Estimation of Panel Correlations for the Canadian Labour Force Survey

HYUNSHIK LEE[1]

ABSTRACT

The Canadian Labour Force Survey uses the rotation panel design. Every month, one sixth of the sample rotates and five sixths remain. Hence, under this rotation scheme, once a rotation panel enters in the sample, it stays 6 months in the sample before it rotates out. Because of this design feature and the way of selecting the rotate-in panel, the estimates based on the panels in the same or different months are correlated. The correlation between two panel estimates is called the panel correlation. Three kinds of panel correlations are defined in this paper: (1) the correlation (denoted by ρ) between estimates for the same characteristic based on the same panel in different months; (2) the correlation (denoted by γ) between estimates of the same characteristic based on geographically neighboring panels in different months; (3) the correlation (denoted by τ) between estimates of different characteristics based on the same panel in the same or different months. This paper describes a methodology for estimating these panel correlations and presents estimated correlations for selected variables using 1980-81 and 1985-87 data with some discussion.

KEY WORDS: Repeated panel survey; Rotation; Taylor method.

1. INTRODUCTION

The Labour Force Survey (LFS) is a continuing monthly household survey which employs rotating panel design. The sample consists of six equal size rotation panels one of which is replaced by a new panel each month. The rotated-in panel stays in the sample for six months before it rotates out from the sample. (For detailed description of the LFS methodology, readers are referred to Platek and Singh (1976) and Singh *et al.* (1990).) Therefore, the estimates based on the same panel consisting of the same sampling units in different months are highly correlated. Moreover, an outgoing rotation panel is usually replaced by a neighboring panel. Because they are geographically close, estimates based on these neighboring rotation panels are also correlated. These correlations are called panel correlations. In this paper, we will describe and discuss how the panel correlations can be estimated and present their estimates for selected variables. The work was originated for the study of composite estimation technique. However, the results are applicable in any situation where the panel correlation plays a role.

The paper is structured as follows. In Section 2, necessary definitions, notations and assumptions are given. Methodology is described in Section 3 and results and discussion are given in Section 4.

2. DEFINITIONS OF PANEL CORRELATION COEFFICIENTS

To define various panel correlations we need to define common panels and the predecessor panel. A panel is identified by the panel number which indicates the duration of the panel in the sample. Thus, Panel 1 in month m, becomes Panel 2 in month $m + 1$, Panel 3 in month $m + 2$,

---

[1] H. Lee, Business Survey Methods Division, Statistics Canada, 11th Floor, R.H. Coats Building, Tunney's Pasture, Ottawa, Ontario, K1A 0T6.

Table 1

Common and Predecessor Panels Pertaining to Months m and $m - j$

m	$m-1$	$m-2$	$m-3$	$m-4$	$m-5$	$m-6$	$m-7$	$m-8$	$m-9$	$m-10$	$m-11$
1	(6)	(5)	(4)	(3)	(2)	(1)	((6))	((5))	((4))	((3))	((2))
2	1	(6)	(5)	(4)	(3)	(2)	(1)	((6))	((5))	((4))	((3))
3	2	1	(6)	(5)	(4)	(3)	(2)	(1)	((6))	((5))	((4))
4	3	2	1	(6)	(5)	(4)	(3)	(2)	(1)	((6))	((5))
5	4	3	2	1	(6)	(5)	(4)	(3)	(2)	(1)	((6))
6	5	4	3	2	1	(6)	(5)	(4)	(3)	(2)	(1)

Note: Single and double parentheses indicate *single* and *double* predecessors, respectively.

and so on. Another term *rotation group* is often used to identify a panel regardless of its duration in the sample. For instance, Rotation Group 1 which rotates in in January is identified as Rotation Group 1 throughout its stay in the sample until it rotates out in July. Then, Panel 1 in January indicates Rotation Group 1 and Panel 2 in February indicates the same rotation group which is now two months old and so on.

