Statistics Canada Statistique Canada

Methodology Branch

Household Survey
Methods Division

Direction de la méthodologie

Division des méthodes
d'enquêtes des ménages

Canada

WORKING PAPER
METHODOLOGY BRANCH

Estimating the Effect of the New Labour Force Survey Questionnaire on Temporary Layoffs

Wesley Yung and Ritu Kaushal

HSMD-97-002E

Survey and Analysis Methods Development Section
Household Survey Methods Division, Statistics Canada

March, 1997

Estimating the Effect of the New Labour Force Survey Questionnaire on Temporary Layoffs

Wesley Yung and Ritu Kaushal

Survey and Analysis Methods Development Section
Household Survey Methods Division, Statistics Canada

ABSTRACT

Previous research suggests that the Canadian Labour Force Survey (LFS) questionnaire adopted in 1976 underestimates the number of persons on temporary layoff. Some persons on temporary layoff have likely been identified as *not in the labour force* instead of being *unemployed*, with a corresponding underestimation of the unemployment rate. As part of the LFS redesign a new questionnaire, to be phased in from September 1996 to January 1997, includes improved identification of persons on temporary layoff. Prior to the phase-in of the redesigned questionnaire, a temporary layoff supplement was administered from August 1995 to August 1996 to measure the effect of the improved identification of temporary layoffs. In this paper, we report our findings on the effect of the improved identification of temporary layoffs on the unemployment rate and on the historical time series of unemployment rates. This paper also illustrates the use of the linearized jackknife variance estimator under a nonresponse weighting adjustment.

Estimation de l'effet du nouveau questionnaire de l'Enquête sur la population active sur les mises à pied temporaires

Wesley Yung et Ritu Kaushal

Section de l'élaboration des méthodes d'enquête et d'analyse
Division des méthodes d'enquêtes-ménages

RÉSUMÉ

Les recherches antérieures font ressortir que le questionnaire de l'Enquête sur la population active adopté en 1976 sous-estime le nombre de personnes mises à pied de façon temporaire. Il semble que certaines de ces personnes aient été identifiées comme *ne faisant pas partie de la population active* plutôt qu'*en chômage,* ce qui a mené à une sous-estimation correspondante du taux de chômage. Dans le cadre du remaniement de l'EPA, un nouveau questionnaire, qui sera adopté de façon échelonnée de septembre 1996 à janvier 1997, permet une meilleure identification des personnes mises à pied de façon temporaire. En attendant l'adoption du nouveau questionnaire, un supplément sur les mises à pied temporaires a été utilisé d'août 1995 à août 1996, en vue de mesurer l'effet d'une meilleure détermination des mises à pied temporaires. Dans le présent document, nous rendons compte des résultats de notre recherche sur l'effet de la détermination améliorée des mises à pied temporaires sur le taux de chômage et sur les séries chronologiques s'y rapportant. Le document illustre en outre le recours à l'estimateur de variance jackknife linéarisé après rajustement des poids pour la non-réponse.

Table of Contents

1.0 INTRODUCTION

The August 1996 Labour Force Survey (LFS) marks the last occasion in which the LFS estimates are based solely on the questionnaire adopted in 1976. Consultation with the survey's user community has led to the design of a new questionnaire which will provide new information on a variety of topics such as wages, union status, job turnover and better information on seasonal and other temporary jobs. Beginning in September 1996, all households rotating into the LFS sample will receive the new questionnaire. By January 1997, the new questionnaire will be fully implemented.

The new questionnaire is consistent with most of the fundamental concepts and question style of the old questionnaire. While the definitions and the associated questions measuring usual work hours, involuntary part-time and discouragement were changed to yield more accurate and relevant information, these changes are not expected to affect the major indicators of labour force status. On the other hand, changes were made to improve the identification of persons on temporary layoff. The improved question set is expected to lead to a slight increase in the number of persons identified as *unemployed* and consequently, *in the labour force*. These changes are expected to lead to a small increase in the unemployment rate.

Considerable effort has been made to estimate the effects of the improved identification of persons on temporary layoff. A temporary layoff supplement, consisting of questions that reflect the new approach to identifying temporary layoffs, has been given to selected LFS respondents from August 1995 to August 1996. This paper presents the anticipated impact of the new questions on the unemployment rate based on the analysis of the temporary layoff supplement. Upper bounds on the increase in the unemployment rate are obtained using a linearized jackknife variance estimator.

2.0 MISIDENTIFICATION OF TEMPORARY LAYOFFS

The primary function of the Labour Force Survey is to classify persons as *employed*, *unemployed* or *not in the labour force*. This classification is straightforward for most persons but can be problematic for some individuals. In particular, many persons on temporary layoff fail to identify themselves as having job attachment in response to the current question "Last week, did ... have a job or business at which he/she did work?" A negative response to this question prevents classification as a temporary layoff and the respondent is not considered to be unemployed unless he/she has searched for work in the preceding four weeks. Since most persons on temporary layoff do not search for employment, the result may be an underestimation of persons *unemployed* and consequently, persons *in the labour force*.

The problem of underestimating the temporary layoffs has been previously studied. Robinson (1989) compared the LFS estimates of temporary layoffs with those obtained from administrative data. According to the administrative data, about a third of all "unemployed" persons return to their former employer while according to the LFS estimates, only about 5% of unemployed persons are temporary layoffs. While much of the difference can be accounted for by seasonal returns (not removed from the administrative data), it is likely that some temporary layoffs are being missed by the LFS.

Within Statistics Canada, Kinack (1991) performed a longitudinal analysis of the LFS data to identify the presence of response inconsistencies associated with the measurement of temporary layoffs. Utilizing the rotating panel design of the LFS, Kinack analyzed individual records to identify logical inconsistencies and recurring code changes at the respondent level. These inconsistencies and code changes are both indicative of misunderstood questions or misapplied concepts. Kinack's study found that many non-employed respondents who were identified as permanent layoffs from their last job actually returned to work at the same job sometime during their subsequent months in the LFS.

To further assess this potential problem of misclassification, a small follow-up survey was conducted in March 1992. A small sample of respondents who were classified as either temporary layoffs or permanent layoffs in the LFS were reinterviewed one week later using a short test questionnaire that identified job attachment differently. In particular, respondents who were not currently employed were asked for specific reasons for leaving their last job. If job loss was because of business conditions or if layoff was specifically mentioned, the respondent was asked about the expectation of recall. The result of this alternative questioning was a doubling in the number of persons classified as temporary layoffs. While it was recognized that the small sample size of the follow-up survey could lead to large coefficients of variation, it was felt that an improvement of temporary layoff estimates from the LFS could be obtained by incorporating this alternative questioning approach into the redesigned LFS questionnaire.

