for the diagonal cells, and downwards for the off-diagonal cells.

The risk of over-adjustment is very high in the case of the NLFS mechanism. Take for instance EN and NE which are the two largest observed off-diagonal flows. The NLFS adjustments here are quite similar to those of PS, according to which these two observed flows are almost entirely spurious on several occasions. Also, as noted earlier, the misclassification rate of latent UU seems improbably high under the NLFS, leading to excessive adjustments for the observed UU.

The SCE models yield much more moderate adjustments with low risks of over-adjustment. The observed proportions of UE, EU, UU, NU and UN are hardly affected at all under the SCE1 model. The adjustments for the remaining four observed flows are almost identical under SCE1 and SCE2. Thus, according to the SCE estimates, misclassification affects mainly the four corner cells of the 3×3 turnover table (i.e., EE, EN, NE and NN). In addition, the SCE estimates preserve the quarterly stocks without any further adjustment — see Figure 3 for the SCE2 estimates, providing empirical evidence for the assumption of unbiased stocks.

Table 2. Some estimates of classification probabilities

Classification	Method						
Latent → Measured	SCE1	SCE2	SR	CF	PS	BB	NLFS
EE → EE	0.9920	0.9916	0.9845	0.9734	0.9553	0.9748	0.9789
EE → EU or UE	0.0003	0	0.0023	0.0032	0.0053	0.0037	0.0036
EE → EN or NE	0.0037	0.0069	0.0055	0.0101	0.0168	0.0088	0.0042
UU → UU	1	0.8912	0.8064	0.7792	0.7184	0.5538	0.2751
UU → UE or EU	0	0.0544	0.0218	0.0311	0.0320	0.0638	0.0895
NN → NN	0.9814	0.9786	0.9686	0.9450	0.9643	0.9604	0.9401
NN → EN or NE	0.0081	0.0070	0.0075	0.0156	0.0114	0.0111	0.0233
NN → UN or NU	0.0012	0.0037	0.0081	0.0116	0.0063	0.0085	0.0072
Data source	Norway	Norway	Canada	USA	USA	USA	Norway
Data time (year)	1999–2002	1999–2002	1989	1982	1981	1996	1999–2002
Reinterview data	No	No	Yes	Yes	Yes	No	No

Table 3. Mean adjustments for misclassification in 12 Norwegian LFS panels

Method	EE	UE	NE	EU	UU	NU	EN	UN	NN
SCE1	0.006	−0.000	−0.005	−0.000	0.000	−0.000	−0.004	−0.000	0.004
SCE2	0.006	−0.001	−0.005	−0.000	0.001	−0.001	−0.004	−0.001	0.005
SR	0.010	−0.001	−0.006	−0.001	0.001	−0.002	−0.006	−0.002	0.006
CF	0.018	−0.002	−0.011	−0.002	0.002	−0.003	−0.011	−0.003	0.012
PS	0.031	−0.003	−0.016	−0.004	0.002	−0.001	−0.015	−0.002	0.007
BB	0.016	−0.001	−0.011	−0.002	0.005	−0.002	−0.010	−0.003	0.007
NLFS	0.011	−0.000	−0.015	−0.002	0.017	−0.002	−0.013	−0.004	0.009

