Statistics Canada Statistique Canada

Methodology Branch

Social Survey Methods Division

Direction de la méthodologie

Division des méthodes d'enquêtes sociales

Canada

[redacted]

C 2

WORKING PAPER NO. SSMD-91-013 E

METHODOLOGY BRANCH

CLASSIFICATION ERROR ADJUSTMENTS FOR GROSS FLOW ESTIMATES

SSMD-91-013 E

A.C. Singh[1] and J.N.K. Rao[2]

June 1991

[1] Social Survey Methods Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6
[2] Department of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada K1S 5B6

CLASSIFICATION ERROR ADJUSTMENTS FOR GROSS FLOW ESTIMATES

ABSTRACT

The problem of estimating gross flows from repeated surveys is considered when an individual's response at successive points in time is subject to classification error. A popular method for correcting classification errors is based on the assumption of independent classification errors and it uses interview-reinterview data for estimating error rates. In this paper a generalized model based on ϵ-response contamination is proposed which includes the model of independent classification errors as a special case. Using interview-reinterview data, this model can provide a range for bias adjustments for each flow as ϵ moves away from 1 (the value corresponding to independent classification errors model) to a lower bound ϵ_0 bounded away from zero so that a sensitivity analysis to the assumption of indepdendent classification errors can be made. Some numerical examples based on the Canadian Labour Force Survey are presented. It is seen that biases for some cells are fairly stable as ϵ varies but, for others, they show monotonic upward trends in magnitude as ϵ increases. However, for a wide range of ϵ values (between .5 and 1), the assumption of independent classification errors seems fairly robust. Moreover, the biases for flow differences corresponding to symmetric cells seem to be quite insensitive as ϵ varies. Chi-square tests of symmetry and quasi-symmetry for an adjusted flow table are also presented.

Key Words: Independent classification errors, Epsilon-contamination, Bias stability, Monotonic trend, Bias range.

RÉSUMÉ

On examine le problème de l'estimation des flux bruts d'enquêtes répétées lorsque la réponse d'un individu à des instants successifs dans le temps fait l'objet d'erreurs de classification. Une méthode populaire de correction des erreurs de classification repose sur l'hypothèse d'erreurs de classification indépendantes et elle utilise des donnés d'interview-réinterview pour estimer les taux d'erreur. On propose ici un modèle généralisé basé sur la contamination de réponse ϵ qui reprend le modèle des erreurs de classification indépendantes comme cas spécial. Utilisant des données d'interview-réinterview, ce modèle peut fournir une étendue des corrections de biais pour chaque flux à mesure que ϵ passe de 1 (la valeur correspondant au modèle des erreurs de classification indépendantes) à un minorant ϵ_0 (...), de sorte qu'il est possible de procéder à une analyse de sensibilité de l'hypothèse des erreurs de classification indépendantes. La communication contient des exemples numériques tirés de l'Enquête sur la population active du Canada. On peut voir que les biais de certaines cellules sont assez stables à mesure que ϵ varie, mais que pour d'autres, leur grandeur suit une tendance à la hausse monotone à mesure que ϵ augmente. Toutefois, pour un vaste éventail de valeurs de ϵ (entre 0.5 et 1), l'hypothèse d'erreurs de classification indépendantes semble assez robuste. De plus, les biasis des différences de flux correspondant à des cellules symétriques peuvent être assez insensibles à mesure que ϵ varie. On présente également des tests du khi carré de la symétrie et de la quasi-symétrie d'une table de flux corrigée.

Mots clés: erreurs de classification indépendantes; contamination de réponse ϵ; stabilité de biais; tendance monotone; étendue du biais.

1. INTRODUCTION

Gross flows represent transition counts between a finite number of states for individuals in a population from one point in time to the next. For example, in the monthly Canadian Labour Force Survey (LFS) with a rotating panel design, gross flows provide month to month movements in labour force status. These flows are important for researchers and policy analysts for understanding labour market dynamics. For example, one can answer interesting questions (cf. Veevers and Macredie, 1983) such as

(i) How much of the increase in unemployment is due to persons losing or leaving job and how much is due to persons formerly not in the labour force starting to look for jobs?

(ii) How many unemployed persons become discouraged and leave the labour force?

For question (i), the flow differences EU-UE and NU-UN are required while for (ii), the flow UN is required, where E, U, N denote respectively the three labour force statuses, namely, employed, unemployed, and not in labour force, and the flow EU, for example, denotes the number of individuals making transitions from E at time t-1 to U at time t.

In this paper we consider the impact of response error due to misclassification in reporting on gross flow estimates. A somewhat surprising phenomenon occurs in the context of gross flow data. If the response error gives rise to only response variability but not response bias, then the margins (or stocks) of the two dimensional gross flow table corresponding to two successive points in time remain unbiased. However, the interior cell counts (or gross flows) could be seriously biased. The problem of adjusting gross flows with regard to this type of bias is addressed here. There are several approaches to this problem as described by Fuller (1987, p. 272), namely, regression models for discrete variables with measurement error, latent class models, and right-wrong models. Here we will consider right-wrong approach because the underlying model is very general, requires very few assumptions and is easily applicable to flow data arising from complex surveys. Under this model, it is assumed that at time t every element (or responding unit) truly belongs to one of the states (or categories) and that there exists a $m \times m$ response probability (or transition) matrix $B(t)$ which governs an individual's movement between states at time t in the sense that

$$\underset{\sim}{\pi}^0(t) = B(t)\ \underset{\sim}{\pi}(t) \tag{1.1}$$

where $\underset{\sim}{\pi}(t)$ and $\underset{\sim}{\pi}^0(t)$ are m-vectors defining respectively the true proportions and the expectation of observed proportions. If stock estimates are unbiased, we must have $\pi^0(t) = \underset{\sim}{\pi}(t)$, i.e.,

$$\underset{\sim}{\pi}(t) = B(t)\ \underset{\sim}{\pi}(t). \tag{1.2}$$

For flow data corresponding to two points t-1 and t, we have

$$\text{vec}\ \pi^0(t-1,t) = A(t-1,t)\ \text{vec}\ \pi(t-1,t) \tag{1.3}$$

where $\pi(t-1,t)$ and $\pi^0(t-1,t)$ denote $m \times m$ matrices corresponding to true flow proportions and expectation of observed flow proportions respectively, and vec notation means that the columns are stacked one below the other. The size of the response probability matrix $A(t-1,t)$ in (1.3) is $m^2 \times m^2$. Note that the margins $\underset{\sim}{\pi}^0(t-1)$, $\underset{\sim}{\pi}^0(t)$ are related to $\pi^0(t-1,t)$ as

$$\pi_i^0(t-1) = \sum_j \pi_{ij}^0(t-1,t), \quad \pi_j^0(t) = \sum_i \pi_{ij}^0(t-1,t) \tag{1.4}$$

Thus the classification error adjustment problem reduces to estimating $A(t-1,t)$ under a suitable model based on some auxiliary information and then obtaining adjusted $\pi^0(t-1,t)$, i.e. estimate $\pi(t-1,t)$ as

$$\text{vec } \hat{\pi}(t-1,t) = A(t-1,t)^{-1} \text{ vec } P(t-1,t), \tag{1.5}$$

where $P(t-1,t)$ denotes the matrix of observed flow proportions and $\hat{\pi}(t-1,t)$ is the matrix of adjusted flow proportions. The matrix $A(t-1,t)$ is assumed to be nonsingular. It may be noted that some of the entries $\hat{\pi}_{ij}(t-1,t)$ could occasionally be negative. However, this is unlikely for large samples because for a given $A(t-1,t)$ the consistency of $\hat{\pi}$ follows from that of P and (1.3). Further, it can be shown that $\hat{\pi}_{ij}$'s will always sum to unity.

Tables 1, 2, and 3 about here

The problem of potentially serious biases in gross flow data can be explained by means of an example from LFS. This example will be used throughout this paper to introduce various ideas and methods. Consider the 3×3 flow data for the months of October and November 1989 as shown in Table 1 where E, U, N denote the three labour force statuses. The entries in the table are gross flow counts weighted suitably according to the sampling design. The numbers in the parentheses are simply the flow proportions. The total count of 20,226,000 represents the civilian noninstitutionalized population aged 15 or over in the month of October 89. The stocks (or the margins of Table 1) are customarily assumed to be approximately unbiased. This implies that misclassifications from one status to another are compensated by opposite movements. Notice that the stock proportions for E and N are much higher than that for U. However, as seen from the interview-reinterview data (see Tables 2 and 3), higher rates of misclassification for individuals with true status U could produce the desired level of compensation in order for stocks to be unbiased. On the other hand, the interior cell counts of a gross flow table are believed to behave quite differently, i.e. they could be seriously biased. To see this, note that individuals truly belonging to diagonal cells mostly consist of people who tend to retain their statuses over long periods, and therefore, are unlikely to respond in error on both occasions. This implies that, in the event of an error, individuals on the diagonal are likely to go off the diagonal. However, individuals belonging to off-diagonal cells generally consist of people who tend to change

their statuses frequently and therefore response errors on both occasions could very well happen in addition to those on single occasions. This implies that in the event of an error, some individuals from off-diagonal are not unlikely to move to the diagonal. Now since diagonal counts overwhelm off-diagonals (about 95% versus 5% in the example of Table 1), there will be more movements towards off-diagonal than on-diagonal on account of response error. Based on this heuristic reasoning, it follows that upward biases in the off-diagonal cells could be serious because these cell proportions are relatively very small. There is empirical evidence which strongly supports this observation, see e.g. discussions in Poterba and Summers (1986) based on interview-reinterview data, and Lemaitre (1988) based on unemployment insurance beneficiaries data.

In view of the bias problem described above, Statistics Canada does not publish gross flow data whereas the monthly LFS stock data are published regularly. In United States, the Current Population Survey was used to publish flows from 1949-52 but was suspended in 1953 until 1982 when the publication was resumed for unadjusted flow data in order to meet user demands. Clearly, it is important to find a suitable adjustment procedure for the bias problem which has now been around for over four decades. A conference on Gross Flows in labour force Statistics was held in 1984 at Washington, D.C. and several important research papers ensued, see e.g., Fuller and Chua (1984), Chua and Fuller (1987), Abowd and Zellner (1985), and Poterba and Summers (1986). They all used interview-reinterview data as an auxiliary source of information. In Statistics Canada, the problem was identified by Fellegi (1979) and some developmental work was carried out by Wong (1983), Veevers and Macredie (1983), and Gentleman (1988).