3.0 TEMPORARY LAYOFF SUPPLEMENT

The Temporary Layoff Supplement (TLOS) was administered during regular LFS interviewing but was given only to specific LFS respondents who were rotating out of the LFS sample. Using the specifications for the computer assisted interviewing application, as outlined in Kinack (1995), the computer application identified respondents who were "potential temporary layoffs" and should receive the TLOS. Based on the respondent's answers to the TLOS, their labour force status (*unemployed* or *not in the labour force*) was determined. To estimate the effect of the improved identification of temporary layoffs, two unemployment rates were calculated; one based on the original LFS labour force status and one based on the labour force status determined by the TLOS. For those individuals who did not receive a TLOS, their LFS labour force status was used for both rates. While the LFS sample was "clean" (i.e. adjusted for nonresponse, imputation performed for missing values, etc...) the nonresponse to the TLOS sample had to be addressed.

The nonresponse to the TLOS was divided into three sources: 1) true nonresponse where the respondent was contacted but refused to answer the questions on the TLOS; 2) Individuals who were not identified in the field as "potential temporary layoffs" but once their responses were edited, they became "potential temporary layoffs". (Note, that some individuals were identified as "potential temporary layoffs" in the field but once their responses were edited, they were no longer "potential temporary layoffs". These individuals were omitted from the TLOS sample); 3) As the supplement ran, the TLOS was not being generated for some of the "potential temporary layoffs" because some interviewers did not have the correct TLO application on their machines. As interviewers picked up new versions of the case management software, the TLO application was being lost. This problem was corrected as of the May survey. Sources 2 and 3 are actually not nonresponse problems, but are in fact frame problems. Since "potential temporary layoffs" could not be properly identified until after editing, our "frame" was a conceptual frame and individuals who did not receive a supplement due to not being identified as a "potential temporary layoff" are part of the undercovered population. Although individuals could be from any one of the three sources, they were all treated as nonrespondents and were compensated for by using a nonresponse weighting adjustment.

Preliminary investigation of the TLOS data indicated that eligible individuals had different response profiles depending on their original labour force status. That is, individuals with an original labour force status of *unemployed* responded to the TLOS at a different rate then those individuals with an original labour force status of *not in the labour force*. Therefore, two nonresponse classes were defined on the basis of the original labour force status and weighting adjustments were applied independently within each class. Details of the nonresponse weighting adjustments are given in section 4.1

Along with the nonresponse/undercoverage problems experienced, it was felt that difficulties due to the learning curve experienced by interviewers and start-up problems with the processing systems could put the quality of the data at risk. Because of this, the August 1995 data was not used for analysis and the September 1995 data was used with caution.

4.0 ANALYSIS

4.1 ESTIMATION

The LFS uses a stratified multi-stage design consisting of n_h clusters sampled from N_h clusters in the *h-th* stratum, *h=1,...L*. Within each sampled cluster, further subsampling is performed according to some probability sampling design. Associated with the *k-th* sampled element within the *i-th* sampled cluster of the *h-th* stratum is the subweight, w_{hik}, the variable of interest, y_{hik}, and a vector of auxiliary variables, z_{hik}. The vector of auxiliary variables is used for benchmarking purposes through the generalized regression estimator.

To estimate the total number of unemployed persons, let y_{hik} be the indicator variable for unemployment status defined as

$$y_{hik} = \begin{cases} 1 & \text{if } (hik)\text{-}th \text{ individual is unemployed} \\ 0 & \text{otherwise.} \end{cases}$$

Assuming full response, an estimator of the total number of unemployed persons is

$$\hat{Y}_{UNP} = \sum_{(hik)\in s} \tilde{w}_{hik} y_{hik},$$

where $(hik)\in s$ denotes all units in the sample, s, and $\tilde{w}_{hik}$ is the regression adjusted weight defined as

$$\tilde{w}_{hik} = w_{hik} z_{hik}^T \hat{A}^{-1} Z,$$

with

$$\hat{A} = \sum_{(hik)\in s} w_{hik} z_{hik} z_{hik}^T$$

and Z is the vector of known population totals corresponding to the auxiliary variables z_{hik}. For the LFS, the population control totals consist of counts for 30 age-sex groups, for LFS economic regions and for large urban centres. For more on the benchmarking of LFS weights, see Kennedy (1996). Similarly, to estimate the total number of persons in the labour force, let

$$x_{hik}=\begin{cases}1 & \text{if } (hik)\text{-}th \text{ individual is in the labour force}\\ 0 & \text{otherwise.}\end{cases}$$

Then an estimator for the total number of persons in the labour force is

$$\hat{Y}_{INLF}=\sum_{(hik)\in s}\tilde{w}_{hik}x_{hik},$$

and the unemployment rate is estimated by

$$\hat{UR}=\frac{\hat{Y}_{UNP}}{\hat{Y}_{INLF}}.$$

These expressions are valid for the LFS sample but need to be modified for the TLOS sample due to the nonresponse adjustment.

For notational simplicity, the LFS sample will be partitioned into two parts; individuals who, after editing, are identified as "potential temporary layoffs" (regardless of whether they received or responded to the TLOS) will make up s_2, and the remaining individuals, who should not receive a TLOS, will make up s_1. That is, the LFS sample, s, is divided into "TLO eligible", s_2, and "TLO ineligible", s_1. Now, within the s_2 sample there are nonrespondents who will be handled through a nonresponse weighting adjustment. Before defining the nonresponse weighting adjustment, we first define the response indicator variable, a_{hik}, as

$$a_{hik}=\begin{cases}1 & \text{if } (hik)\text{-}th \text{ individual responds}\\ 0 & \text{otherwise.}\end{cases}$$

We will assume that all individuals in s_1 respond and will suppress the response indicator variable when dealing with s_1.

As previously mentioned, nonresponse classes were formed on the basis of an individual's original labour force status. Again for notational simplicity we define the weighting class indicator variable, $\delta_{hik,l}$, as

$$\delta_{hik,l}=\begin{cases}1 & \text{if } (hik)\text{-}th \text{ individual is in the } l\text{-}th \text{ weighting class}\\ 0 & \text{otherwise.}\end{cases}$$

The nonresponse weighting adjustment for the *l-th* weighting class is then defined as

$$d_l = \frac{\sum_{(hik)\in s_2} w_{hik}\,\delta_{hik,l}}{\sum_{(hik)\in s_2} w_{hik}\,a_{hik}\,\delta_{hik,l}},$$

where $(hik)\in s_2$ denotes all sampled units in s_2. Note that the nonresponse adjustment is calculated using the subweights, w_{hik}, and not the regression adjusted weights, $\tilde{w}_{hik}$. The TLO adjusted estimator of the total number of persons unemployed is then

$$\hat{Y}^{/}_{UNP} = \sum_{(hik)\in s_1} \tilde{w}_{hik}\,y_{hik} + \sum_l \sum_{(hik)\in s_2} d_l\,\tilde{w}_{hik}\,a_{hik}\,y^{/}_{hik}\,\delta_{hik,j},$$

where $y^{/}_{hik}$ is the indicator variable for *unemployed* based on the TLOS. Note that the regression adjusted weight is the same as defined previously. Similarly, let

$$\hat{Y}^{/}_{INLF} = \sum_{(hik)\in s_1} \tilde{w}_{hik}\,x_{hik} + \sum_l \sum_{(hik)\in s_2} d_l\,\tilde{w}_{hik}\,a_{hik}\,x^{/}_{hik}\,\delta_{hik,j},$$

where $x^{/}_{hik}$ is the indicator variable for *in the labour force* based on the TLOS. The TLO adjusted unemployment rate is then given as

$$\hat{UR}^{/} = \frac{\hat{Y}^{/}_{UNP}}{\hat{Y}^{/}_{INLF}}.$$

Two analyses of the TLOS data were performed: 1) A monthly comparison of the LFS unemployment rate and the TLO adjusted unemployment rate and 2) A comparison of the two unemployment rates over six month periods. Since the TLOS is given to respondents who are rotating out of the LFS sample, for each month we have data for only one rotation group. Thus when comparing monthly unemployment rates, we are actually comparing the unemployment rates for a single rotation group. Similarly, when comparing rates over the six month period, we are actually comparing rates for six consecutive rotation groups. The two analyses will be discussed separately.