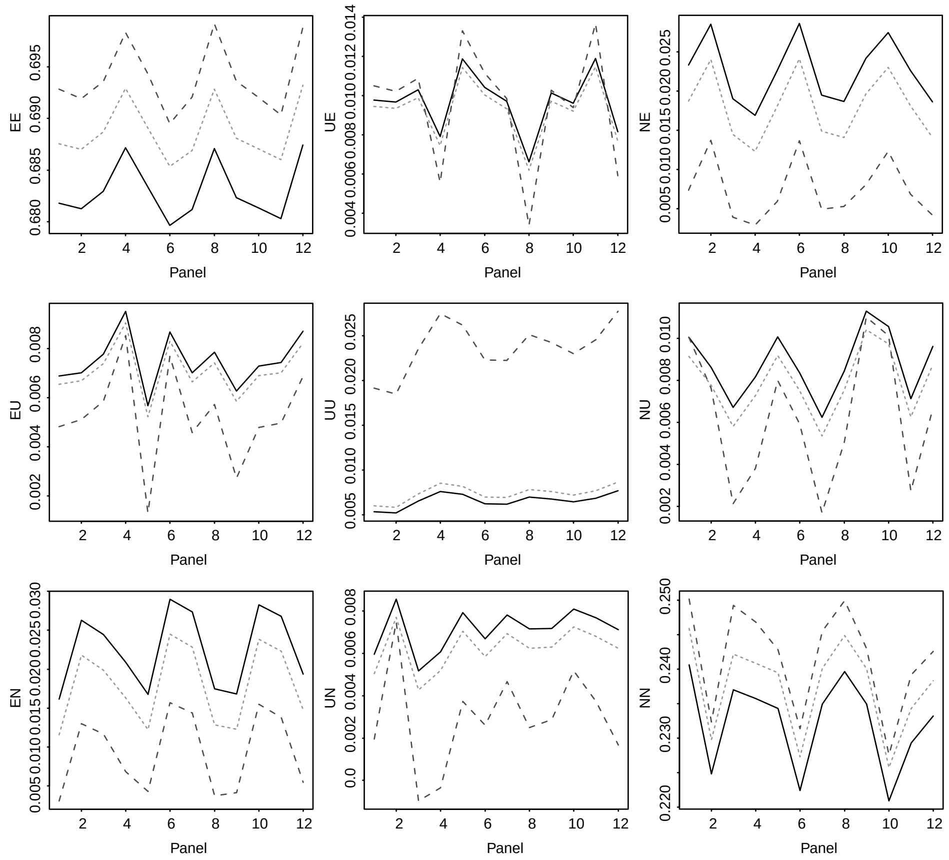

Fig. 2. Quarter-to-quarter labour flow estimates in the Norwegian LFS sample from 1999 to 2002: post-stratified (solid), SCE2 (dotted), and NLFS (dashed)

3.6. Discussion

Formal assessment of the sampling variation of the parameter estimators based on a single panel is complicated here because of the use of the marginal constraints and the adjustments for nonresponse in advance. Nevertheless, one can be quite certain that the estimates are associated with relatively large standard errors as compared to the bias of the nonadjusted flow estimates, due to the lack of degrees of freedom in the data. In deriving the adjusted flow estimates above, I have treated the average error mechanism as known and applied it for the whole period, as in Singh and Rao (1995).

The lack of degrees of freedom also imposes limitations on the SCE models which we can examine. The SCE1 model has clear weaknesses. In the SCE2 model, I assume the probability of classifying EE as EU or UE to be zero based on the results of fitting the SCE1 model, and allow UU to be classified as UE or EU. The findings, however, are basically the same under both models. We actually rejected the SCE model where I set the probability of classifying UU as UE or EU to be zero and allowed UU to be classified as UN or NU instead, because it did not fit the data as well as the SCE2 model. Of course, I

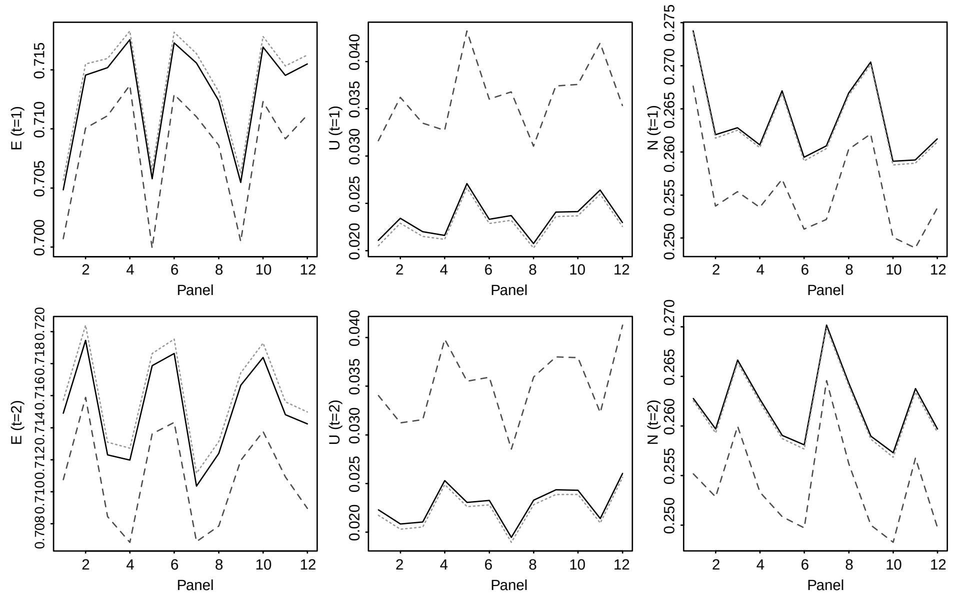

Fig. 3. Sample estimates for the marginal proportions of labour force characteristics in the Norwegian LFS from 1999 to 2002: post-stratified (solid), SCE2 (dotted), and NLFS (dashed)

need to treat standard goodness-of-fit statistics with caution when dealing with nonsampling errors, especially when there is no degree of freedom left in the data under the SCE models above. In the future, we plan to look at more flexible SCE models by alternative ways of exploiting the sample overlap in time.