Two panels in two different months which represent the same rotation group are called *common panels*. When a rotation group rotates out, it is usually replaced by a rotation group consisting of neighboring households and given the same rotation group number. A panel associated with the out-going rotation group is called a *predecessor panel* of a panel associated with the in-coming rotation group. Therefore, in the above example, Panel 6 in June which is associated with Rotation Group 1 is a predecessor panel of Panel 1 in July. Table 1 shows schematically the common and predecessor panels pertaining to given months m and $m - j$.

Since each panel can be identified by two components, month and panel number, let P(month, panel number) denote a panel. Then $P(m, 4)$ and $P(m - 1, 3)$, for instance, are common panels 1 month apart. Similarly, $P(m, 4)$ and $P(m - 2, 2)$ are common panels 2 month apart. The correlation coefficient of estimates of a characteristic based on common panels that are j months apart is denoted by ρ_j. Obviously, there are no common panels which are more than 5 months apart and thus, the subscript j can be at most 5. We assume that ρ_j is independent of m and panel number. However, it is a function of j and varies between characteristics.

The correlation coefficient of estimates based on a panel and its predecessor that are j months apart is denoted by γ_j. But in this case, j can go up to 11, *i.e.* γ_{11} is the last correlation coefficient in this series and it is the correlation between $P(m, 6)$ and $P(m - 11, 1)$. We assume again that γ's are independent of m and panel number. They do, however, depend on characteristic as well as j as ρ-correlations do.

The third type of panel correlation is defined as the correlation between estimates for two different characteristics based on common panels and denoted by τ_j for common panels that are j months apart. Now j can take values from 0 to 5. The same assumptions as for the ρ's and γ's apply here as well.

The formal definitions of ρ's, γ's and τ's are as follows:

Let $y_{m,l}$ be the LFS estimate of a characteristic of interest obtained from $P(m,l)$. We assume that $V(y_{m,l}) = \sigma_y^2$ regardless of m and l. Then, ρ_j's are defined by

$$\text{Cov}(y_{m,l}, y_{m-j,l-j}) = \rho_j \sigma_y^2, \quad 1 \le j \le 5, \quad j < l \le 6,$$

and γ_j's by

$$\text{Cov}(y_{m,l}, y_{m-j,6+l-j}) = \gamma_j \sigma_y^2,$$

where $1 \leq l \leq j$ if $1 \leq j \leq 6$ and $j - 5 \leq l \leq 6$ if $7 \leq j \leq 11$.

It would be natural to conjecture that ρ_j's and γ_j's decrease as the subscript j increases and that ρ_j's are larger than γ_j's because ρ_j's are correlations pertaining common households while γ_j's are those pertaining neighboring households. We can also define the correlation between a panel and the predecessor of the panel's predecessor (denoted by double parentheses and called *double predecessor* in Table 1) in a similar way, say δ, and thus, we have $\delta_7, \delta_8, \ldots,$ δ_{17}. They will be smaller than γ_j's but could be quite close to them for the same subscript because double and single predecessors are close geographically. However, the δ-correlations are not considered here due to time and resource constraints.

We assume that $\text{Cov}(y_{m,l}, y_{m,l'}) = 0$ if $l \neq l'$ and $\text{Cov}(y_{m,l}, y_{m-j,l'}) = 0$ if $P(m - j, l')$ is not a common panel nor a predecessor of $P(m,l)$.

In order to define τ-correlations, let $x_{m,l}$ be the LFS estimate of another characteristic obtained from $P(m,l)$ and let $V(x_{m,l}) = \sigma_x^2$ be independent of m and l. Then τ-correlations are defined by

$$\text{Cov}(y_{m,l}, y_{m-j,l-j}) = \tau_j \sigma_x \sigma_y, \quad 0 \leq j \leq 5, \quad j < l \leq 6.$$

3. ESTIMATION OF THE PANEL CORRELATIONS

Since a variance estimation computer program was available, the method described here was geared to use this program with minimum modification. The methodology used in the program is the generalized Keyfitz method (Choudhry and Lee 1987; Lee 1989a) better known as the Taylor method. The program can compute variance estimates of linear combinations of monthly estimates.