4.2 MONTHLY COMPARISONS

The first analysis focused on the difference between the unemployment rate based on the LFS sample and the TLOS sample for a rotation group. Because the comparisons occurred for a rotation group, variances cannot be estimated for this analysis. Results of the monthly comparisons appear in Table 1. Note that August 1995 does not appear in Table 1 since it was felt that the data from August 1995 was of poor quality due to reasons previously mentioned.

From Table 1, we can see that the unemployment rate increases slightly for all months except September 1995 with an average increase of 0.12 percentage points. The months of December and January have the largest differences, possibly due to the seasonal nature of temporary layoffs. It is worth noting, when considering the differences in this analysis, that since the estimates are based on a single rotation group they will have more variability than the estimates in the following analysis.

Table 1
Monthly Unemployment Rates (%) for the LFS and TLO Supplement

Month	LFS Rate	TLO Rate	Difference
September 1995	7.56	7.56	0.00
October 1995	8.23	8.35	0.13
November 1995	8.05	8.14	0.09
December 1995	8.88	9.12	0.24
January 1996	9.50	9.77	0.27
February 1996	10.46	10.64	0.17
March 1996	10.10	10.24	0.13
April 1996	9.40	9.52	0.12
May 1996	9.36	9.49	0.13
June 1996	8.83	8.89	0.06
July 1996	10.08	10.14	0.06
August 1996	8.98	8.99	0.01
Average	9.12	9.24	0.12

4.3 SIX MONTH COMPARISON

In the six month analysis, data from six consecutive months were combined to mimic the rotating panel design of the LFS and to allow for the calculation of variance estimates. Unemployment rates were calculated from the LFS data and the TLOS data and a variance estimator for the difference of the two rates was obtained. The variance of the difference was used to produce an upper bound for the difference of the two rates using a one-sided 97.5% confidence interval. A one-sided interval was used because it was expected that the new questionnaire would produce an increase in the unemployment rate. In order to simplify variance calculations, the estimate of the total number of persons in the labour force based on the original LFS was used for both the LFS unemployment rate and the TLO adjusted unemployment rate. This will result in a slightly higher TLO unemployment rate than the correct unemployment rate which uses the TLO estimate of the total number of persons in the labour force. Details of the variance estimator are given in Appendix 1. Results of the six month analysis are presented in Table 2.

Table 2
Unemployment Rates (%) for the LFS and TLO Supplement for Six Month Periods

Time Period	LFS	TLO	Difference	SE of Diff.	Upper Bound
09/95 to 02/96	8.83	8.99	0.16	0.039	0.22
10/95 to 03/96	9.25	9.43	0.18	0.036	0.24
11/95 to 04/96	9.42	9.61	0.19	0.038	0.25
12/95 to 05/96	9.66	9.85	0.19	0.038	0.25
01/96 to 06/96	9.66	9.81	0.16	0.035	0.22
02/96 to 07/96	9.73	9.85	0.12	0.036	0.19
03/96 to 08/96	9.48	9.56	0.08	0.032	0.15
Average			0.15	0.036	0.22

The average difference between the two unemployment rates is approximately 0.15 percentage points and the average upper bound for the difference is 0.22 percentage points. There appears to be an increasing trend in the differences for the first time period to the fourth time period but note that the first time period includes the month of September which showed no difference between the two rates in the monthly analysis. It is felt that the data from September is suspect due to interviewers becoming familiar with the TLO application and that this monthly rate is reducing the true difference in the first time period. In addition, the first four time periods include the peak months of December and January.

5.0 EFFECT ON HISTORICAL TIME SERIES

Of particular concern to Statistics Canada is the effect of the new questionnaire on the major seasonally adjusted time series. A possible option is to adjust the historical time series using an adjustment factor based on the phase-in data. As the phase-in data are not available, the temporary layoff numbers from the TLOS were used to simulate the phase-in of the new questionnaire. We stress that the effect as estimated from the TLOS will not be used to adjust the historical LFS series and that if the historical series is adjusted then the adjustment factors should be based on the phase-in data. The following study is intended only as an example of how the redesigned questionnaire may affect the historical LFS series.

Unadjusted (i.e. not seasonally adjusted) employment and unemployment series were obtained for the years 1984 to 1995 for six age-sex groups: 15 to 24 years old, 25 to 54 years old and 55+ for both males and females. The increase in temporary layoffs, as obtained from the TLOS, for the six age-sex groups appear in Table 3. These numbers represent the total increases in temporary layoffs for each month. Recall that the TLOS produced estimates of the increase in temporary layoffs for a rotation group. To obtain the increase for each month, the estimated rotation group increases from the TLOS were simply multiplied by a factor of six to represent the six rotation groups of the LFS design.

Table 3
Monthly Increases in Temporary Layoffs by Age-Sex Groups

		Male			Female		
Month	Year	15-24	25-54	55+	15-24	25-54	55+
September	1995	282	3,924	0	0	-1,782	-3,300
October	1995	348	8,370	0	1,866	5,298	636
November	1995	648	2,958	0	0	6,006	0
December	1995	3,684	19,404	0	-1,314	9,576	0
January	1996	2,670	14,184	774	306	16,512	0
February	1996	144	7,314	2,334	0	9,318	42
March	1996	3,048	2,178	2,196	528	5,298	1,464
April	1996	1,650	10,050	0	0	762	0
May	1996	5,886	12,960	0	0	1,122	-2,388
June	1996	0	3,870	-438	798	2,094	-1,350
July	1996	3,684	2,958	774	0	1,122	1,464
August	1996	3,684	2,958	774	0	1,122	1,464

From Table 3 we see that the number of temporary layoffs can actually decrease due to better identification of temporary layoffs. Also, note that the month of July and August 1996 are equivalent due to the fact that at the time of the analysis, August 1996 data were not available.

The phase-in of the new questionnaire will begin in September 1996 with all households rotating into the LFS sample receiving the new questionnaire. Thus in September 1996 one rotation group will receive the new questionnaire, two rotation groups in October, three rotation groups in November, four rotation groups in December and finally all rotation groups starting in January 1997. To simulate this phase-in, the temporary layoff increases were introduced as outlined in Table 4. For the months of February to August, the full temporary layoff increases were used.