The adjustment methods based on interview-reinterview data as proposed in the literature employ the assumption of Independent Classification Errors (ICE); see Section 3. An important question that naturally arises is how stable the bias adjustments would be due to uncertainty in ICE assumption. Questions relating to stability of bias adjustments and its possible differential patterns for different cells in the flow table were also raised by Fellegi (1982). The main purpose of this paper is to study robustness of ICE based adjustment methods.

We use an ε-response contamination model to propose a parametrization of error mechanisms alternative to ICE which can be used to study the impact of uncertainty of ICE assumption on the stablity of biases. The proposed method requires only interview-reinterview data and provides a range for bias adjustments for each flow as the

model parameter ε varies from ε_0 (a lower bound) to 1. The ICE model is obtained as a special case when ε=1. Biases at $\varepsilon = \varepsilon_0$ and $\varepsilon = 1$ for each flow provide bounds for bias adjustment because the bias can be shown to be a monotonic function of ε under very mild regularity conditions. Empirically it was observed that there is a differential pattern in the stability of flow bias adjustments. However, the flow differences of symmetric cells, rather surprisingly, do not show any such differential pattern. In particular, the application of this method to the LFS example in conjunction with the 1989 interview-reinterview data leads to the following important observations.

(i) The bias ranges for EU and UE are obtained respectively as (5%, 10.4%) and (8.2%, 15.5%) of the observed flows as ε varies from ε_0 to 1. (For the given reinterview data, ε_0 is obtained as .098.) Thus, the biases are not very large but there is some instability. The bias range for the difference EU-UE (the increase in the number of unemployed due to individuals losing or leaving job) is (-4.5%, -4.9%) of the observed difference indicating small and stable bias adjustments.

(ii) The bias ranges for EN and NE are respectively (48.2%, 48.9%) and (65.7%, 66.6%) of the observed flows. Thus the biases are very substantial but the narrow intervals indicate stability of bias adjustments. The bias range for the difference EN-NE (the increase in the number of not in labour force due to individuals who lost or left job) is (3.5%, 3.8%) of the observed difference indicating again small and stable bias adjustments.

(iii) The bias ranges for UN and NU are respectively (25.4%, 61.5%) and (24.8%, 48.1%) of the observed flows. Thus the biases in these flows are quite high as well as quite unstable because of wide intervals. However, the bias range for the difference NU-UN (the increase in the number of unemployed due to individuals who started looking for work) is (18.5%, 20.3%) indicating large but stable bias.

(iv) The differential stability patterns in flow biases as ε varies indicates that no single value of the model parameter ε could be recommended for analysis with flow data. However, if ε is believed to lie in the interval [.5,1] which is quite wide, then the adjustments based on ICE (i.e. ε=1) seem fairly robust. In practice multiple analyses of flow tables such as tests of symmetry and quasi-symmetry, could be performed with several values of ε (say, ε_0, .5 and 1) and each analysis scenario could be further investigated for practical relevance using external considerations.

It should be mentioned that there are a number of data defects other than the classification errors that arise before the 3×3 gross flow data of Table 1 can be formed from the raw data. First, there are problems in matching individuals from the two time points because of missing information at one or both time points. The matching is required for determining which cell in the flow-table the individual belongs. The nonmatching may occur either due to nonresponse, processing error, the individual being outside the population of interest or due to the rotation in or out of the sample. Second, there is difference in flow margins and published monthly stocks (based on the full sample) because flow data is based on only part (at most 5/6th in LFS) of the sample due to rotating panel survey design. Third, there is difference in sampling weights at two time points due to changes in the underlying population. Finally, there is the problem of rotation group bias which might contribute bias in the adjustments for partial nonresponse of rotate ins and outs. This bias problem is not considered in this paper. However, for other data defects, a fairly reasonable and practically convenient procedure for data cleaning is outlined below.

First the problem of differing sampling weights at the two time points can be resolved by using the convention that the weights at the previous time point will be used. The changes in population at the two time points are not likely to have much impact on the general pattern of flow data. Moreover, the stock estimates obtained under this approach would be easier to interpret in relation to the published stocks than the one in which the common population for the two time points is ascertained and then the corresponding flow table is generated. Next turning to the problem of missing data due to nonmatching, one can divide all nonmatches into two groups — those with partial information at one of the two points and those with no partial information. For the former group, a suitable imputation procedure such as hot deck can be used to fill in the missing values under the missing at random assumption. For those with no partial information, a suitable weight adjustment procedure such as raking can be employed after post-stratification by age, sex, and region. The missing at random assumption used for individuals with partial information may not be reasonable. However, excluding nonmatches due to rotate ins/outs, only a very small proportion of cases require imputation and so the impact of this assumption is not expected to be serious. The proportion of nonmatches caused by rotation, on the other hand, would not be small and the missing at random assumption not suitable due to rotation group bias. This problem can be alleviated to some extent by performing a margin adjustment via raking in order that flow margins match the published (or full sample) stock estimates at both time points. This will have the added benefit that the unpleasant situation that would have resulted from discrepancies between flow margins and published stocks would no longer

exist. Alternative processing methods such as the one described in Veevers and Macredie (1983) can also be used. Since we are mainly interested in the classification error problem, we assume here that the flow data is suitably processed and reported in the form of a 3×3 table before adjustments for classification errors are considered.

Section 2 describes the nature of auxiliary information that might be used for the problem under consideration. Methods previously proposed are reviewed in Section 3 and the proposed method is described in Section 4 along with its application to LFS. Section 5 presents an example on data analysis with flow data in which impact of classification error bias on chi-square tests of symmetry and quasi-symmetry is investigated. An interesting finding was that conclusions could be reversed if bias-adjusted flow data were analysed as observed data under the given sampling design. Finally, Section 6 contains summary and concluding remarks.

2. USE OF AUXILIARY INFORMATION

To make any adjustment for classification error bias, we need extra information. For this purpose, a special survey such as reinterview is often conducted to assess the nature of response error. With reinterview data, there are two types that can be used: unreconciled and reconciled portions. Fuller and Chua (1984) use only unreconciled data to estimate response probability matrices $B(t-1)$ and $B(t)$ via modelling while Poterba and Summers (1986) describe a procedure to estimate these matrices from both unreconciled and reconciled portions. Abowd and Zellner (1985) consider another alternative in which only reconciled portion of reinterview data is used. The basic idea underlying the use of reinterview data is to estimate the two transition matrices $B(t-1)$ and $B(t)$. The method proposed in Section 4 does not restrict the use to either the unreconciled or reconciled data. We illustrate, however, a method for producing B matrices from reconciled reinterview data for LFS. Generally, interview-reinterview data is cumulated over several months of past data to obtain reliable estimates of response probabilities. For example, Table 2 shows unweighted counts for the period January-November of 1989.

Each column of interview-reinterview data (Table 2) can be scaled to sum to unity in order to produce a proper response probability matrix. However, under the assumption of unbiased stocks, the matrices $B(t-1)$ and $B(t)$ must satisfy

$$B(t-1)\ \underset{\sim}{\pi}(t-1) = \underset{\sim}{\pi}(t-1),\ B(t)\ \underset{\sim}{\pi}(t) = \underset{\sim}{\pi}(t), \tag{2.1}$$

where $\underset{\sim}{\pi}(t-1)$ is replaced by its estimate $\underset{\sim}{P}(t-1)$, $\underset{\sim}{\pi}(t)$ by $\underset{\sim}{P}(t)$, the vectors $\underset{\sim}{P}(t-1)$, $\underset{\sim}{P}(t)$ are respectively the row and column margins (or observed stocks) of $P(t-1,t)$, the matrix of observed flow proportions . The matrices $B(t-1)$ and $B(t)$ can be estimated from interview-reinterview data by raking Table 2 two times in succession so that both margins match $\underset{\sim}{P}(t-1)$ first and then both margins match $\underset{\sim}{P}(t)$. The first raked table after scale adjustment (for the constraint that each column sums to unity) gives an estimate of $B(t-1)$. Similarly, the second raked table yields an estimate of $B(t)$. These are shown in Table 3.

Tables 4 and 5 about here

It may be noted that smoothing of interview-reinterview data via raking to produce B matrices is a convenient and reasonable option which preserves the two-way associations. The B matrices would be treated as known for analysis of adjusted gross flows. This is a convenient assumption which may be reasonable because past reinterview data is cumulated in estimating B matrices. To check this assumption, an alternative interview-reinterview data based on three years (January 1987-November 1989) will also be considered. Table 4 shows the counts for this interview-reinterview data and Table 5 the corresponding B matrices. It is seen that fluctuations in the entries of B matrices due to the alternative reinterview data are fairly small.

3. METHODS BASED ON ICE

There are three main methods of bias adjustment proposed respectively by Fuller and Chua (1984), Abowd and Zellner (1985), and Poterba and Summers (1986). As mentioned earlier in Section 2, these methods differ with respect to how B matrices are estimated from the interview-reinterview data. However, they all use the key assumption of ICE. This can be defined as

$$\phi_{ij|k\ell}(t-1,t) = \beta_{i|k}(t-1)\ \beta_{j|\ell}(t);\ i,j,k,\ell = E,U,N, \tag{3.1}$$

where $\phi_{ij|k\ell}(t-1,t)$ is in the (i,j)th row and (k,ℓ)th column of $A(t-1,t)$, i.e., for a randomly chosen individual it is the conditional probability of being observed in (i,j) when the true t-1 to t labour force status is (k,ℓ), and $\beta_{i|k}(t-1)$ is the (i,k)th element of $B(t-1)$, i.e. it is the conditional probability of observing i at time t-1 when the true state

is k. The $\beta_{j|\ell}(t)$ is similarly defined. Note that ICE does not imply the independence of responses at two consecutive time points i.e. (3.1) does not imply $\Pi^0_{ij}(t-1,t) = \Pi^0_i(t-1)\ \Pi^0_j(t)$ where Π^0 was defined earlier in (1.3) and (1.4). In fact, the condition (3.1) is equivalent to two simpler conditions, namely, first the usual condition of independence of observed states at times t and t-1 given the true states, and second, given the true state at time t-1 (or t), the probability of observing a state does not depend on the true state at other time t (or t-1). Thus, ICE can also be defined by

$$\phi_{ij|k\ell}(t-1,t) = \beta_{i|k\ell}(t-1)\ \beta_{j|k\ell}(t), \tag{3.2a}$$

$$\beta_{i|k\ell}(t-1) = \beta_{i|k}(t-1),\ \beta_{j|k\ell}(t) = \beta_{j|\ell}(t), \tag{3.2b}$$

where β's as before denote conditional probabilities for a single point in time. It now follows that under ICE, the response probability matrix $A(t-1,t)$ can be obtained from B matrices as

$$A(t-1,t) = B(t) \otimes B(t-1) \tag{3.3}$$

where $\otimes$ denotes the Kronecker product.