We employ the following basic equality to estimate the desired correlations using the existing variance program:

$$\text{Cov}(A,B) = \frac{V(A) + V(B) - V(A - B)}{2}. \tag{1}$$

From the program, $V(A - B), V(A)$ and $V(B)$ can be obtained and so can $\text{Cov}(A,B)$ using (1). An expression for $V(A - B)$ from which (1) can be obtained is also given in Kish (1965).

3.1 Estimation of ρ-Correlations

Let $A = \sum_{l=2}^{6} y_{m,l}$ and $B = \sum_{l=1}^{5} y_{m-1,l}$. A and B are obtained by eliminating Panel 1 from month m and Panel 6 from month $m - 1$, respectively. Note that the eliminated panels are uncommon and the remaining ones are all common. Using the variance program we compute estimates of $V(A - B)$, $V(A)$ and $V(B)$ and obtain estimates of $\text{Cov}(A,B)$ by (1). From the assumptions given in Section 2, it is easy to see that

$$\text{Cov}(A,B) = 5\rho_1 \sigma_y^2,$$

$$V(A) = V(B) = 5\sigma_y^2,$$

and thus,

$$\rho_1 = \frac{\mathrm{Cov}(A,B)}{\sqrt{V(A)V(B)}}. \tag{2}$$

An estimate of ρ_1 is then obtained by substituting estimates of $\mathrm{Cov}(A,B), V(A)$ and $V(B)$. Estimates of ρ_2, ρ_3 and ρ_4 can be obtained in the same way by putting $A = \sum_{l=j+1}^{6} y_{m,l}$, and $B = \sum_{l=1}^{6-j} y_{m-j,l}, j = 2, 3, 4$. But there is some problem in estimating ρ_5 this way. When we drop all uncommon panels from months m and $m - 5$, only one panel is left in each month and this causes problem in variance estimation for Self-Representing Units (SRUs). SRUs are large cities each of which is represented in the survey by independent sampling. There is no such problem for Non-Self-Representing Units (NSRUs) which are the areas outside of the SRUs, containing rural areas and small urban centers. In NSRUs, each Primary Sampling Unit (PSU), which becomes a replicate for variance estimation, has all rotation panels and thus, even after eliminating 5 uncommon panels, there is still one panel remaining in the PSU so that variance can be computed. In SRUs, however, rotation panels form replicates and if there is only one panel left, then there is only one replicate in each stratum and thus, variance can not be computed in the usual way. Therefore, $\hat{\rho}_5$ was obtained by prediction using a nonlinear regression $\rho = a + bt + ce^{-t}, t = 1, \ldots, 4$. Another way to estimate ρ_5 will be discussed later in Subsection 4.1.

3.2 Estimation of γ-Correlations

It is easy to see that $\mathrm{Cov}(A,B) = (5\rho_1 + \gamma_1)\sigma_y^2$ if $A = \sum_{l=1}^{6} y_{m,l}$ and $B = \sum_{l=1}^{6} y_{m-1,l}$. In general,

$$\mathrm{Cov}(A,B) = \{(6 - j)\rho_j + j\gamma_j\}\sigma_y^2,$$

where

$$A = \sum_{l=1}^{6} y_{m,l},$$

$$B = \sum_{l=1}^{6} y_{m-j,l}, \quad j = 1, \ldots, 4.$$

Then, an estimate of γ_j can be obtained from the following equation:

$$\gamma_j = \frac{1}{j}\left[6\frac{\mathrm{Cov}(A,B)}{\sqrt{V(A)V(B)}} - (6 - j)\rho_j\right], \tag{3}$$

by substituting estimated values on the right. There is a direct way to estimate these γ-correlations including γ_5 by

$$\gamma_j = \frac{\mathrm{Cov}(A_j,B_j)}{\sqrt{V(A_j)V(B_j)}}, \tag{4}$$

where $A_j = \sum_{l=1}^{j} y_{m,l}$ and $B_j = \sum_{l=7-j}^{6} y_{m-j,l}, j = 2, \ldots, 5$. In Section 4, the two methods were compared by using empirical data.