Table 4
Temporary Layoffs During Simulated Phase-in

Month	Year	Temporary Layoff Increase
September	1995	1/6 of total September TLO increase
October	1995	1/3 of total October TLO increase
November	1995	1/2 of total November TLO increase
December	1995	2/3 of total December TLO increase
January	1996	100% total January TLO increase

Using the temporary layoff numbers in Table 3, a September 1989 phase-in of the new questionnaire was simulated and the resulting series were seasonally adjusted using the X-11 ARIMA model. Two phase-in strategies were used: 1) phase-in the increases obtained from the TLOS and 2) adjust the historical series before the phase-in using an adjustment (Morry, 1996) recommended by the Time Series Research and Analysis Division (TSRAD).

The TSRAD adjustment is defined as follows. In each age-sex group g, calculate a new series

$$\alpha_{ij}^{g} = \frac{TL_{ij}^{g}}{UE_{ij}^{g}},$$

where TL_{ij}^{g} = the number of temporary layoffs in age-sex group g, year i and month j and
UE_{ij}^{g} = the number of unemployed in age-sex group g, year i and month j.

Next, calculate the average monthly effect of the temporary layoffs as a proportion of unemployed for age-sex group g, as

$$\alpha_j^g = \sum_{i=1}^{N} \frac{\alpha_{ij}^g}{N}.$$

Now, define the increase due to the new questionnaire as

$$\beta_j = \frac{TL_j^{/}}{TL_j},$$

where $TL_j^{/}$ = the number of temporary layoffs in month j based on the new questionnaire and
TL_j = the number of temporary layoffs in month j based on the old questionnaire.

The adjustment factor for month j and age-sex group g is defined as

$$\begin{aligned} F_j^g &= (\alpha_j^g \cdot \beta_j) + (1 - \alpha_j^g) \\ &= 1 - \alpha_j^g (1 - \beta_j). \end{aligned}$$

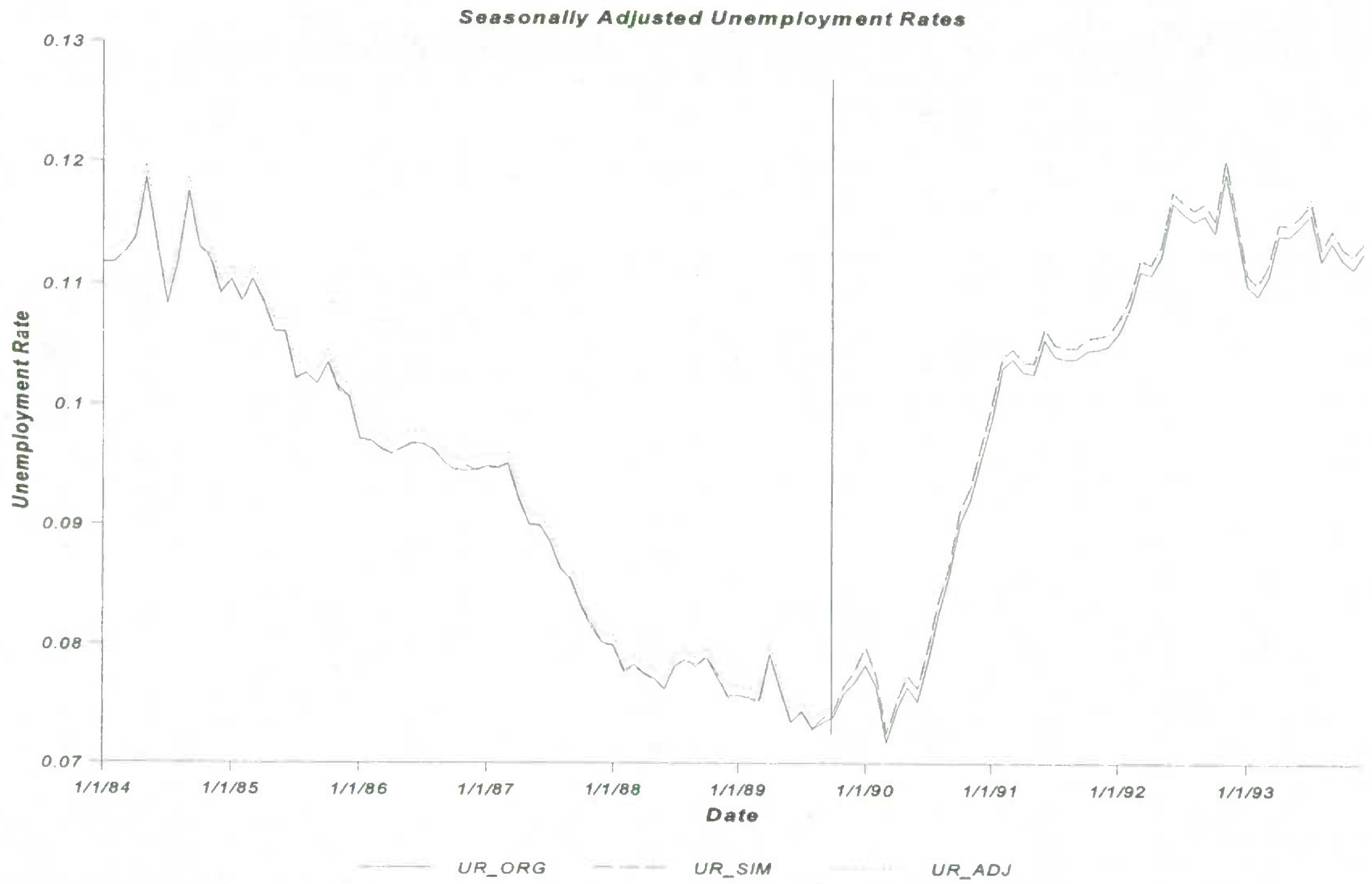
Note that if $\beta_j = 1$ (i.e. no effect of new questionnaire) then $F_j^g = 1$ and there is effectively no adjustment. The values of F_j^g appear in Table 5. Once these adjustments were calculated, the series before the September 1989 phase-in was adjusted.

The seasonally adjusted unemployment rates are shown in Graph 1. Three unemployment rates are shown: the original unemployment rate, the unemployment rate with the simulated phase-in, that is for all j and g, and the F_j^g adjusted unemployment rate with the simulated phase-in. From the graph we see that before the phase-in, the simulated and original series are almost equivalent with the adjusted series slightly higher. After the phase-in period, as expected the simulated and the adjusted series are similar while the original series is slightly lower. During the phase-in period, the simulated series makes a smooth transition from the original series, at the beginning of the phase-in, to the adjusted series, at the end of the phase-in period. On a whole, there is very little difference between the three series with the simulated series being slightly higher than the original series.