The ICE assumption defined above by (3.1) or (3.3) can also be explained in terms of correlations between response (or classification) errors. As suggested by Fuller (1991), we define the 3-vector of response errors for the discrete variable of labour force status as

$$\underset{\sim}{e}_t = B(t)^{-1}\ \underset{\sim}{y}^0_t - \underset{\sim}{y}_t, \tag{3.4}$$

where $\underset{\sim}{y}^0_t$, $\underset{\sim}{y}_t$ denote respectively the observed and true response vectors at time t. The vector $\underset{\sim}{y}_t$, for example, has three indicator variables denoting respectively the occurrence or nonoccurrence of E, U, and N. Notice that the response error is not defined in the usual way as $\underset{\sim}{y}^0_t - \underset{\sim}{y}_t$; instead, $\underset{\sim}{y}^0_t$ is first transformed to $B^{-1}(t)\ \underset{\sim}{y}^0_t$. The main reason for the transformation is that $\underset{\sim}{e}_t$ has mean $\underset{\sim}{0}$ and is uncorrelated with the true value $\underset{\sim}{y}_t$. This property simplifies the treatment of response errors for multivariate discrete data, see Fuller (1987, p. 278). The covariance matrix between error vectors $\underset{\sim}{e}_{t-1}$ and $\underset{\sim}{e}_t$ can be obtained as

$$\mathrm{Cov}(\underset{\sim}{e}_{t-1}, \underset{\sim}{e}_t) = B(t-1)^{-1}\ (\Pi^0(t-1,t) - B(t-1)\ \Pi(t-1,t)B'(t))B'(t)^{-1} \tag{3.5}$$

It is easily seen from (1.3) and (3.3) that $\mathrm{Cov}(\underset{\sim}{e}_{t-1},\underset{\sim}{e}_t)$ is a null matrix under ICE as expected.

Once the A matrix has been specified, the adjusted flows ($\hat{\Pi}$) are obtained from the general result (1.5) for right-wrong models. In particular, under ICE we get

$$\Pi^0(t-1,t) = B(t-1)\ \Pi(t-1,t)\ B'(t),\ \hat{\Pi}(t-1,t) = B^{-1}(t-1)\ P(t-1,t)\ B'(t)^{-1} \quad (3.6)$$

Tables 6 and 7 about here

We will illustrate this for the LFS example considered earlier. The B matrices obtained in Section 2 will be used here although the methods originally proposed use somewhat different B matrices. Tables 6 and 7 (columns labelled ICE) show adjusted flows ($\hat{\Pi}$) respectively for the two types of interview-reinterview data based on two different periods of cumulation. Notice that the flows for cells EN, NE, NU and UN show very high magnitudes for positive biases as their relative adjustments are over 50%. The relative adjustment (RA) is simply defined as

$$\text{Relative Adjustment} = (\text{Adjusted-Observed})/\text{Observed}. \quad (3.7)$$

Therefore, bias for a cell will be estimated by negative of the relative adjustment multiplied by the observed proportion. For flows EU and UE, biases are positive but not very high (RA is 10% to 15%). Among diagonals, only the cell UU seems to have moderately high negative bias (RA is 20% or so) whereas the cells EE and NN show very low negative biases (around 1 to 3% for RA). Tables 6 and 7 also indicate that the impact of error in estimating B matrices on adjusted flows is not likely to be serious because the two tables show very similar bias adjustments. It, therefore, follows that treating B matrices (or the A matrix) as known from past data may not be unreasonable. This would then allow, among other conveniences, computation of covariance matrix of adjusted flows, vec $\hat{\Pi}(t-1,t)$, quite simply from that of vec $P(t-1,t)$ using the relation (1.5).

It is seen that the above approach based on ICE does provide bias corrections in the directions as expected on heuristic grounds outlined in the Introduction. The differential effects of bias for different cells are also observed. The stability (or reliability) of biases with regard to sampling variability can, of course, be ascertained for a given $A(t-1,t)$ from the covariance matrix of observed flows $P(t-1,t)$ under a given sampling design. However, their stability with respect to uncertainty in the assumption of ICE is unknown. We propose to investigate this aspect under a particular class of alternatives to ICE which induces correlations between response errors as defined by (3.4). The proposed model provides a bias range for each cell of the flow table so that a sensitivity analysis to the assumption of ICE can be made, as illustrated in the next section for the LFS data.

4. GENERALIZED ICE MODEL

We now consider a generalized ICE model which enables us to study the sensitivity of bias corrections to the global ICE assumption. We assume that a randomly chosen individual is either error-free at time points t-1 and t with probability $1-\varepsilon$ or error-prone with probability ε, $0 \leq \varepsilon < 1$. We also assume that ICE hold for an error-prone individual. The response probability matrix for time points t-1 and t is, therefore, given by

$$A(t-1,t) = (1-\varepsilon)\ I_9 + \varepsilon\ B_\varepsilon(t) \otimes B_\varepsilon(t-1), \tag{4.1}$$

where $B_\varepsilon(t)$ is the response probability matrix for time point t for an error-pone individual, and I_m denotes a mxm identity matrix. Note that ICE holds trivially for an error-free individual since $I_9 = I_3 \times I_3$. We thus have a generalized independent classification errors (GICE) model for gross flows which includes ICE as a special case when $\varepsilon=1$. Similar ε-contamination models have been used in other contexts, for example, in robustness studies (Huber, 1981) and in Bayesian inference (see Perrichi and Walley, 1991 for a review).

The marginal response probability matrices $B(t-1)$ and $B(t)$ are obtained from (4.1) as

$$B(t-1) = (1-\varepsilon)\ I_3 + \varepsilon\ B_\varepsilon(t-1) \tag{4.2a}$$

$$B(t) = (1-\varepsilon)\ I_3 + \varepsilon\ B_\varepsilon(t) \tag{4.2b}$$

To see this, note that the elements of $B(t-1)$ and $B(t)$ are defined by $A(t-1,t)$ and $\Pi(t-1,t)$ as

$$\beta_{i|k}(t-1) = \textstyle\sum_\ell \sum_j \phi_{ij|k\ell}(t-1,t)\ \Pi_{k\ell}(t-1,t)\ /\ \Pi_k(t-1) \tag{4.3a}$$

$$\beta_{j|\ell}(t) = \textstyle\sum_k \sum_i \phi_{ij|k\ell}(t-1,t)\ \Pi_{k\ell}(t-1,t)\ /\ \Pi_\ell(t). \tag{4.3b}$$

It follows from (1.1) and (4.2b) that under the assumption of unbiased stocks as defined by (1.2), we have

$$\begin{aligned} \underset{\sim}{\Pi}^0(t) &= (1-\varepsilon)\ \underset{\sim}{\Pi}(t) + \varepsilon\ B_\varepsilon(t)\ \underset{\sim}{\Pi}(t) \\ &= \underset{\sim}{\Pi}(t) \end{aligned} \tag{4.4a}$$

which implies that

$$\underset{\sim}{\Pi}(t) = B_\varepsilon(t)\ \underset{\sim}{\Pi}(t) \tag{4.4b}$$

We can write $\underset{\sim}{\Pi}^0(t)$ alternatively as

$$\underset{\sim}{\Pi}^0(t) = (1-\varepsilon)\underset{\sim}{\Pi}^0_{1-\varepsilon}(t) + \varepsilon\underset{\sim}{\Pi}^0_{\varepsilon}(t) = \underset{\sim}{\Pi}(t), \tag{4.5}$$

where $\underset{\sim}{\Pi}^0_{1-\varepsilon}(t)$ and $\Pi_\varepsilon^0(t)$ are the expected proportions for error-free and error-prone individuals respectively. Note that (4.4) and (4.5) do not necessarily imply that $\underset{\sim}{\Pi}^0_{1-\varepsilon}(t) = \Pi^0_\varepsilon(t) = \underset{\sim}{\Pi}(t)$.

The B_ε-matrices are estimated, for a given ε, from (4.2a) and (4.2b) using the estimates $\hat{B}(t-1)$ and $\hat{B}\ (t)$ obtained from interview-reinterview data:

$$\hat{B}_\varepsilon(t-1) = \varepsilon^{-1}\ [\hat{B}(t-1) - (1-\varepsilon)I_3] \tag{4.6a}$$

and

$$\hat{B}_\varepsilon(t) = \varepsilon^{-1}\ [\hat{B}(t) - (1-\varepsilon)I_3] \tag{4.6b}$$

It can be seen that in order for the estimated entries of B_ε to be nonnegative, we must have

$$\varepsilon \geq \max\ \{1-\min_{i,j}\ (\hat{\beta}_{i|i}(t-1),\ \hat{\beta}_{j|j}(t)),\ \max_{i\neq k\ j\neq \ell}\ (\hat{\beta}_{i|k}(t-1),\ \hat{\beta}_{j|\ell}(t))\}$$

$$= \varepsilon_0\ (\text{say}). \tag{4.7}$$

The adjusted proportions $\hat{\Pi}(t-1,\ t)$, for a given ε, are obtained from (4.1) and (1.5) by substituting the estimated matrices $\hat{B}_\varepsilon(t-1)$ and $\hat{B}_\varepsilon(t)$ in (4.1). Thus the sensitivity of bias corrections to ICE assumption can be studied by varying ε in the range $[\varepsilon_0,\ 1]$.

The assumption of ICE is clearly not satisfied by GICE for ε below 1 because (3.2a) is violated. The condition (3.2b) is, however, satisfied by GICE. The GICE model gives rise to dependent classification errors analogous to intra-class correlations induced by mixture models. The response error covariances can be computed from (3.5) by substituting $\hat{\Pi}(t-1,t)$ obtained under GICE for ech $\varepsilon \geq \varepsilon_0$. To get correlations, standard deviations of elements of $\underset{\sim}{e}_{t-1}$ and $\underset{\sim}{e}_t$ are of course needed which can be calculated from diagonal elements of $\text{Cov}(\underset{\sim}{e}_{t-1})$ and $\text{Cov}(\underset{\sim}{e}_t)$ where

$$\mathrm{Cov}(\underset{\sim}{e}_t) = B^{-1}(t)\ \mathrm{Diag}\ \{\underset{\sim}{\Pi}^0(t)\}\ B'(t)^{-1} - \mathrm{Diag}\ \{\underset{\sim}{\Pi}(t)\}, \tag{4.8}$$

and $\mathrm{Cov}(\underset{\sim}{e}_{t-1})$ is similarly defined.