Other γ-correlations ($\gamma_j, j = 6, \ldots, 10$) are obtained by (4) with

$$A_j = \sum_{l=j-5}^{6} y_{m,l},$$

$$B_j = \sum_{l=1}^{12-j} y_{m-j,l}.$$

There is no simple way of estimating γ_{11} directly or indirectly. Both $\hat{\gamma}_5$ and $\hat{\gamma}_{11}$ were predicted by a log-linear model $\gamma = \exp(a + bt)$, $t = 1, \ldots, 4, 6, \ldots, 10$.

3.3 Estimation of τ-Correlations

These correlations can be estimated by the same way as the ρ-correlations just by replacing $y_{m,l}$ by $x_{m,l}$. Let $A = \sum_{l=j+1}^{6} x_{m,l}$ and $B = \sum_{l=1}^{6-j} y_{m-j,l}, j = 0, 1, \ldots, 4$. Then we have

$$\text{Cov}(A,B) = (6 - j)\tau_j \sigma_x \sigma_y,$$

$$V(A) = (6 - j)\sigma_x^2,$$

$$V(B) = (6 - j)\sigma_y^2,$$

from which we get

$$\tau_j = \frac{\text{Cov}(A,B)}{\sqrt{V(A)V(B)}}, \quad j = 0, 1, \ldots, 4. \tag{5}$$

All τ's can be estimated using (5) except τ_5 which is predicted by a log-linear model, $\tau = \exp(a + bt)$, $t = 1, \ldots, 4$.

4. RESULTS AND DISCUSSION

By using the methods discussed in the previous section, estimates of ρ- and γ-correlations were computed from the 1980-81 and 1985-87 LFS data for 5 characteristics: In Labour Force (IN LF), Employed (EMP), Employed Agriculture (EMP AG), Employed Non-Agriculture (EMP NON-AG), Unemployed (UNEMP). The panel correlations were estimated for only 3 provinces, Nova Scotia (NS), Ontario (ONT), and British Columbia (BC) from the 1980-81 data. However, the estimation was extended to all provinces when more recent data (March 1985 – February 1987) were used. Moreover, 4 more characteristics, the employed and the unemployed of two age groups, 15-24 and 25+ (EMP 15-24, EMP 25+, UNEMP 15-24, UNEMP 25+), were added. The estimation of τ-correlations was done only for those additional characteristics for NS, ONT and Alberta (ALT) from the 1985-87 data.

In the following, only part of these results will be presented and discussed. All the results are available in Lee (1989b).

4.1 Estimates of ρ-Correlations

The results of estimated ρ-correlations are given in Table 2. Even though estimates for the 5 characteristics (IN LF, EMP, EMP AG, EMP NON-AG, UNEMP) from the 1985-87 data are available for all provinces, the results for only 3 provinces, NS, ONT and BC, are presented for a historical comparison. Table 2 also shows the results for the other 4 characteristics (EMP 15-24, EMP 25 +, UNEMP 15-24, UNEMP 25 +) from the provinces of NS and ONT.

The ρ-correlations are generally high as expected because they are correlations for the common panels. The correlations for EMP AG are the highest and those for UNEMP are the lowest. It seems that the size of the ρ-correlation indicates the degree of mobility of the labour force with a particular characteristic. For instance, the high ρ-correlation for EMP AG shows a low mobility of the labour force in agriculture while a high mobility of unemployed labour force is demonstrated in its low ρ-correlation. The different levels of mobility of labour force in two age groups are also evident. The younger group (15-24) is more mobile than the older one (25 +).

The decreasing trend of the ρ-correlations over time is clearly demonstrated in the results. The trend was extremely well fitted by a nonlinear regression model $\rho_t = a + bt + ce^{-t}$. The R-squares (multiple correlations) are close to 1 (> 0.98). Therefore, the predicted values for ρ_5 seem to be very good. In Lee (1989a and 1989b), $\hat{\rho}_5$ was obtained by extrapolating $\hat{\rho}_3$ and $\hat{\rho}_4$ instead. The differences between the predicted and extrapolated values for $\hat{\rho}_5$, however, are very small. They are less than 0.01 for all characteristics except for UNEMP, UNEMP 15-24 and UNEMP 25 + where the largest difference is 0.03.