Table 5
Monthly Adjustment Factor by Age-Sex Groups

	Male			Female		
Month	15-24	25-54	55+	15-24	25-54	55+
January	1.0025	1.0029	1.0010	1.0003	1.004	1.000
February	1.0015	1.0015	1.0029	1.000	1.0024	1.0001
March	1.0028	1.0005	1.0027	1.0005	1.0014	1.0032
April	1.0015	1.0020	1.000	1.000	1.0002	1.000
May	1.0047	1.0026	1.000	1.000	1.0003	0.9948
June	1.000	1.0001	0.9995	1.0007	1.0005	0.9971
July	1.0026	1 0006	1.0009	1.0000	1.0003	1.0033
August	1.0026	1 0006	1.0009	1.0000	1.0003	1.0033
September	1.0002	1.0008	1.0000	1.0000	0.9996	0.9929
October	1.0003	1.0016	1.0000	1.0017	1.0013	1.0014
November	1.0006	1.0006	1.0000	1.0000	1.0015	1.0000
December	1.0033	1.0039	1.0000	0.9989	1.0024	1.0000

Graph 1


Seasonally Adjusted Unemployment Rates
Unemployment Rate
0.13
0.12
0.11
0.1
0.09
0.08
0.07
1/1/84
1/1/85
1/1/86
1/1/87
1/1/88
1/1/89
1/1/90
1/1/91
1/1/92
1/1/93
Date
UR_ORG
UR_SIM
UR_ADJ

6.0 CONCLUSIONS

The main purpose of the Temporary Layoff Supplement was to estimate the effect of the improved identification of TLO's on the major LFS estimates. As expected the new questionnaire led to an increase in the total number of persons on temporary layoff although the increase produced a minimal rise in the unemployment rate. The results from the monthly comparison showed an average increase of 0.12 percentage points and from the six month analysis, an average increase of 0.15 percentage points was observed. It should be noted that both of these increases fall within the bounds of the sampling variability for the unemployment rate based on the existing questionnaire. The observed monthly increase in TLO's was used to investigate the impact on the historical time series for the unemployment rate. Results of this investigation indicate that the impact will be minimal and if the effect estimated from the phase-in of the new questionnaire is similar to that of the TLOS, adjustment of the historical data will not be necessary.

REFERENCES

Caron, N. and Sunter, D. (1996). Phasing in the New LFS Questionnaire: The Impact of Improved Measurement of Persons on Temporary Layoff. *The Labour Force*, Statistics Canada Publication 71-001-XPB, Statistics Canada.

Kennedy, B. (1997). Weighting and Estimation Methodology of the Canadian Labour Force Survey. Methodology Branch Working Paper, in press, Statistics Canada.

Kinack, M. (1991). Temporary Layoffs in the LFS: An Examination of Longitudinal Data. Unpublished memo, Statistics Canada.

Kinack, M. (1995). Labour Force Survey Supplement on Temporary Layoffs: Specifications for Computer Assisted Interviewing. Unpublished memo, Statistics Canada.

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979). *Multivariate Analysis*. London: Academic Press Limited.

Morry, M. (1996). Private communication.

Robinson, M. (1989). Temporary Layoffs and Unemployment in Canada. *A Journal of Economy and Society*, **28**, 1, 82-90.

Sunter, D., Kinack, M., Akyeampong, E. and Charrette, D. (1995). Redesigning the Canadian Labour Force Survey Questionnaire. Household Surveys Division Working Paper, Statistics Canada.

Yung, W., and Rao, J.N.K. (1996). Jackknife Linearization Variance Estimators Under Stratified Multi-Stage Sampling. *Survey Methodology*, 22, 23-31.

Yung, W. (1996). Contributions to Poststratification in Stratified Multi-Stage Samples. Unpublihsed Ph.D. thesis, Carleton University, Ottawa, Canada.

APPENDIX 1

A linearized jackknife variance estimator was used to estimate the variance of the difference between the LFS unemployment rate and the TLO-adjusted unemployment rate. For more on the linearized jackknife variance estimator we refer the reader to Yung and Rao (1996) and Yung (1996). As mentioned in section 4.2, the estimate of the total number of persons in the labour force obtained from the LFS sample was used in both unemployment rates. We wish to derive an estimator of the variance of

$$\tilde{UR} - \hat{UR} = \frac{\hat{Y}^{/}}{\hat{X}} - \frac{\hat{Y}}{\hat{X}}$$

$$= \frac{1}{\hat{X}} (\hat{Y}^{/} - \hat{Y})$$

where $\tilde{UR}$ is the TLO- adjusted unemployment rate with the total number of persons in the labour force estimated from the LFS sample. Now letting $\hat{D} = \hat{Y}^{/} - \hat{Y}$, the variance of $\tilde{UR} - \hat{UR}$ can be approximated as

$$V(\hat{D}/\hat{X}) = \frac{1}{\hat{X}^2} V(\hat{D} - (D/X)\hat{X})$$

$$= \frac{1}{\hat{X}^2} V(\hat{D}) + \frac{1}{\hat{X}^2} (D/X)^2 V(\hat{X}) - \frac{2}{\hat{X}^2} (D/X)\, Cov(\hat{D},\hat{X})$$

$$\cong \frac{1}{\hat{X}^2} V(\hat{D})$$

where we have assumed that the true difference D/X is small enough so that the contribution of the $(D/X)\hat{X}$ term is negligible. The effect of the $(D/X)\hat{X}$ term was investigated and was found to have an insignificant contribution to the variance of $\tilde{UR} - \hat{UR}$. Thus as a variance for $\tilde{UR} - \hat{UR}$ we use the approximation

$$V(\tilde{UR} - \hat{UR}) \cong \frac{1}{\hat{X}^2} V(\hat{Y}^{/} - \hat{Y}),$$

and will derive a linearized jackknife variance estimator for

$$V(\hat{Y}^{/} - \hat{Y}) = V(\hat{Y}^{/}) + V(\hat{Y}) - 2\, COV(\hat{Y}^{/}, \hat{Y}).$$

To derive a linearized jackknife variance estimator for $V(\hat{Y})$ we first consider the jackknife variance estimator for $\hat{Y}$, where

$$\hat{Y} = \sum_{(hik)\in s} \tilde{w}_{hik} y_{hik},$$

with $\tilde{w}_{hik} = w_{hik} z_{hik}^T \hat{A}^{-1} \mathbf{Z}$ being the regression adjusted weight, $\hat{A} = \sum_{(hik)\in s} w_{hik} z_{hik} z_{hik}^T$ and $\mathbf{Z}$ is the vector of known population totals corresponding to the auxiliary variables z_{hik}.