Figures 1(a) and (b) about here

The 3×3 matrix of response error correlations defined by (3.5) and (4.8) was estimated from the LFS in conjunction with the 1989 reinterview data. The diagonal elements of the resulting correlation matrix were found to be positive and off-diagonal ones to be negative. (Their magnitudes are shown in Figures 1(a) and (b)). Note that it is reasonable to expect different signs for correlations because response errors correspond to same categories for the diagonal and different categories for the off-diagonal. An interesting finding is that these correlations are monotonic decreasing to zero in magnitude as ε increases to 1. This is an important phenomenon which can be used to explain the behaviour of flow biases obtained under GICE as ε increases. It is seen that flow biases (see the discussion below) also have monotonic patterns, but in the opposite direction, as ε increases. The monotonicity of correlations can be proved under mild regularity conditions about B and Π matrices. It may be noted that sensitity analyses of biases as ε varies would also shed light on their behaviour as response error correlations change because these correlations are monotonic functions of ε.

To interpret the impact of dependent classification errors induced by GICE, it might be useful to consider the following inequalities in addition to the response error correlations presented above. Let indices i,j,k,ℓ be defined as before in (3.1). It can be shown that under GICE, for $\varepsilon \geq \varepsilon_0$

$$\begin{aligned}
&P\{\text{no error at } t \mid \text{no error at } t-1,k,\ell\} \geq P\{\text{no error at } t \mid k,\ell\},\\
&P\{\text{error at } t \mid \text{error at } t-1,k,\ell\} \geq P\{\text{error at } t \mid k,\ell\},\\
&P\{\text{error at } t \mid \text{no error at } t-1,k,\ell\} \leq P\{\text{error at } t \mid k,\ell\},\\
&P\{\text{no error at } t \mid \text{error at } t-1,k,\ell\} \leq P\{\text{no error at } t \mid k,\ell\},
\end{aligned} \tag{4.9}$$

with equality holding if and only if $\varepsilon=1$. To see this, note that the LHS is $\phi_{ij|k\ell}(t-1,t)/\beta_{i|k}(t-1)$, the RHS is $\beta_{j|\ell}(t)$ and then the results follow after some algebra. The above inequalities imply that biases should be less in magnitude under GICE than those under ICE because there is more chance for correct response under GICE and also more chance for

exchange between cells i.e. for compensatory moves. Moreover, the absolute differences between the two sides in (4.9) can be easily shown to be monotonic nonincreasing in ε by noting that the derivative with respect to ε is nonpositive for all $\varepsilon \geq \varepsilon_0$. This provides another explanation of why the biases under GICE attain a maximum in magnitudes at $\varepsilon=1$ and decrease to a minimum as ε decreases to the lower bound ε_0.

Figures 2(a) and (b) and Figure 3 about here

Under very mild regularity conditions (which are expected to be met in practice) about B matrices and the nature of the Π matrix, it can be shown that the flow biases are monotonic functions of ε. Thus, bias ranges can be obtained by computing biases at $\varepsilon = \varepsilon_0$ and $\varepsilon = 1$. The true value of ε is in general unknown and difficult to estimate. However, the bias range can be easily computed using only interview-reinterview data. For the LFS example, Figures 2(a) and (b) show the variation of biases for different flows as ε increases from ε_0(=.098) to 1. In the figures, absolute relative adjustments (ARA) are plotted against ε. Note that the three cells UU,UN and NU are especially sensitive to ε whereas the other cells are quite insensitive. This behaviour may also be explained in terms of response error correlations (given in Figures 1(a) and 1(b)) in the sense that the absolute correlation remains small (less than 0.10) for all ε for the insensitive cells (except for the cell NN).

The behaviour of flow differences for symmetric cells (i.e. EN-NE, EU-UE, NU-UN) as observed from Figure 3 shows a rather striking and interesting result. Their biases are almost constant as ε varies. Thus, under GICE, different flows or their differences have different types of stability patterns. These patterns should be investigated in practice for each flow table before drawing any conclusions regarding bias magnitudes.

Table 8 about here

The bias ranges for flow proportions and their differences for the LFS example using 1989 interview-reinterview data can also be seen from Table 6 by comparing columns of adjusted proportions labelled GICE(ε=.098) and GICE(ε = 1). The corresponding ranges for flow counts and flow differences are shown in Table 8.

The bias ranges when the three year interview-reinterview data is used, are shown in Table 7. Comparing Tables 6 and 7, it is seen that the impact on biases due to variability in interview-reinterview data is small. Thus, B matrices (and hence A) can be treated as fixed in finding covariance matrix of $\hat{\Pi}$ for a given ε. Note that the covariance of $\hat{\Pi}$ depends on ε.

5. CHI-SQUARED TESTS OF SYMMETRY AND QUASI-SYMMETRY

We now turn to analysis of adjusted gross flow tables, in particular chi-squared or likelihood ratio tests of symmetry and the weaker hypothesis of quasi-symmetry. Such tests are well-known for multinomial cell counts (e.g. Bishop, Fienberg and Holland, 1975, Chapter 8), but adjusted gross flow tables do not satisfy the multinomial assumption.

The chi-squared tests will be developed in the context of gross flows for labour force status. The hypothesis of symmetry on the true gross flows is given by H_S: $\Pi_{ij}(t-1,t) = \Pi_{ji}(t-1,t)$ for all $i \neq j$. It can also be expressed as H_S: $u_{1(i)} = u_{2(i)}$ and $u_{12(ij)} = u_{12(ji)}$ for all $i \neq j$ in the saturated loglinear representation of $\Pi_{ij}(t-1,t)$:

$$\ln\Pi_{ij}(t-1,t) = \tilde{u} + u_{1(i)} + u_{2(j)} + u_{12(ij)}, \tag{5.1}$$

where the parameters $u_{1(i)}$, $u_{2(j)}$ and $u_{12(ij)}$ are subject to the constraints $\sum_i u_{1(i)} = 0$ $\sum_i u_{12(ij)} = 0$ for all j, $\sum_j u_{12(ij)} = 0$ for all i, and $\tilde{u}$ is a normalizing factor to ensure that $\sum_i\sum_j \Pi_{ij}(t-1,t) = 1$ (see e.g. Bishop, Fienberg and Holland. 1975, p. 24). Under H_S, the reduced model may be written in matrix notation as

$$\ln\text{vec}\Pi(t-1,t) = u\underset{\sim}{1} + X_S\theta_S, \tag{5.2}$$

where ln vec $\Pi(t-1,t)$ is the vector of logarithms of elements of vec $\Pi(t-1,t)$, $\underset{\sim}{1}$ is a 9×1 vector of 1's, $\theta_S = (u_{1(1)}, u_{1(2)}, u_{12(11)}, u_{12(12)}, u_{12(22)})$ and

$$X_S = \begin{bmatrix} 2 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ -1 & 0 & 0 & -1 & -1 \\ 0 & -1 & -1 & -1 & 0 \\ -1 & 0 & 0 & -1 & -1 \\ -2 & -2 & 1 & 2 & 1 \end{bmatrix} \tag{5.3}$$

A weaker hypothesis than symmetry is the hypothesis of quasi-symmetry given by H_{QS}: $u_{12(ij)} = u_{12(ji)}$ (see e.g., Bishop, Fienberg and Holland, 1975, p. 286). Under H_{QS}, the reduced model may be written as

$$\ln\text{vec}\Pi(t-1,t) = \tilde{u}1 + X_{QS}\theta_{QS}, \tag{5.4}$$

where $\theta_{QS} = (u_{1(1)}, u_{1(2)}, u_{2(1)}, u_{2(2)}, u_{12(11)}, u_{12(12)}, u_{12(22)})'$ and

$$X_{QS} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & 1 & 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ -1 & -1 & 0 & 1 & 0 & -1 & -1 \\ 1 & 0 & -1 & -1 & -1 & -1 & 0 \\ 0 & 1 & -1 & -1 & 0 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 2 & 1 \end{bmatrix} \tag{5.5}$$

In order to develop valid chi-squared tests of H_S and H_{QS}, we need to express models (5.2) and (5.4) in terms of $\pi^0(t-1,t)$ using (1.3). The models do not retain the standard log-linear form since $\ln \mathrm{vec}\pi(t-1,t) = \ln[A(t-1,t)^{-1}\mathrm{vec}\pi^0(t-1,t)] = g(\mathrm{vec}\pi^0)$, say, and hence the standard methods for loglinear models are not applicable. Scott, Rao and Thomas (1990) considered models of the form $g(\mathrm{vec}\pi^0) = \tilde{u}\underset{\sim}{1} + X_1\theta_1$ for general $g(\mathrm{vec}\pi^0)$, and proposed an iterative scheme involving weighted least squares to calculate the maximum likelihood estimates (m.l.e.) of $\tilde{u}$, θ_1 and hence of π^0, assuming that the observed proportions, P, are obtained from a simple random sample drawn with replacement (i.e. multinomial sampling with probabilities π). Rao and Thomas (1990) give details of the iterative scheme for the special case of $g(\mathrm{vec}\pi^0) = \ln(A^{-1}\mathrm{vec}\pi^0)$.

Since no single ε can be selected for the flow data under GICE model, we will do multiple analyses using several values of ε, in particular, $\varepsilon = \varepsilon_0$, 0.5 and 1. The matrix $A(t)$ for the three values of ε under GICE, is obtained from (4.1) using interview -reinterview data. We treat the resulting $A(t)$ as fixed, and obtain the m.l.e. of $\pi^0(t-1,t)$ under H_S and H_{QS} using the iterative scheme, assuming that the observed gross flows, $P_{ij}(t-1,t)$, are obtained from a multinomial sample. Denote the m.l.e. of $\pi^0_{ij}t-1,t)$ by $\hat{\pi}^0_{ij}(St)$ and $\hat{\pi}^0_{ij}(QSt)$ under H_S and H_{QS} respectively. Then, an asymptotically correct test of H_S is obtained by treating either the chi-squared statistic

$$X_S^2 = n \sum\sum [P_{ij}(t-1,t) - \hat{\pi}^0_{ij}(St)]^2 / \hat{\pi}^0_{ij}(St) \tag{5.6}$$

or the log-likelihood ratio statistic

$$G_S^2 = 2n\sum\sum P_{ij}(t-1,t) \ln[P_{ij}(t-1,t)/\hat{\pi}^0_{ij}(St)] \tag{5.7}$$

as a χ^2 variable with 3 degrees of freedom (d.f.). Similarly, an asymptotically correct test of H_{QS} is obtained by treating either

$$X^2_{QS} = n\sum\sum[P_{ij}(t-1,t) - \hat{\pi}^0_{ij}(t-1,t)]^2/\hat{\pi}^0_{ij}(QSt) \tag{5.8}$$

or

$$G^2_{QS} = 2n\sum\sum P_{ij}(t-1,t)\ln[p_{ij}(t-1,t)/\hat{\pi}^0_{ij}(QSt)] \tag{5.9}$$

as a χ^2 variable with 1 d.f.