Table 2

Estimates of ρ-Correlations (1980-81 and 1985-87 Data)

Prov	Characteristic	80–81 Data					85–87 Data				
		$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{\rho}_5$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$	$\hat{\rho}_4$	$\hat{\rho}_5$
NS	IN LF	0.862	0.797	0.744	0.679	0.622	0.845	0.769	0.730	0.696	0.670
	EMP	0.866	0.783	0.714	0.651	0.590	0.863	0.768	0.713	0.686	0.660
	EMP AG	0.913	0.837	0.756	0.678	0.598	0.912	0.867	0.825	0.802	0.773
	EMP NON-AG	0.865	0.774	0.710	0.649	0.594	0.873	0.779	0.724	0.697	0.670
	UNEMP	0.590	0.455	0.333	0.243	0.145	0.703	0.546	0.426	0.415	0.375
	EMP 15-24						0.773	0.632	0.556	0.495	0.446
	EMP 25 +						0.878	0.800	0.754	0.729	0.705
	UNEMP 15-24						0.618	0.454	0.364	0.300	0.246
	UNEMP 25 +						0.695	0.554	0.443	0.440	0.406
ONT	IN LF	0.843	0.782	0.717	0.674	0.622	0.846	0.781	0.732	0.681	0.635
	EMP	0.852	0.779	0.709	0.664	0.611	0.853	0.771	0.706	0.648	0.592
	EMP AG	0.955	0.926	0.901	0.861	0.827	0.962	0.948	0.944	0.937	0.934
	EMP NON-AG	0.861	0.791	0.724	0.678	0.625	0.866	0.795	0.746	0.701	0.660
	UNEMP	0.580	0.445	0.334	0.286	0.222	0.579	0.436	0.328	0.291	0.238
	EMP 15-24						0.747	0.605	0.500	0.429	0.356
	EMP 25 +						0.888	0.824	0.777	0.732	0.691
	UNEMP 15-24						0.468	0.339	0.257	0.219	0.178
	UNEMP 25 +						0.622	0.468	0.365	0.313	0.256
BC	IN LF	0.849	0.767	0.705	0.665	0.622	0.817	0.753	0.701	0.647	0.597
	EMP	0.835	0.755	0.695	0.651	0.607	0.851	0.770	0.711	0.651	0.597
	EMP AG	0.896	0.809	0.733	0.656	0.582	0.938	0.886	0.847	0.828	0.805
	EMP NON-AG	0.855	0.769	0.715	0.661	0.616	0.857	0.784	0.730	0.679	0.632
	UNEMP	0.516	0.407	0.334	0.320	0.294	0.634	0.524	0.459	0.363	0.290

4.2 Estimates of γ-Correlations

As mentioned in Subsection 3.2, there are two ways of estimating γ_2, γ_3 and γ_4, that is, by formulae (3) and (4). We will call the method by (3) as Method 1 and that by (4) as Method 2. Only Method 1 can be used to estimate γ_1 while direct estimation of γ_5 is feasible only by Method 2. The two methods are compared in Table 3 using empirical data. In the table, $\hat{\gamma}_5$'s for Method 1 are predicted values by a log-linear model. The table shows that the two methods produced somewhat different results. The correlations produced by Method 2 clearly show an increasing trend contrary to our intuition while Method 1 gave more acceptable results. Moreover, if we compare these correlations with $\hat{\gamma}_1$ in Table 4A (which had to be estimated by Method 1), Method 1 seems to produce more reasonable results than Method 2. Therefore, we adopted Method 1. However, if everything is correct, the two methods should be equivalent and produce similar results. It seems that the real data do not conform to some extent with the assumptions we made to derive the formulae.

Estimates of the γ-correlations are presented in Tables 4A and 4B. The size of γ-correlations is much smaller than that of ρ-correlations as we expected. But it also reflects differences in mobility of the labour force with different characteristics as seen from the results of ρ-correlations.