To construct a jackknife variance estimator for $\hat{Y}$, we first define the jackknife weights, $w_{hik(gj)}$, when the *(gj)-th* cluster has been deleted as

$$w_{hik(gj)} = \begin{cases} 0 & \text{if } (hi)=(gj) \\ \dfrac{n_g}{n_g - 1} w_{gik} & \text{if } h=g \text{ and } i \neq j \\ w_{hik} & \text{if } h \neq g. \end{cases}$$

The regression adjusted jackknife weights are then defined as

$$\tilde{w}_{hik(gj)} = w_{hik(gj)} z_{hik}^T \hat{A}_{(gj)}^{-1} \mathbf{Z},$$

where $\hat{A}_{(gj)} = \sum_{(hik)\in s} w_{hik(gj)} z_{hik} z_{hik}^T$. The estimate of the total number of unemployed persons based on the LFS sample when the *(gj)-th* cluster has been deleted is then given by

$$\hat{Y}_{(gj)} = \sum_{(hik)\in s} \tilde{w}_{hik(gj)} y_{hik},$$

and the jackknife variance estimator for $\hat{Y}$ is

$$v_J(\hat{Y}) = \sum_g \frac{n_g - 1}{n_g} \sum_j \left(\hat{Y}_{(gj)} - \hat{Y}\right)^2.$$

To obtain a linearized jackknife variance estimator, we approximate the difference $\hat{Y}_{(gj)} - \hat{Y}$ as follows. Using the definition of the jackknife weights, see equation (1), rewrite $\hat{A}_{(gj)}$ as

$$
\begin{aligned}
\hat{A}_{(gj)} &= \sum\nolimits_{(hik)\in s} w_{hik(gj)}\, z_{hik} z_{hik}^{T} \\
&= \sum\nolimits_{(hik)\in s, h\neq g} w_{hik}\, z_{hik} z_{hik}^{T} + \frac{n_g}{n_g - 1} \sum\nolimits_{(gik)\in s, i\neq j} w_{gik}\, z_{gik} z_{gik}^{T} \\
&= \sum\nolimits_{(hik)\in s} w_{hik}\, z_{hik} z_{hik}^{T} + \frac{n_g}{n_g - 1} \sum\nolimits_{(gik)\in s, i\neq j} w_{gik}\, z_{gik} z_{gik}^{T} - \sum\nolimits_{(gik)\in s} w_{gik}\, z_{gik} z_{gik}^{T} \\
&= \hat{A} + \frac{n_g}{n_g - 1} \left(\frac{1}{n_g} \sum\nolimits_{(gik)\in s} w_{gik}\, z_{gik} z_{gik}^{T} - \sum\nolimits_{(gjk)\in s} w_{gjk}\, z_{gjk} z_{gjk}^{T} \right) \\
&= \hat{A} + D_{gj}.
\end{aligned}
$$

The notation $(gik)\in s$ denotes summation over i and k while g is fix. Similarly, $(gjk)\in s$ denotes summation over k while g and j are fixed. In general, whenever g or j appear as indices of a summation, they are fixed. Using the matrix identity (see equation A.2.4f on page 459 in Mardia et al, 1979)

$$
(\boldsymbol{I} + \boldsymbol{P}\boldsymbol{Q})^{-1} = \boldsymbol{I} - \boldsymbol{P}(\boldsymbol{I} + \boldsymbol{Q}\boldsymbol{P})^{-1}\boldsymbol{Q},
$$

we get (see the Appendix in Yung and Rao, 1996),

$$
\hat{A}_{(gj)}^{-1} \cong \hat{A}^{-1} - \hat{A}^{-1} D_{gj} \hat{A}^{-1}.
$$

Putting this expression for $\hat{A}_{(gj)}^{-1}$ in equation (2) gives

$$
\begin{aligned}
\tilde{w}_{hik(gj)} &\cong w_{hik(gj)}\, z_{hik}^{T} \left(\hat{A}^{-1} - \hat{A}^{-1} D_{gj} \hat{A}^{-1} \right) \boldsymbol{Z} \\
&= w_{hik(gj)}\, z_{hik}^{T} \hat{A}^{-1} \boldsymbol{Z} - w_{hik(gj)}\, z_{hik}^{T} \hat{A}^{-1} D_{gj} \hat{A}^{-1} \boldsymbol{Z}.
\end{aligned}
$$

Using this approximation of $\tilde{w}_{hik(gj)}$ in $\hat{Y}_{(gj)}$ and the definition of the jackknife weights gives

$$
\begin{aligned}
\hat{Y}_{(gj)} &\cong \sum\nolimits_{(hik)\in s} w_{hik(gj)} z_{hik}^{T} \hat{A}^{-1} \boldsymbol{Z} y_{hik} - \sum\nolimits_{(hik)\in s} w_{hik(gj)} z_{hik}^{T} \hat{A}^{-1} D_{gj} \hat{A}^{-1} \boldsymbol{Z} y_{hik} \\
&= \sum\nolimits_{(hik)\in s} w_{hik} z_{hik}^{T} \hat{A}^{-1} \boldsymbol{Z} y_{hik} - \sum\nolimits_{(hik)\in s} w_{hik(gj)} z_{hik}^{T} \hat{A}^{-1} D_{gj} \hat{A}^{-1} \boldsymbol{Z} y_{hik} \\
&\quad \frac{n_g}{n_g - 1}\left(\frac{1}{n_g} \sum\nolimits_{(gik)\in s} w_{gik} z_{gik}^{T} \hat{A}^{-1} \boldsymbol{Z} y_{gik} - \sum\nolimits_{(gjk)\in s} w_{gjk} z_{gjk}^{T} \hat{A}^{-1} \boldsymbol{Z} y_{gjk} \right) \qquad (3) \\
&= \hat{Y} - \beta^{T} D_{gj} \hat{A}^{-1} \boldsymbol{Z} + \frac{n_g}{n_g - 1}\left(\frac{1}{n_g} \sum\nolimits_{(gik)\in s} \tilde{w}_{gik} y_{gik} - \sum\nolimits_{(gjk)\in s} \tilde{w}_{gjk} y_{gjk} \right) + \\
&\quad \text{lower order terms}
\end{aligned}
$$

where $\beta = \hat{A}^{-1} \sum_{(hik)\in s} w_{hik} z_{hik} y_{hik}$. Note that in obtaining equation (3) we have used the fact that

$$
\begin{aligned}
\sum\nolimits_{(hik)\in s} w_{hik(gj)} z_{hik}^{T} \hat{A}^{-1} D_{gj} \hat{A}^{-1} \boldsymbol{Z} y_{hik} &= \sum\nolimits_{(hik)\in s} w_{hik} z_{hik}^{T} \hat{A}^{-1} D_{gj} \hat{A}^{-1} \boldsymbol{Z} y_{hik} + \text{lower order terms} \\
&= \beta^{T} D_{gj} \hat{A}^{-1} \boldsymbol{Z} + \text{lower order terms.}
\end{aligned}
$$