We can also test the hypothesis of marginal homogeneity, H_{MH}: $\underset{\sim}{\pi}_i(t-1) = \underset{\sim}{\pi}_i(t)$ for all i, given that H_{QS} is true. An asymptotically correct test of $H_{MH}|H_{QS}$ is obtained by treating $G^2_S - G^2_{QS}$ as a χ^2 variable with 2 d.f., noting that $H_S = H_{QS} \cap H_{MH}$ (see e.g., Bishop, Fienberg and Holland, 1975, p. 287).

It is possible to account for the effect of the survey design if the estimated covariance matrix of the $P_{ij}(t-1,t)$ is also known. Corrections to X^2_S and X^2_{QS} that take account of the survey design could then be obtained along the lines of Rao and Thomas (1990).

We will apply the above X^2 tests to the LFS example given in the introduction. We will treat, for convenience, the data of the weighted flow proportions of Table 1 as a multinomial sample with sample size n = 78,007, the matched common respondents in the months of October and November 1989. Later, we will consider implications of adjusting the X^2 tests by the average design effect. The matrix $A(t-1,t)$ is specified from Jan-Nov '89 interview-reinterview data. The m.l.e. under the hypothesis of quasi-symmetry, H_{QS}, is obtained for different values of ε as

For ε = .098,

$$\text{vec } \hat{\pi}^0_{QS}(t-1,t) = (.59905, .00845, .01008, .01203, .02759, .00790, .01256, .00798, .31436)',$$

For ε = .5,

$$\text{vec } \hat{\pi}^0_{QS}(t-1,t) = (.59906, .00854, .00988, .01194, .02762, .00782, .01274, .00803, .31437)',$$

and for $\varepsilon = 1$,

$$\text{vec } \hat{\pi}^0_{QS}(t-1,t) = (.59906, .00856, .00984, .01192, .02763, .00781, .01277, .00803, .31438)'. \tag{5.10}$$

The resulting values of $X^2(\varepsilon,QS)$ are given by

$$X^2(\varepsilon = \varepsilon_0,QS) = 10.24, \qquad G^2(\varepsilon = \varepsilon_0,QS) = 10.28,$$

$$X^2(\varepsilon = .5,QS) = 7.72, \qquad G^2(\varepsilon = .5,QS) = 7.73,$$

$$X^2(\varepsilon = 1, QS) = 7.23, \qquad G^2(\varepsilon = 1,QS) = 7.23, \tag{5.11}$$

where $\varepsilon_0 = .098$. The sample design is ignored here, but a crude adjustment can be made by the average design effect for the LFS which is expected to be around 1.75. By treating $X^2(\varepsilon,QS)/1.75$ or $G^2(\varepsilon,QS)/1.75$ as a χ^2 variable with 1 d.f., we see that even with the most liberal test when ε is at its lower bound, the observed X^2 or G^2 value adjusted for design effect lies between $\chi^2_1(.05)$ or 3.84 and $\chi^2_1(.01)$ or 6.64. Given the large sample size (n = 78,007), we may conclude that quasi-symmetry model provides a reasonably good fit to the observed proportions.

If we ignore the fact that adjusted proportions, $\hat{\pi}(t-1,t)$, do not satisfy the multi-nomial assumption, and fit the loglinear model (5.4) to $\hat{\pi}(t-1,t)$, we get

$$\tilde{X}^2(\varepsilon = \varepsilon_0,QS) = 21.28, \qquad \tilde{G}^2(\varepsilon=\varepsilon_0,QS) = 21.44,$$

$$\tilde{X}^2(\varepsilon = .5,QS) = 20.61, \qquad \tilde{G}^2(\varepsilon = .5,QS) = 20.71,$$

$$\tilde{X}^2(\varepsilon = 1,QS) = 20.57, \qquad \tilde{G}^2(\varepsilon = 1,QS) = 20.66, \tag{5.12}$$

where $\varepsilon_0 = .098$. The $\hat{\pi}(t-1,t)$ values for $\varepsilon = .098$ and 1 are given in Table 6(a). For $\varepsilon = .5$, they are obtained as

$$\text{vec } \hat{\pi}(t-1,t) = (.60665, .00752, .00314, .01056, .03323, .00381, .00670, .00325, .32515)' \tag{5.13}$$

Treating $\tilde{X}^2(\varepsilon,QS)/1.75$ or $\tilde{G}^2(\varepsilon,QS)/1.75$ as a $\chi 2$ variable with 1 d.f. will lead to the erroneous conclusion that H_{QS} is not tenable.

We also calculated the values of X^2 and G^2 under the hypothesis of symmetry, H_S as

$$X^2(\varepsilon = \varepsilon_0, S) = 77.03, \quad G^2(\varepsilon = \varepsilon_0, S) = 77.32,$$

$$X^2(\varepsilon = .5, S) = 77.00, \quad G^2(\varepsilon = .5, S) = 77.29,$$

$$X^2(\varepsilon = 1, S) = 77.00, \quad G^2(\varepsilon = 1, S) = 77.29, \tag{5.14}$$

where $\varepsilon_0 = .098$. Treating $X^2(\varepsilon,S)/1.75$ or $G^2(\varepsilon,S)/1.75$ as a $\chi 2$ variable with 3 d.f., we may conclude that H_S is not tenable because $\chi 2(.01) = 11.34$.

Treating the adjusted proportions $\hat{\Pi}(t-1,t)$ as multinomial, we get for $\varepsilon_0 = .098$,

$$\tilde{X}^2(\varepsilon = \varepsilon_0,S) = 138.76, \quad G^2(\varepsilon = \varepsilon_0,S) = 141.13,$$

$$\tilde{X}^2(\varepsilon = .5,S) = 143.82, \quad G^2(\varepsilon = .5, S) = 146.33,$$

$$\tilde{X}^2(\varepsilon = 1, S) = 144.54, \quad G^2(\varepsilon = 1,S) = 147.06. \tag{5.15}$$

As was the case with H_{QS}, both $\tilde{X}^2$ and $\tilde{G}^2$ values get inflated as compared to correct X^2 and G^2 values.

We also considered X^2 and G^2 tests for H_{QS} and H_S when the alternative value of A matrix based on the Jan '87-Nov '89 reinterview data was used. Although the values of test statistics changed somewhat because of some fluctuations in A, the overall conclusions remained the same except for the case of $X^2(\varepsilon_0,QS)$ or $G^2(\varepsilon_0,QS)$ when ε_0 is the revised value of the lower bound as .093. The actual values of X^2 and G^2 in the case of H_{QS}, for example, were obtained for $\varepsilon = .093$ as

$$X^2(\varepsilon=\varepsilon_0, QS) = 12.31, \quad G^2(\varepsilon=\varepsilon_0, QS) = 12.39,$$

$$X^2(\varepsilon=.5, QS) = 10.55, \quad G^2(\varepsilon=.5, QS) = 10.59,$$

$$X^2(\varepsilon=1, QS) = 10.17, \quad G^2(\varepsilon=1, QS) = 10.20. \tag{5.16}$$

Note that the values of X^2 or G^2 adjusted for design effect are somewhat larger than those given by (5.11). The value of ε_0 for (5.16) is also different from that in (5.11). For H_S, with $\varepsilon_0 = .093$, the X^2 and G^2 values were obtained as

$$X^2(\varepsilon=\varepsilon_0,S) = 76.96, \qquad G^2(\varepsilon=\varepsilon_0,S) = 77.24,$$

$$X^2(\varepsilon=.5,S) = 76.92, \qquad G^2(\varepsilon=.5,S) = 77.21,$$

$$X^2(\varepsilon=1,S) = 76.92, \qquad G^2(\varepsilon=1,S) = 77.21. \tag{5.17}$$

It is seen that the general trends observed earlier in (5.11) and (5.14) continue to persist with the alternative reinterview data.

6. CONCLUDING REMARKS

In this paper, the problem of adjusting classification error bias in gross flows using interview-reinterview data was considered. Under the assumption of independent classification errors (ICE), it is known that biases for different cells could be computed which indicate right directions (i.e. + or -) as expected under heuristic considerations and that some cells show substantial amounts of bias compared to other cells. To check the stability of bias adjustments with respect to uncertainty in the assumption of ICE, a class of models termed generalized independent classification errors (GICE) was proposed which parametrizes different scenarios of error mechanism by means of a mixing parameter ε. The ICE model is obtained as a special case of GICE when ε takes the maximum value of 1. The case $\varepsilon = 0$ corresponds to no error and then the response probability matrix based on the interview-reinterview data must be identity. If not, then ε must be bounded away from 0 i.e. $\varepsilon \geq \varepsilon_0$ where the lower bound ε_0 is specified by the interview-reinterview data.

Under the proposed GICE model, one can construct a bias range for each flow as ε varies from ε_0 to 1 because biases can be shown to be monotonic functions of ε under fairly mild regularity conditions. If bias ranges are narrow for some flows and wide for others, it indicates differential patterns of stability for different cells. For the particular LFS example considered, it was seen that the three cells UU, UN and NU were especially sensitive to ICE as the corresponding bias ranges were quite wide. This observation can be attributed to the nature of the response probability matrices as given by the interview-reinterview data. The cells EU and UE were found to be only marginally sensitive to ICE while the other flows were quite insensitive to ICE i.e. their bias ranges were quite narrow. However a rather striking phenomenon was observed for the three flow differences of symmetric cells, namely EU-UE, EN-NE, and NU-UN. Their biases were seen to be very stable with respect to

uncertainty in ICE. Since these flow differences are very useful in understanding movements in labour market, this finding should have important practical implications. Another important finding was that if ε is restricted to a fairly wide range, namely, [.5,1] then the adjustments under ICE seem quite robust to departures from ICE.

In view of the differential stability patterns of biases for different cells, it is not feasible, in general, to choose a single value of ε for data analysis purposes such as testing of hypothesis. Consequently, it is suggested that multiples analyses be performed by choosing several values of ε e.g. the two boundaries ε_0 and 1 and a third one near the middle such as 0.5. If the resulting analyses agree with each other reasonably well, then one would have confidence in the conclusions. However, if they are discrepant, then the results must be reported with caution after further investigations based on external considerations.

It may be remarked that the GICE modelling approach proposed in this paper can be easily extended to flow data for domains defined by age, sex, and region, for example. Separate interview-reinterview data and hence different ε_0's would be required. In this paper, we did not examine the effect of rotation group bias on gross flows. This is an important problem and should be investigated in future.