4. Summary

Both nonresponse and misclassification cause bias in the gross labour flow estimates. Below is a summary of the empirical findings of the present study of the Norwegian LFS.

Nonresponse has very little effect on the proportions of people with quarter-to-quarter changes in the labour force status. This explains the previous finding that nonresponse does not affect the net change estimates for "employed" based on panel data. For the people without change in status, the adjustment of gross flow estimates for nonresponse should be downwards for EE, and upwards for UU and NN. For the level estimates at any particular point in time, therefore, the adjustment for nonresponse should be downwards for E and upwards for U and N.

Misclassification has turned out to be more difficult to handle. The MLCA approach produced improbably large adjustments for the Norwegian data. A reason for this may be that the Markov assumption for latent transitions is too restrictive. The SCE models are *ad hoc* in the sense that they have only been designed for the particular type of data which we are considering here. Despite the fact that they probably cannot remove all the bias due to misclassification, the adjustments by the SCE models are plausible in several respects, from which the following conclusions emerge: (i) adjustments for misclassification boost the labour market stability, i.e., proportions of people without change in status; (ii) there seems to be empirical support for the assumption of unbiased stocks, from which it follows that misclassification has no effect on the net change estimates.

Finally, the combined effects of nonresponse and misclassification are as follows.

- The observed labour market mobilities are biased upwards. The observed flows UU and NN are biased downwards. Both nonresponse and misclassification cause large bias in the observed flow EE, but points in opposite directions and may be mutually compensating.
- Neither nonresponse nor misclassification affects the net change estimates.
- The observed level estimates are biased upwards for E and downwards for U and N.

5. References

Abowd, J.M. and Zellner, A. (1985). Estimating Gross Labor Force Flows. Journal of Business and Economic Statistics, 3, 254–283.

Biemer, P.P. and Bushery, J.M. (2000). On the Validity of Markov Latent Class Analysis for Estimating Classification Error in Labor Force Data. Survey Methodology, 26, 139–152.

Chua, T.C. and Fuller, W.A. (1987). A Model for Multinomial Response Error Applied to Labor Flows. Journal of the American Statistical Association, 82, 46–51.

Clarke, P.S. and Chambers, R.L. (1998). Estimating Labour Force Gross Flows from Surveys Subject to Household-level Nonignorable Nonresponse. Survey Methodology, 24, 123–129.

Clarke, P.S. and Tate, P.F. (2002). An Application of Non-ignorable Non-response Models for Gross Flows Estimation in the British Labour Force Survey. Australian and New Zealand Journal of Statistics, 44, 413–425.

Little, R.J.A. and Rubin, D.B. (1987). Statistical Analysis with Missing Data. New York: Wiley.

Poterba, J. and Summers, L. (1986). Reporting Errors and Labor Market Dynamics. Econometrics, 54, 1319–1338.

Poterba, J. and Summers, L. (1995). Unemployment Benefits and Labor Market Transitions: A Multinomial Logit Model with Errors in Classification. The Review of Economics and Statistics, 77, 207–216.

Singh, A.C. and Rao, J.N.K. (1995). On the Adjustment of Gross Flow Estimates for Classification Error with Application to Data from the Canadian Labour Force Survey. Journal of the American Statistical Association, 90, 478–488.

Stasny, E.A. (1986). Estimating Gross Flows Using Panel Data with Nonresponse: An Example from the Canadian Labour Force Survey. Journal of the American Statistical Association, 81, 42–47.

Stasny, E.A. (1988). Modeling Nonignorable Nonresponse in Categorical Panel Data with an Example in Estimating Gross Labor-force Flows. Journal of Business and Economic Statistics, 6, 207–219.

Thomsen, I. and Zhang, L.-C. (2001). The Effects of Using Administrative Registers in Economic Short Term Statistics: The Norwegian Labour Force Survey as a Case Study. Journal of Official Statistics, 17, 285–294.

Van de Pol, F. and de Leeuw, J. (1986). A Latent Markov Model to Correct for Measurement Error. Sociological Methods and Research, 15, 118–141.

Van de Pol, F. and Langeheine, R. (1997). Separating Change and Measurement Error in Panel Surveys with an Application to Labor Market Data. Measurement Error and Process Quality, L. Lyberg et al. (eds). New York: John Wiley and Sons, 671–687.

Received February 2003
Revised October 2005