The overall trend of $\hat{\gamma}$'s is somewhat fuzzy, especially for the results from the 1985-87 data. There are about 25% of cases – a case is a row entry in the tables – in Table 4B which show an increasing trend. In those cases, the log-linear regression lines have a positive slope even though it is fairly small in magnitude. Moreover, in most of those cases, R-squares are small, which indicates that fittings by the log-linear model are not good. This does not mean, however, that there are other models which can fit the data better. Rather it means that no clear trend is exhibited. Among the cases that show a decreasing trend, about half of the cases have an R-square greater than 0.5.

The results from the 1980-81 data show a quite different picture. There is only one case that shows an increasing trend and most of the cases have R-squares > 0.5. In fact, the results for NS and BC look more reasonable than those for ONT as far as the trend is concerned.

Table 3

Comparison of Estimates of γ_2, γ_3, γ_4 and γ_5 Obtained by Different Methods
(Ontario, 1980-81)

Characteristic	Method	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\hat{\gamma}_5$
IN LF	1	0.141	0.128	0.133	0.135
	2	0.107	0.105	0.116	0.120
EMP	1	0.136	0.142	0.142	0.147
	2	0.100	0.115	0.126	0.133
EMP AG	1	0.483	0.474	0.486	0.451
	2	0.321	0.370	0.407	0.448
EMP NON-AG	1	0.150	0.147	0.157	0.163
	2	0.117	0.134	0.145	0.149
UNEMP	1	0.074	0.076	0.063	0.080
	2	0.043	0.056	0.046	0.043

Note: Methods 1 and 2 are defined by the formulae (3) and (4) in Section 3, respectively.

Table 4A

Estimates of γ-Correlations
(1980-81 Data)

Prov	Characteristic	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\hat{\gamma}_5$	$\hat{\gamma}_6$	$\hat{\gamma}_7$	$\hat{\gamma}_8$	$\hat{\gamma}_9$	$\hat{\gamma}_{10}$	$\hat{\gamma}_{11}$
NS	IN LF	0.288	0.263	0.265	0.250	0.236	0.233	0.211	0.199	0.193	0.167	0.164
	EMP	0.262	0.219	0.228	0.226	0.219	0.239	0.210	0.200	0.188	0.161	0.172
	EMP AG	0.351	0.308	0.283	0.237	0.205	0.190	0.141	0.113	0.063	0.021	0.007
	EMP NON-AG	0.238	0.187	0.189	0.180	0.164	0.151	0.123	0.121	0.136	0.091	0.086
	UNEMP	0.106	0.176	0.091	0.097	0.091	0.076	0.066	0.063	0.066	0.032	0.031
ONT	IN LF	0.161	0.141	0.128	0.133	0.135	0.136	0.125	0.127	0.124	0.122	0.117
	EMP	0.164	0.136	0.142	0.142	0.147	0.149	0.148	0.150	0.153	0.141	0.146
	EMP AG	0.477	0.483	0.474	0.486	0.451	0.474	0.459	0.429	0.394	0.323	0.368
	EMP NON-AG	0.184	0.150	0.147	0.157	0.163	0.167	0.166	0.169	0.174	0.156	0.165
	UNEMP	0.141	0.074	0.076	0.063	0.080	0.051	0.045	0.060	0.077	0.136	0.074
BC	IN LF	0.177	0.137	0.117	0.119	0.119	0.112	0.101	0.112	0.094	0.066	0.070
	EMP	0.211	0.146	0.133	0.107	0.101	0.083	0.050	0.068	0.058	−0.033	−0.015
	EMP AG	0.380	0.311	0.301	0.272	0.241	0.216	0.198	0.170	0.122	0.078	0.071
	EMP NON-AG	0.207	0.166	0.161	0.129	0.108	0.093	0.069	0.038	0.023	−0.004	−0.020
	UNEMP	0.126	0.125	0.114	0.103	0.091	0.076	0.062	0.092	0.032	0.040	0.031

Table 4B

Estimates of γ-Correlations
(1985-87 Data)