ACKNOWLEGEMENTS

We are grateful to G.J. Brackstone and G.E. Lemaitre for helpful comments and especially to W.A. Fuller for very useful suggestions and discussions. This research was partially supported by grants from the Natural Sciences and Engineerng Research Council of Canada.

REFERENCES

Abowd, J.M., and Zellner, A. (1985). Estimating gross flows, *Journal of Business and Economic Statistics*, 3, 254-283.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975). *Discrete Myltivariate Analysis: Theory and Practice*, Cambridge, Mass.: MIT Press.

Chua, T.C., and Fuller, W.A. (1987). A model for multinomial response error applied to labour flows, *Journal of the American Statistical Association*, 82, 46-51.

Fellegi, I.P. (1979). Internal memorandum to I. Macredie, Statistics Canada, Census and Household Surveys Division.

Fellegi, I.P. (1982). Internal memorandum to I. Macredie, Statistics Canada, Labour Force Activity Section.

Fuller, W.A. (1987). *Measurement Error Models*, N.Y.: John Wiley.

Fuller, W.A. (1991). Personal communication.

Fuller, W.A. and Chua, T.C. (1984). Gross change estimation in the presence of response error. *Proceedings of the Conference on Gross Flows in Labor Force Statistics*, Washington, D.C. 65-79.

Gentleman, J.F. (1988). Assessing bias due to classification error in labour force survey data. Proceedings of the *Section on Survey Research Methods*, American Statistical Association.

Huber, P.J. (1981). *Robust Statistics*, New York: John Wiley & Sons.

Lemaître, G.E. (1988). The measurement and analysis of gross flows. *Methodology Branch Working Paper No. SSMD-88-1E*. Statistics Canada.

Pericchi, L.R., and Walley, P. (1981). Robust Bayesian credible intervals and prior ignorance. *Int. Statist. Rev.*, 59, 1-23.

Poterba, J.M., and Summers, L.H. (1986). Reporting Errors and labour market dynamics, *Econometrica*, 54, 1319-1338.

Rao, J.N.K. and Thomas, D.R. (1990). Chi-squared tests with complex survey data subject to misclassification error. Paper presented at the *International Conference on Measurement Errors in Surveys*, Tucson, Arizona.

Scott, A.J., Rao, J.N.K. and Thomas, D.R. (1990). Weighted least-squares and quasi-likelihood estimation for categorical data under singular models, *Linear Algebra and its Applications*, 127, 427-447.

Veevers, R. and Macredie, I. (1983). Estimating gross flows from the Canada Labour Force Survey. Internal publication, Statistics Canada, Economic Characteristics Division.

Wong, F. (1983). Technique to correct response bias in the 4×4 labour force gross flow matrix. Internal publication, Statistics Canada, Labour Force Activity Section.

Table 1: Weighted Gross Flow Counts and Proportions (in Brackets) for October–November 1989, LFS

October 1989	November 1989 E	November 1989 U	November 1989 N	Stocks (Oct'89)
E	12,117,397 (.5991)	236,644 (.0117)	264,961 (.0131)	12,619,001 (.6239)
U	177,989 (.0088)	558,238 (.0276)	153,717 (.0076)	889,944 (.0440)
N	190,124 (.0094)	167,876 (.0083)	6,359,054 (.3144)	6,717,055 (.3321)
Stocks (Nov'89)	12,485,510 (.6173)	962,758 (.0476)	6,777,732 (.3351)	20,226,000 (1)

Table 2: Reconciled Interview-Reinterview Data (Jan'89–Nov'89)

Interview	Reinterview E	Reinterview U	Reinterview N	Row Margin
E	4,304	10	26	4,340
U	10	370	28	408
N	24	32	3,354	3,410
Column Margin	4,338	412	3,408	8,158

Table 3: Response Probability Matrices B(t-1) and B(t) from Interview-Reinterview Data of Table 2

	October '89 E	October '89 U	October '89 N	November '89 E	November '89 U	November '89 N
E	.9933	.0267	.0091	.9931	.0256	.0090
U	.0021	.9022	.0090	.0022	.9056	.0093
N	.0046	.0711	.9819	.0047	.0688	.9817
Column Total	1	1	1	1	1	1

Table 4: Reconciled Interview-Reinterview Data (Jan'87-Nov'89)

Interview	Reinterview E	Reinterview U	Reinterview N	Row Margin
E	8,221	16	41	8,278
U	25	746	44	815
N	63	60	5,830	5,953
Column Margin	8,309	822	5,915	15,046

Table 5: Response Probability Matrices B(t-1) and B(t) from Interview — Reinterview data of Table 4

	October '89			November '89		
	E	U	N	E	U	N
E	.9927	.0263	.0103	.9925	.0252	.0102
U	.0022	.9072	.0081	.0023	.9105	.0084
N	.0051	.0665	.9816	.0052	.0643	.9814
Column Total	1	1	1	1	1	1

Table 6: % Relative Adjustments in the LFS Observed Flow Proportions and their Differences using Jan'89-Nov'89 Reinterview Data

	Observed	GICE(ε=.098) Adjusted	GICE(ε=.098) % RA	ICE or GICE(ε=1) Adjusted	ICE or GICE(ε=1) % RA
EE	.5991	.60600	+1.2	.60673	+1.3
UE	.0088	.00808	-8.2	.00744	-15.5
NE	.0094	.00322	-65.7	.00314	-66.6
EU	.0117	.01111	-5.0	.01048	-10.4
UU	.0276	.03025	+9.6	.03364	+21.9
NU	.0083	.00624	-24.8	.00348	-58.1
EN	.0131	.00679	-48.2	.00669	-48.9
UN	.0076	.00567	-25.4	.00292	-61.6
NN	.3144	.32264	+2.6	.32548	+3.5
EU-UE	.0029	.00303	+4.5	.00304	+4.9
EN-NE	.0037	.00357	-3.5	.00356	-3.8
NU-UN	.0007	.00057	-18.5	.00056	-20.3

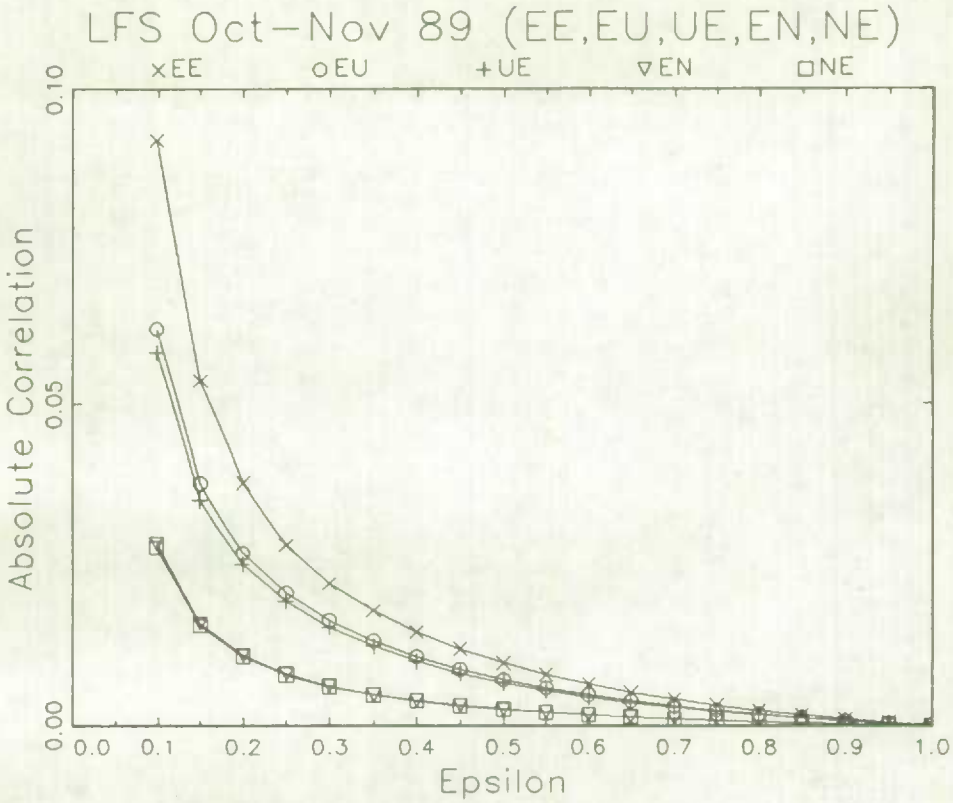
Table 7: % Relative Adjustment in the LFS Observed Flows and their Differences using Jan'87-Nov'89 Reinterview Data

		GICE(ε=.093)		ICE or GICE(ε=1)	
	Observed	Adjusted	% RA	Adjusted	% RA
EE	.5991	.60658	+1.2	.60747	+1.4
UE	.0088	.00797	-9.5	.00735	-16.5
NE	.0094	.00275	-70.7	.00248	-73.7
EU	.0117	.01098	-6.1	.01038	-11.3
UU	.0276	.03012	+9.1	.03327	+20.6
NU	.0083	.00650	-21.7	.00395	-52.5
EN	.0131	.00633	-51.6	.00605	-53.8
UN	.0076	.00591	-22.2	.00338	-55.6
NN	.3144	.32285	+2.7	.32568	+3.6
EU-UE	.0029	.00302	+3.4	.00303	+4.5
EN-NE	.0037	.00358	-3.1	.00357	-3.5
NU-UN	.0007	.00058	-16.5	.00057	-18.6