Now, consider the last two terms in equation (3). By the definition of D_{gj} we have

$$
\begin{aligned}
&-\beta^{T} D_{gj} \hat{A}^{-1} \boldsymbol{Z} + \frac{n_g}{n_g - 1}\left(\frac{1}{n_g} \sum\nolimits_{(gik)\in s} \tilde{w}_{gik} y_{gik} - \sum\nolimits_{(gjk)\in s} \tilde{w}_{gjk} y_{gjk} \right) \\
&= \frac{n_g}{n_g - 1}\left(\frac{1}{n_g} \sum\nolimits_{(gik)\in s} \left(\tilde{w}_{gik} y_{gik} - \beta^{T} w_{gik} z_{gik} z_{gik}^{T} \hat{A}^{-1} \boldsymbol{Z} \right) - \sum\nolimits_{(gjk)\in s} \left(\tilde{w}_{gjk} y_{gjk} - \beta^{T} w_{gjk} z_{gjk} z_{gjk}^{T} \hat{A}^{-1} \boldsymbol{Z} \right) \right) \\
&= \frac{n_g}{n_g - 1}\left(\frac{1}{n_g} \sum\nolimits_{(gik)\in s} \tilde{w}_{gik} (y_{gik} - \beta^{T} z_{gik}) - \sum\nolimits_{(gjk)\in s} \tilde{w}_{gjk} (y_{gjk} - \beta^{T} z_{gjk}) \right) \\
&= \frac{n_g}{n_g - 1} \left(\bar{e}_g - e_{gj} \right)
\end{aligned}
$$

where

$$
\begin{aligned}
e_{gj} &= \sum\nolimits_{(gjk)\in s} \tilde{w}_{gjk} e_{gjk}, \\
e_{gjk} &= y_{gjk} - \beta^{T} z_{gjk} \quad \text{and} \\
\bar{e}_g &= \frac{1}{n_g} \sum\nolimits_{i} e_{gi}.
\end{aligned}
$$

That is,

$$\hat{Y}_{(gj)} - \hat{Y} \cong \frac{n_g}{n_g - 1} (\bar{e}_g - e_{gj}),$$

and the jackknife variance estimator can be approximated by

$$v_{JL}(\hat{Y}) = \sum_h \frac{n_h - 1}{n_h} \sum_i (e_{hi} - \bar{e}_h)^2.$$

Now turning to the variance of $\hat{Y}'$, recall that

$$\hat{Y}' = \sum_{(hik)\in s_1} \tilde{w}_{hik} y'_{hik} + \sum_l \sum_{(hik)\in s_2} d_l \tilde{w}_{hik} y'_{hik} a_{hik} \delta_{hik,l}.$$

This estimator when the *(gj)-th* cluster has been deleted is

$$\hat{Y}'_{(gj)} = \sum_{(hik)\in s_1} \tilde{w}_{hik(gj)} y'_{hik} + \sum_l \sum_{(hik)\in s_2} d_{l(gj)} \tilde{w}_{hik(gj)} y'_{hik} a_{hik} \delta_{hik,l},$$

where $\tilde{w}_{hik(gj)}$ are the regression adjusted jackknife weights given in equation (2) and $d_{l(gj)}$ is the nonresponse adjustment in the *l-th* weighting class when the *(gj)-th* cluster is deleted and is defined as

$$d_{l(gj)} = \frac{\sum_{(hik)\in s_2} w_{hik(gj)} \delta_{hik,l}}{\sum_{(hik)\in s_2} w_{hik(gj)} a_{hik} \delta_{hik,l}}.$$

The jackknife variance estimator is then given by

$$v_J(\hat{Y}') = \sum_g \frac{n_g - 1}{n_g} \sum_j \left(\hat{Y}'_{(gj)} - \hat{Y}'\right)^2.$$

To obtain a linearized jackknife variance estimator, rewrite $\hat{Y}'_{(gj)}$ as

$$\begin{aligned}\hat{Y}'_{(gj)} &= \sum_{(hik)\in s_1} \tilde{w}_{hik(gj)} y'_{hik} + \sum_l \sum_{(hik)\in s_2} d_{l(gj)} \tilde{w}_{hik(gj)} y'_{hik} a_{hik} \delta_{hik,l} \\ &= \hat{Y}'_{1(gj)} + \hat{Y}'_{2(gj)}.\end{aligned}$$

We will work with the two parts of $\hat{Y}'_{(gj)}$ separately since the nonresponse adjustment must be included in $\hat{Y}'_{2(gj)}$. Working with $\hat{Y}'_{1(gj)}$ and using the expression for the regression adjusted jackknife weights, equation (2), gives

$$\hat{Y}^{/}_{1(gj)} \cong \hat{Y}^{/}_{1} + \frac{n_g}{n_g - 1}\left(\frac{1}{n_g}\sum\nolimits_{(gik)\in s_1} \tilde{w}_{gik}\, y^{/}_{gik} - \sum\nolimits_{(gjk)\in s_1} \tilde{w}_{gjk}\, y^{/}_{gjk}\right) - \sum\nolimits_{(hik)\in s_1} w_{hik}\, z_{hik}\, \hat{A}^{-1} D_{gj}\, \hat{A}^{-1} \boldsymbol{Z}\, y^{/}_{hik}$$

$$= \hat{Y}^{/}_{1} + \frac{n_g}{n_g - 1}\left(\frac{1}{n_g}\sum\nolimits_{(gik)\in s_1} \tilde{w}_{gik}\, y^{/}_{gik} - \sum\nolimits_{(gjk)\in s_1} \tilde{w}_{gjk}\, y^{/}_{gjk}\right) -$$

$$\frac{n_g}{n_g - 1}\left(\frac{1}{n_g}\sum\nolimits_{(gik)\in s} \tilde{w}_{gik}\, \beta_1^T z_{gik} - \sum\nolimits_{(gjk)\in s} \tilde{w}_{gjk}\, \beta_1^T z_{gjk}\right)$$

$$= \hat{Y}^{/}_{1} + \frac{n_g}{n_g - 1}\left(\frac{1}{n_g}\sum\nolimits_{(gik)\in s_1} \tilde{w}_{gik}\, e^{/}_{1,gik} - \sum\nolimits_{(gjk)\in s_1} \tilde{w}_{gjk}\, e^{/}_{1,gjk}\right) -$$

$$\frac{n_g}{n_g - 1}\left(\frac{1}{n_g}\sum\nolimits_{(gik)\in s_2} \tilde{w}_{gik}\, \beta_1^T z_{gik} - \sum\nolimits_{(gjk)\in s_2} \tilde{w}_{gjk}\, \beta_1^T z_{gjk}\right)$$

where

$$e^{/}_{1,gik} = y^{/}_{gik} - \beta_1^T z_{gik} \quad \text{and}$$

$$\beta_1 = \hat{A}^{-1} \sum\nolimits_{(hik)\in s_1} w_{hik}\, z_{hik}\, y^{/}_{hik}.$$