Prov	Characteristic	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\hat{\gamma}_5$	$\hat{\gamma}_6$	$\hat{\gamma}_7$	$\hat{\gamma}_8$	$\hat{\gamma}_9$	$\hat{\gamma}_{10}$	$\hat{\gamma}_{11}$
NS	IN LF	0.250	0.238	0.247	0.230	0.216	0.204	0.181	0.196	0.189	0.162	0.160
	EMP	0.170	0.183	0.205	0.196	0.185	0.157	0.158	0.194	0.198	0.219	0.198
	EMP AG	0.326	0.296	0.246	0.245	0.265	0.267	0.234	0.217	0.259	0.269	0.231
	EMP NON-AG	0.146	0.168	0.199	0.201	0.178	0.153	0.152	0.189	0.199	0.216	0.201
	UNEMP	0.233	0.267	0.241	0.211	0.206	0.168	0.171	0.176	0.157	0.187	0.147
	EMP 15-24	0.107	0.127	0.140	0.133	0.112	0.105	0.099	0.107	0.090	0.074	0.082
	EMP 25 +	0.088	0.075	0.117	0.108	0.100	0.099	0.090	0.103	0.099	0.137	0.118
	UNEMP 15-24	0.051	0.080	0.042	0.024	0.054	0.061	0.079	0.081	0.058	0.011	0.049
	UNEMP 25 +	0.155	0.129	0.177	0.171	0.148	0.159	0.158	0.127	0.102	0.134	0.124
ONT	IN LF	0.162	0.138	0.141	0.134	0.132	0.135	0.127	0.116	0.111	0.103	0.101
	EMP	0.114	0.122	0.121	0.122	0.117	0.124	0.119	0.108	0.110	0.112	0.111
	EMP AG	0.508	0.518	0.553	0.561	0.571	0.569	0.582	0.617	0.668	0.650	0.672
	EMP NON-AG	0.133	0.140	0.132	0.140	0.157	0.156	0.168	0.182	0.204	0.205	0.210
	UNEMP	0.030	0.047	0.055	0.047	0.043	0.048	0.039	0.030	0.039	0.048	0.041
	EMP 15-24	0.012	−0.006	0.018	0.031	0.017	0.023	0.011	0.011	0.016	0.044	0.029
	EMP 25 +	0.354	0.358	0.349	0.343	0.319	0.312	0.298	0.285	0.276	0.240	0.246
	UNEMP 15-24	0.068	0.039	0.038	0.058	0.033	0.026	0.008	0.018	0.011	−0.002	−0.006
	UNEMP 25 +	0.052	0.054	0.033	0.017	0.034	0.033	0.026	0.018	0.021	0.044	0.022
BC	IN LF	0.103	0.095	0.113	0.103	0.090	0.090	0.091	0.083	0.078	0.030	0.055
	EMP	0.125	0.100	0.112	0.111	0.116	0.135	0.123	0.121	0.118	0.095	0.114
	EMP AG	0.394	0.443	0.426	0.401	0.396	0.400	0.401	0.381	0.347	0.334	0.345
	EMP NON-AG	0.080	0.067	0.076	0.072	0.091	0.109	0.111	0.118	0.112	0.106	0.124
	UNEMP	0.096	0.086	0.084	0.080	0.083	0.097	0.068	0.074	0.068	0.083	0.071

Table 5

Estimates of τ-Correlations

x_1: EMP 15-24, x_2: EMP 25 + , x_3: UNEMP 15-24, x_4: UNEMP 25 + ,

(1985-87 Data)