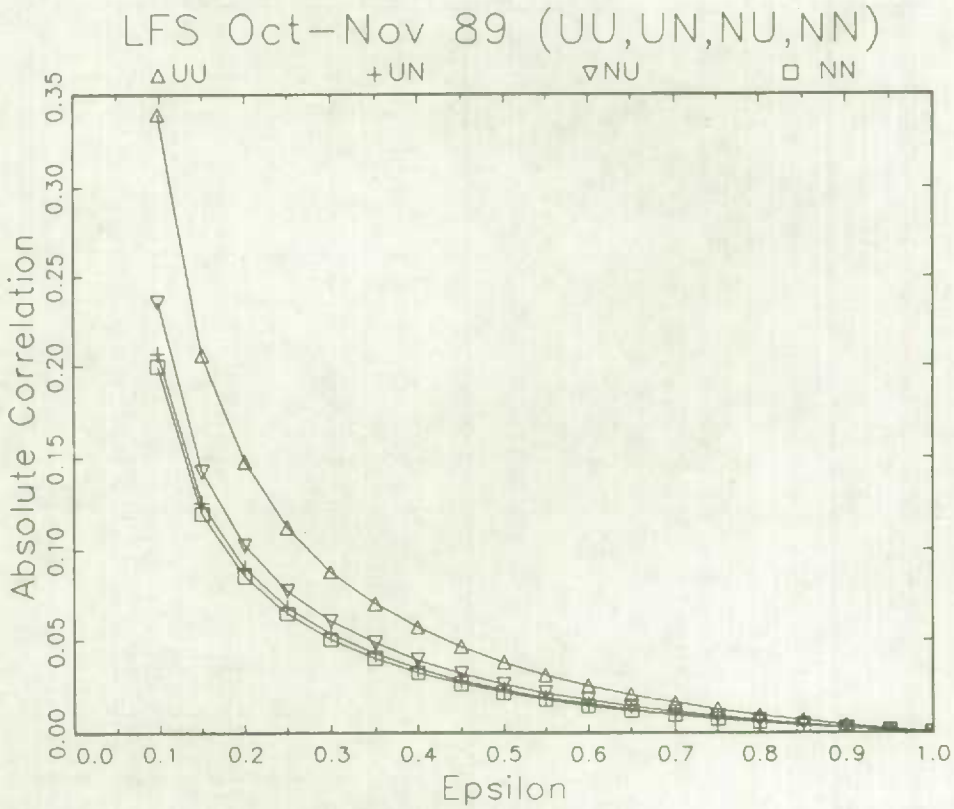
Table 8: Ranges for Bias Adjusted Flow counts and their Differences using Jan'89-Nov'89 Reinterview Data

	Observed	Adjusted ε = .098	Adjusted ε = 1
EE	12,117,397	12,256,868	12,271,614
UE	177,989	163,493	150,484
NE	190,124	65,149	63,412
Subtotal	12,485,510	12,485,510	12,485,510
EU	236,644	224,719	212,015
UU	558,238	611,831	680,362
NU	167,876	126,208	70,381
Subtotal	962,758	962,758	962,758
EN	264,961	137,367	135,372
UN	153,717	114,668	59,099
NN	6,359,054	6,525,697	6,583,261
Subtotal	6,777,732	6,777,732	6,777,732
Total Population	20,226,000	20,226,000	20,226,00
EU-UE	58,655	61,226	61,531
EN-NE	74,837	72,218	71,960
NU-UN	14,159	11,540	11,282

(a)

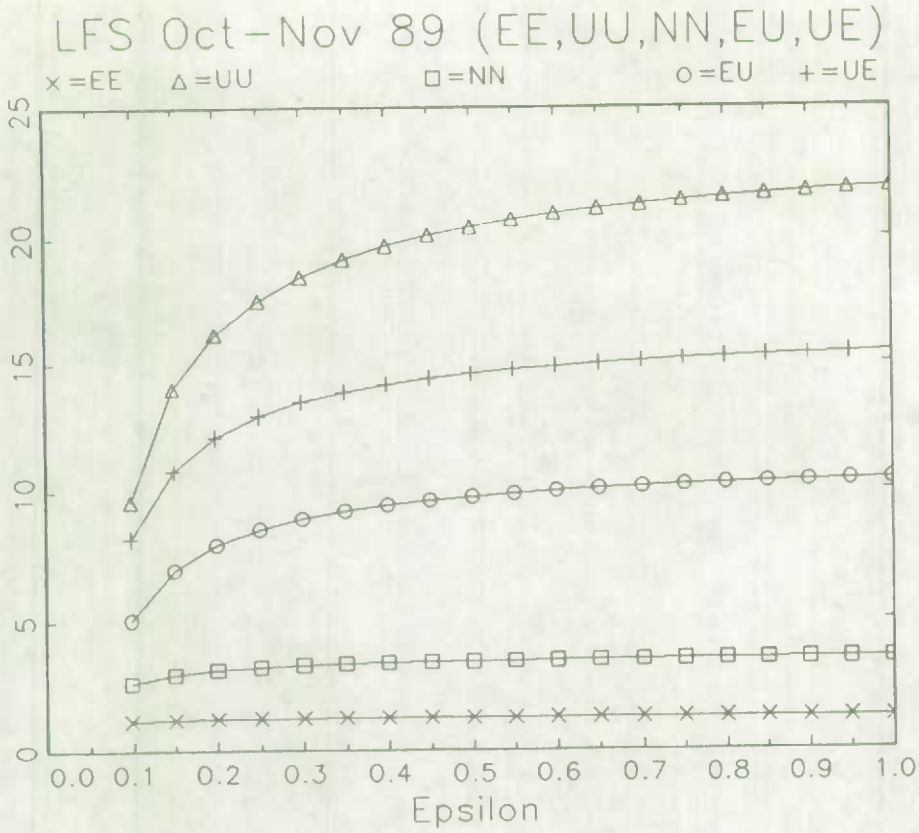


(b)

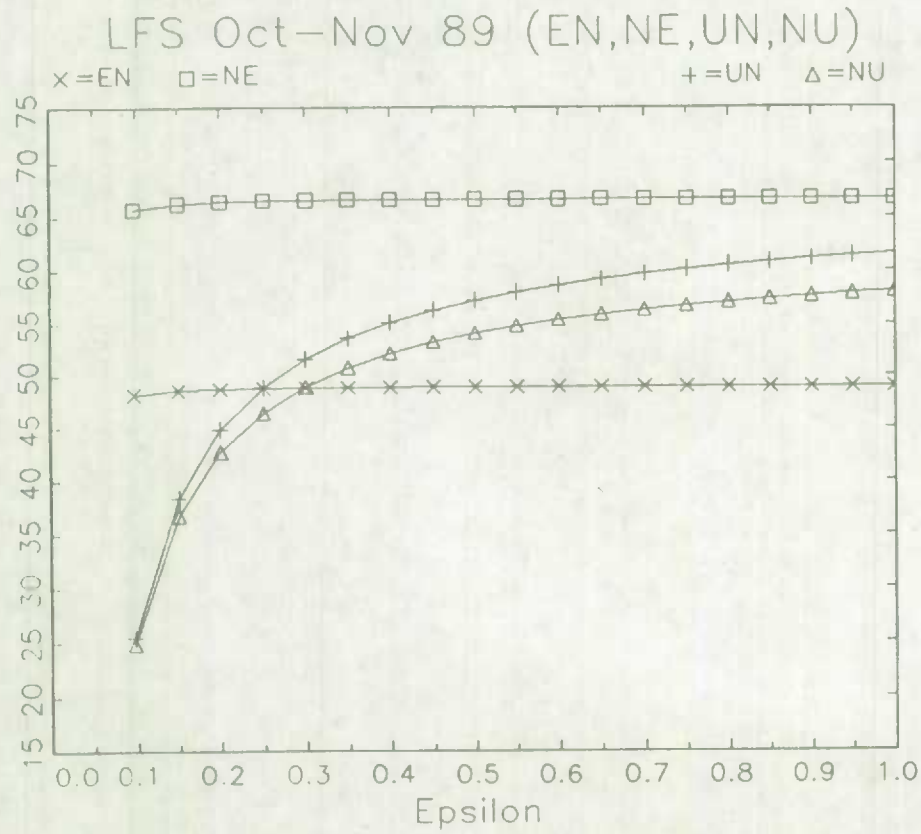


Figures 1 (a) and (b): Absolute Response Error Correlations as ε varies

(a)



(b)



Figures 2 (a) and (b): % Absolute Relative Adjustments in Flows as ε varies

(3)

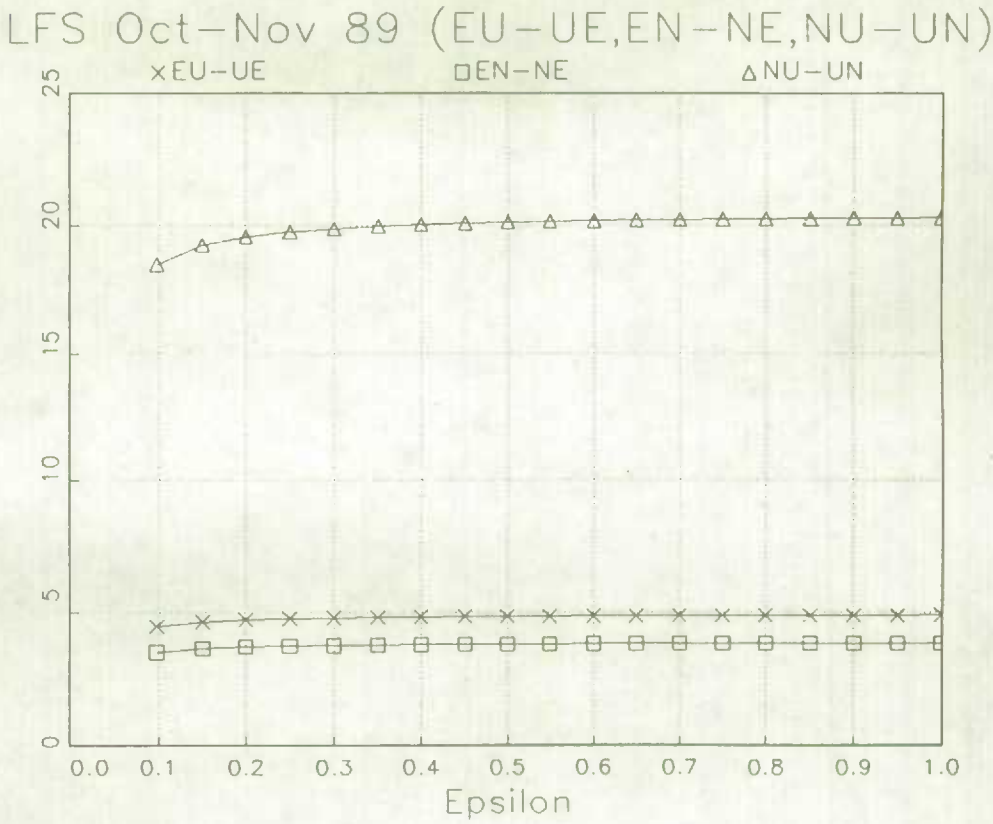


Figure 3: % Absolute Relative Adjustments in Differences of Flows as ε varies

APPENDIX

Proposition 1

$$\sum_i \sum_j \hat{\pi}_{ij}\,(t-1,t) = 1$$

Proof We need the following lemma.