Turning to $\hat{Y}^{/}_{2(gj)}$, in order to simplify calculations we replace $d_{l(gj)}$ with d_l in $\hat{Y}^{/}_{2(gj)}$. This simplification will lead to an underestimation of the true variance but the effect should be negligible. Now using expression (2) in $\hat{Y}^{/}_{2(gj)}$ and after algebraic simplification we have,

$$\hat{Y}^{/}_{2(gj)} \cong \hat{Y}^{/}_{2} + \frac{n_g}{n_g - 1}\left(\frac{1}{n_g}\sum\nolimits_l \sum\nolimits_{(gik)\in s_2} d_l\, \tilde{w}_{gik}\, y^{/}_{gik}\, a_{gik}\, \delta_{gik,l} - \sum\nolimits_l \sum\nolimits_{(gjk)\in s_2} d_l\, \tilde{w}_{gjk}\, y^{/}_{gjk}\, a_{gjk}\, \delta_{gjk,l}\right) -$$

$$\sum\nolimits_l \sum\nolimits_{(hik)\in s_2} d_l\, w_{hik}\, z_{hik}\, \hat{A}^{-1} D_{gj}\, \hat{A}^{-1} \boldsymbol{Z}\, y^{/}_{hik}\, a_{hik}\, \delta_{hik,l}$$

$$= \hat{Y}^{/}_{2} + \frac{n_g}{n_g - 1}\left(\frac{1}{n_g}\sum\nolimits_l \sum\nolimits_{(gik)\in s_2} \tilde{w}_{gik}\, e^{/}_{2,gik,l} - \sum\nolimits_l \sum\nolimits_{(gjk)\in s_2} \tilde{w}_{gjk}\, e_{2,gjk,l}\right) -$$

$$\frac{n_g}{n_g - 1}\left(\frac{1}{n_g}\sum\nolimits_{(gik)\in s_1} \tilde{w}_{gik}\, \beta_2^T z_{gik} - \sum\nolimits_{(gjk)\in s_1} \tilde{w}_{gjk}\, \beta_2^T z_{gjk}\right)$$

where

$$e_{2,gjk,l} = \left(d_l\, y_{gjk}\, a_{gjk} - \beta_2^T z_{gjk}\right)\delta_{gjk,l} \quad \text{for } (gjk)\in s_2 \text{ and}$$

$$\beta_2 = \hat{A}^{-1} \sum\nolimits_l \sum\nolimits_{(hik)\in s_2} d_l\, w_{hik}\, z_{hik}\, y^{/}_{hik}\, a_{hik}\, \delta_{hik,l}.$$

Combining the expressions for $\hat{Y}^{/}_{1(gj)}$ and $\hat{Y}^{/}_{2(gj)}$ gives

$$
\begin{aligned}
\hat{Y}^{/}_{(gj)} &= \hat{Y}^{/}_{1(gj)} + \hat{Y}^{/}_{2(gj)} \\
&\cong \hat{Y}^{/} + \frac{n_g}{n_g-1}\left(\frac{1}{n_g}\sum_{(gik)\in s_1} \tilde{w}_{gik}\left(e^{/}_{1,gik} - \beta_2^T z_{gik}\right) - \sum_{(gjk)\in s_1} \tilde{w}_{gjk}\left(e^{/}_{1,gjk} - \beta_2^T z_{gjk}\right)\right) + \\
&\quad \frac{n_g}{n_g-1}\left(\frac{1}{n_g}\sum_l \sum_{(gik)\in s_2} \tilde{w}_{gik}\left(e^{/}_{2,gik,l} - \beta_1^T z_{gik}\,\delta_{gik,l}\right) - \sum_l \sum_{(gjk)\in s_2} \tilde{w}_{gjk}\left(e^{/}_{2,gjk,l} - \beta_1^T z_{gjk}\,\delta_{gjk,l}\right)\right) \\
&= \hat{Y}^{/} + \frac{n_g}{n_g-1}\left(\frac{1}{n_g}\sum_{(gik)\in s_1} \tilde{w}_{gik}\, e^{*}_{gik} - \sum_{(gjk)\in s_1} \tilde{w}_{gjk}\, e^{*}_{gjk}\right) + \\
&\quad \frac{n_g}{n_g-1}\left(\frac{1}{n_g}\sum_l \sum_{(gik)\in s_2} \tilde{w}_{gik}\, e^{*}_{gik,l} - \sum_l \sum_{(gjk)\in s_2} \tilde{w}_{gjk}\, e^{*}_{gjk,l}\right)
\end{aligned}
$$

where

$$
\begin{aligned}
e^{*}_{gjk} &= y^{/}_{gjk} - \beta^T z_{gjk}, \\
e^{*}_{gjk,l} &= \left(d_l\, y^{/}_{gjk} - \beta^T z_{gjk}\right)\delta_{gjk,l} \quad \text{and} \\
\beta &= \beta_1 + \beta_2 \\
&= \hat{A}^{-1}\left(\sum_{(hik)\in s_1} w_{hik}\, z_{hik}\, y^{/}_{hik} + \sum_l \sum_{(hik)\in s_2} d_l\, w_{hik}\, z_{hik}\, y^{/}_{hik}\, a_{hik}\, \delta_{hik,l}\right).
\end{aligned}
$$

Letting

$$
\begin{aligned}
e^{*}_{gj} &= \sum_{(gjk)\in s_1} \tilde{w}_{gjk}\, e^{*}_{gjk} + \sum_l \sum_{(gjk)\in s_2} \tilde{w}_{gjk}\, e^{*}_{gjk,l} \quad \text{and} \\
\bar{e}^{*}_{g} &= \frac{1}{n_g}\sum_j e^{*}_{gj},
\end{aligned}
$$

we have

$$
\hat{Y}^{/}_{(gj)} - \hat{Y}^{/} \cong \frac{n_g}{n_g-1}\left(\bar{e}^{*}_{g} - e^{*}_{gj}\right),
$$

and the linearized jackknife variance estimator for $\hat{Y}^{/}$ is

$$
v_{JL}(\hat{Y}^{/}) = \sum_g \frac{n_g-1}{n_g}\sum_j \left(e^{*}_{gj} - \bar{e}^{*}_{g}\right)^2.
$$

Recall that we wish to estimate

$$V(\tilde{UR}-UR) \cong \frac{1}{\hat{X}^2} V(\hat{Y}'-\hat{Y})$$

$$= \frac{1}{\hat{X}^2}\left(V(\hat{Y}') + V(\hat{Y}) - 2\,COV(\hat{Y}',\hat{Y})\right).$$

A linearized jackknife variance estimator of $COV(\hat{Y}',\hat{Y})$ is

$$cov_{JL}(\hat{Y}',\hat{Y}) = \sum_g \frac{n_g-1}{n_g} \sum_j \left(e_{gj}-\bar{e}_g\right)\left(e^*_{gj}-\bar{e}^*_g\right),$$

and a linearized jackknife variance estimator for $V(\tilde{UR}-UR)$ is

$$v_{JL}(\tilde{UR}-UR) = \frac{1}{\hat{X}^2}\left(v_{JL}(\hat{Y}') + v_{JL}(\hat{Y}) - 2\,cov_{JL}(\hat{Y}',\hat{Y})\right)$$

$$= \sum_g \frac{n_g}{n_g-1} \sum_j \left((e_{gj}-\bar{e}_g)^2 + (e^*_{gj}-\bar{e}^*_g)^2 - 2\,(e_{gj}-\bar{e}_g)(e^*_{gj}-\bar{e}^*_g)\right).$$