Province	Characteristic	$\hat{\tau}_0$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$	$\hat{\tau}_5$
NS	(x_1,x_2)	0.150	0.140	0.148	0.181	0.187	0.196
	(x_1,x_3)	− 0.440	− 0.275	− 0.187	− 0.135	− 0.039	0.126
	(x_1,x_4)	− 0.036	− 0.040	− 0.043	− 0.015	0.024	0.022
	(x_2,x_3)	− 0.029	− 0.037	− 0.078	− 0.049	− 0.016	− 0.038
	(x_2,x_4)	− 0.437	− 0.374	− 0.276	− 0.182	− 0.231	− 0.094
	(x_3,x_4)	0.136	0.127	0.094	0.055	0.049	0.020
ONT	(x_1,x_2)	0.092	0.070	0.055	0.040	0.028	0.010
	(x_1,x_3)	− 0.420	− 0.267	− 0.205	− 0.161	− 0.145	− 0.010
	(x_1,x_4)	− 0.065	− 0.056	− 0.053	− 0.036	− 0.028	− 0.019
	(x_2,x_3)	− 0.061	− 0.054	− 0.054	− 0.042	− 0.089	− 0.074
	(x_2,x_4)	− 0.392	− 0.303	− 0.230	− 0.187	− 0.181	− 0.077
	(x_3,x_4)	0.058	0.043	0.022	0.013	0.022	0.001

4.3 Estimates of τ-Correlations

Table 5 contains estimates of τ-correlations obtained from the 1985-87 data for all possible combinations of EMP 15-24 (denoted by x_1), EMP 25 + (x_2), UNEMP 15-24 (x_3) and UNEMP 25 + (x_4). The correlations between x_1 and x_2 are positive as well as those between x_3 and x_4. Other correlations are mostly negative. In terms of magnitude, only the correlations pertaining to (x_1, x_3) and (x_2, x_4) are quite different from zero. Others are close to zero. These observations seem to agree with what we understand about the movement of labour force between the employed and the unemployed in the same age group. When the employment increases, the unemployment decreases and vice versa. The trend is obviously upward in these cases.

The data were fit by a log-linear model and τ_5's were predicted. The model fitting seems reasonable except for the correlations between (x_2, x_3) whose R-squares are very small in both provinces NS and ONT.

4.4 Conclusions

The estimation of correlations from complex survey data is a difficult problem. It is so not because the derivation of formulae is difficult – in fact, the formulae given here are elementary – but because there are many practical constraints in applying the formulae. If we had not made the assumptions in Section 3, the estimation of the panel correlations by using the existing computer program would have been impossible. On the other hand, these assumptions should be conformable to the real data to which the formulae are applied. In our case, there seem to be some unconformable elements in the assumptions we made to the real data, which was indicated by the discrepancy in the results obtained by formulae (3) and (4) (see Table 3). Nevertheless, the estimates are not thought to be unreasonable.

In a study of the composite estimator for the LFS, the results given in this paper were successfully used to compare various composite estimators (Kumar and Lee 1983). Recently

Binder and Dick (1990) proposed a method for analyzing Seasonal ARIMA models by taking the survey errors into account. They applied their technique to the LFS data using the estimated panel correlations. However, in cases when the results to be obtained by the use of the estimated panel correlations are sensitive to the accuracy of these estimates, the results should be interpreted carefully.

ACKNOWLEDGEMENT

REFERENCES

CHOUDHRY, G.H., and LEE, H. (1987). Variance estimation for the Canadian Labour Force Survey. *Survey Methodology*, 13, 147-161.

BINDER, D.A., and DICK, J.P. (1990). Analysis of seasonal ARIMA models from survey data. *Survey Methodology*, this issue.

KISH, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.

KUMAR, S., and LEE, H. (1983). Evaluation of composite estimation for the Canadian Labour Force Survey. *Survey Methodology*, 9, 178-201.

LEE, H. (1989a). Variance estimation methodology and general purpose variance estimation system for the Labour Force Survey. Methodology Working Paper Series, SSMD-89-022 E, Statistics Canada.

LEE, H. (1989b). Estimation of panel correlations for the Canadian Labour Force Survey. Methodology Working Paper Series, SSMD-89-023 E, Statistics Canada.

PLATEK, R., and SINGH, M.P. (1976). *Methodology of the Canadian Labour Force Survey*. Catalogue No. 71-526, Statistics Canada.

SINGH, M.P., DREW, J.D., GAMBINO, J., and MAYDA, F. (1990). *Methodology of the Canadian Labour Force Survey*. Catalogue No. 71-526, Statistics Canada.