<u>Lemma</u> Let C be a nonsingular matrix $((c_{ij}))$ whose inverse C^{-1} is denoted by $((c^{ji}))$. We have, for all i, j,

$$\sum_i c_{ij} = 1 \quad \text{iff} \quad \sum_j c^{ji} = 1.$$

To see this, first suppose $\sum_i c_{ij} = 1$. Now denoting the cofactor of c_{ij} by C_{ij}, we have

$$\begin{aligned}
\sum_j c^{ji} &= |C|^{-1} \sum_j C_{ij} \\
&= |C|^{-1} \sum_j \left(\sum_{i'} c_{i'j}\right) C_{ij} \\
&= |C|^{-1} \left[\sum_j c_{ij} C_{ij} + \sum_{i=i'} \sum_j c_{i'j} C_{ij}\right] \\
&= |C|^{-1} \left[|C| + 0\right] = 1
\end{aligned}$$

This proves the necessary part. The sufficiency part follows immediately.

Now to prove proposition 1, note that

$$\text{vec}\ \hat{\pi}\ (t-1,t) = A(t-1,t)^{-1}\ \text{vec}\ P(t-1,t).$$

and each column of $A(t-1,t)^{-1}$ sums to unity in view of the lemma because each column of $A(t-1,t)$ adds to one. We thus have

$$\sum_i \sum_j \pi^0_{ij}(t-1,t) = \underset{\sim}{1}' \; A(t-1,t)^{-1} \quad \text{vec } P(t-1,t),$$

$$= P_{11}(t-1,t) \quad (\text{sum of column 1 of } A(t-1,t)^{-1})$$

$$+ P_{21}(t-1,t) \; (\text{sum of column 2 of } A(t-1,t)^{-1}) + \ldots$$

$$= P_{11}(t-1,t) + P_{21}(t-1,t) + \ldots \quad (\text{by Lemma 1})$$

$$= \sum_i \sum_j P_{ij}\,(t-1,t) = 1. \qquad \text{Q.E.D.}$$

The next proposition considers behaviour of biases in $\Pi^0(t-1,t)$ as ε varies under certain mild regularity conditions about magnitudes of elements in $B(t-1)$, $B(t)$ and $\Pi(t-1,t)$ matrices. These conditions are expected to be met in practice. Let γ_{ij}'s and γ'_{ij}'s denote small positive quantities such that $(1 - \gamma_{ii})$'s, $(1 - \gamma'_{ii})$'s are diagonal elements of $B(t-1)$ and $B(t)$ respectively and γ_{ij} and γ'_{ij} are the corresponding off-diagonals. The diagonal elements of $\Pi(t-1,t)$ are also assumed to be much larger than off-diagonals. The regularity conditions required for the next proposition can be stated in terms of nine inequalities, some examples of which are:

(i) $$\gamma'_{11}\,(\Pi_{11}\gamma_{11} - \Pi_{21}\gamma_{12} - \Pi_{31}\gamma_{13}) + \gamma'_{12}\,(-\Pi_{12}\gamma_{11} + \Pi_{22}\gamma_{12} + \Pi_{32}\gamma_{13})$$
$$+ \gamma'_{13}\,(-\Pi_{13}\gamma_{11} + \Pi_{23}\gamma_{12} + \Pi_{33}\gamma_{13}) \geq 0,$$

(ii) $$\gamma'_{11}\,(\Pi_{11}\gamma_{21} - \Pi_{21}\gamma_{22} + \Pi_{31}\gamma_{23}) + \gamma'_{12}\,(-\Pi_{12}\gamma_{21} + \Pi_{22}\gamma_{22} - \Pi_{32}\gamma_{23})$$
$$+ \gamma'_{13}\,(-\Pi_{13}\gamma_{21} + \Pi_{23}\gamma_{22} - \Pi_{33}\gamma_{23}) \geq 0$$

and so on.

Notice that in the LHS of above inequalities, number of positive terms are nearly half and that they contain at least two out of the three diagonal elements of $\Pi(t-1,t)$. Therefore, these inequalities are expected to hold in practice because the Π_{ij}'s for $i \neq j$, γ's and γ''s are likely to be small.

The bias vector vec $\Pi^0(t-1,t)$ - vec $\Pi(t-1,t)$ under GICE can be expressed in two ways: $A(t-1,t)$ vec $\Pi(t-1,t)$ - vec $\Pi(t-1,t)$ or vec $\Pi^0(t-1,t)$ - $A(t-1,t)^{-1}$ vec $\Pi^0(t-1,t)$ where A varies with ε. The former expression is more convenient mathematically as it does

not require matrix inversion and is, therefore, considered in the next proposition. The estimated bias vector, vec $\hat{\Pi}^0 - A(t-1,t)^{-1}$ vec $\hat{\Pi}^0(t-1,t)$ is expected to behave in a similar manner by consistency arguments.

Proposition 2

Assume that the GICE model holds. Then, under the regularity conditions mentioned above the magnitudes of biases in $\Pi^0(t-1,t)$ are monotonic nondecreasing functions of ε. Moreover, the signs of biases are nonnegative for off-diagonal flows and nonpositive for diagonal flows.

Proof Under GICE, $\Pi(t-1,t)$ is fixed but unknown, while $\Pi^0(t-1,t)$ varies with ε. It is given by

$$\text{vec } \Pi^0(t-1,t) = \text{vec } \Pi(t-1,t) - \varepsilon(I_9 - B_\varepsilon(t) \times B_\varepsilon(t-1)) \text{ vec } \Pi(t-1,t).$$

The biases are therefore given by

$$\text{vec } \Pi^0(t-1,t) - \text{vec } \Pi(t-1,t) = - \varepsilon(I_9 - B_\varepsilon(t) \times B_\varepsilon(t-1)) \text{ vec } \Pi(t-1,t).$$

Now, the bias in the first diagonal element of $\Pi^0(t-1,t)$

$$= \text{1st element of } (-\varepsilon)\ (I_9 - B_\varepsilon(t)) \times B_\varepsilon(t-1) \text{ vec } \Pi(t-1,t)$$

$$\begin{aligned} = &-(\gamma_{11} + \gamma_{11}' + \varepsilon^{-1}\gamma_{11}\gamma_{11}')\ \Pi_{11} + (\gamma_{12} - \varepsilon^{-1}\gamma_{12}\gamma_{11}')\ \Pi_{21} + (\gamma_{13} - \varepsilon^{-1}\gamma_{13}\gamma_{11}')\ \Pi_{31} \\ &+ (\gamma_{12}' - \varepsilon^{-1}\gamma_{11}\gamma_{12}')\ \Pi_{12} + \varepsilon^{-1}\gamma_{12}\gamma_{12}'\ \Pi_{22} + \varepsilon^{-1}\gamma_{13}\gamma_{12}'\ \Pi_{32} \\ &+ (\gamma_{13}' - \varepsilon^{-1}\gamma_{11}\gamma_{13}')\ \Pi_{13} + \varepsilon^{-1}\gamma_{12}\gamma_{13}'\ \Pi_{23} + \varepsilon^{-1}\gamma_{13}\gamma_{13}'\ \Pi_{33}. \end{aligned}$$

Differentiating with respect to ε, we get

$$\begin{aligned} \frac{d}{d\varepsilon}(\text{bias in } \Pi^0_{11}(t-1,t)) = &-\varepsilon^{-2}\ [\gamma_{11}'\ (\Pi_{11}\gamma_{11} - \Pi_{21}\gamma_{12} - \Pi_{31}\gamma_{13}) \\ &+ \gamma_{12}'\ (-\Pi_{12}\gamma_{11} + \Pi_{22}\gamma_{12} + \Pi_{32}\gamma_{13}) + \gamma_{13}'\ (-\Pi_{13}\gamma_{11} + \Pi_{23}\gamma_{12} + \Pi_{33}\gamma_{13})] \end{aligned}$$

Since the above expression inside the square brackets is nonnegative under our regularity conditions, the derivative will be nonpositive for all ε and hence the bias in the first diagonal flow will be a nonincreasing function of ε as ε increases. Moreover, at $\varepsilon = 0$, bias is zero, therefore, bias will remain nonpositive for all $\varepsilon > 0$. The behaviour of biases for other diagonal flows can be similarly established.

Now consider the off-diagonal flow of $\Pi^0(t-1,t)$ in the second row and first column. Its bias is given by

$$\begin{aligned}\text{Bias in } \Pi^0_{21}(t-1,t) = {} & (\gamma_{21} - \varepsilon^{-1}\gamma_{21}\gamma'_{11})\,\Pi_{11} - (\gamma'_{11} + \gamma'_{22} - \varepsilon^{-1}\gamma_{22}\gamma'_{11})\,\Pi_{21} \\ & + (\gamma_{23} - \varepsilon^{-1}\gamma'_{11}\gamma_{23})\,\Pi_{31} + \varepsilon^{-1}\gamma_{12}\gamma_{21}\,\Pi_{12} + (\gamma'_{12} - \varepsilon^{-1}\gamma_{12}\gamma'_{22})\,\Pi_{22} \\ & + \varepsilon^{-1}\gamma'_{12}\gamma_{23}\,\Pi_{32} + \varepsilon^{-1}\gamma'_{13}\gamma_{21}\,\Pi_{13} + (\gamma'_{13} - \varepsilon^{-1}\gamma'_{13}\gamma_{22})\,\Pi_{23} + \varepsilon^{-1}\gamma'_{13}\gamma_{23}\,\Pi_{33}\end{aligned}$$

Differentiating with respect to ε, we get

$$\begin{aligned}\frac{d}{d\varepsilon}(\text{bias in } \Pi^0_{21}(t-1,t)) = \varepsilon^{-2}\,[& \gamma'_{11}(\Pi_{11}\gamma_{21} - \Pi_{21}\gamma_{22} + \Pi_{31}\gamma_{23}) \\ & + \gamma'_{12}(-\Pi_{12}\gamma_{21} + \Pi_{22}\gamma_{22} - \Pi_{32}\gamma_{23}) \\ & + \gamma'_{13}(-\Pi_{13}\gamma_{21} + \Pi_{23}\gamma_{22} - \Pi_{33}\gamma_{23})]\end{aligned}$$

Again, the expression inside the square brackets is nonnegative under the regularity conditions. Therefore, bias in the off-diagonal $\Pi^0_{21}(t-1,t)$ will be a nondecreasing function of ε. Further, bias is zero at $\varepsilon=0$, so for $\varepsilon>0$, it must be nonnegative. For other off-diagonals, bias behaviour can be similarly established.

Corollary: Biases in the flow differences $\Pi^0_{ij}(t-1,t) - \Pi^0_{ji}(t-1,t)$ for $i \neq j$, are monotonic functions of ε. For a given i,j, bias is nondecreasing if the difference in derivatives of biases in $\Pi^0_{ij}(t-1,t)$ and $\Pi^0_{ji}(t-1,t)$ is nonnegative and nonincreasing if nonpositive.