COVERAGE ISSUES IN THE CANADIAN LABOUR FORCE SURVEY

Sander Post[1], Charlene Walker

ABSTRACT

The Canadian Labour Force Survey (LFS) is one of Statistics Canada's most important surveys. It collects data from approximately 52000 households per month about their labour force status, nature of their work or reason for not working, as well as demographic data. Coverage error is a measure of data quality important to any survey. One of the key measures of coverage error in the Canadian Labour Force Survey is the percentage difference between Census of Population estimates and the Canadian Labour Force Survey (LFS) population counts; it is called slippage. A negative value indicates that the LFS has a problem of overcoverage, while a positive value indicates the LFS has an undercoverage problem. In general, slippage is positive, thus the LFS consistently misses people who should be enumerated. The purpose of this study was to determine why slippage is increasing and what can be done to stop it. The study was conducted in two stages. The first stage was a historical review of the projects that have studied and tried to control slippage in the LFS as well as the operational changes that have been implemented over time. The second stage was an analysis of factors such as vacancy rates, nonresponse, demographics, urban/rural status and their impact upon the slippage rate.

KEY WORDS: Slippage, Coverage, Nonresponse

1. INTRODUCTION

1.1 What is Slippage?

Slippage is a measure of survey coverage error. It is defined as the percent difference between the Census population projections and the LFS population counts. A negative slippage rate indicates LFS overcoverage, while positive slippage indicates undercoverage. In general, slippage is positive, thus the LFS consistently misses people who should be enumerated.

Slippage occurs when dwellings or persons within households that should have been covered by the survey are missed. Dwellings may be missed because some areas are difficult to map, because they are erroneously classified as vacant, or because several dwellings may be classified as one. Persons are sometimes missed in listed dwellings because the concept of "usual place of residence" is misunderstood or incorrectly applied, or because information was withheld.

Slippage becomes a problem when the characteristics of missed dwellings and persons are different from those that are enumerated. As a result, slippage can introduce bias into the estimates. Also, because the realized sample size is lower than the planned sample size in certain groups, slippage can increase the variance of estimates.

Due to the potential problems associated with it, slippage has been a concern for a very long time. Slippage is of special concern currently because it has been increasing for years and is now consistently around 10%. Slippage differs by region, reaching as high as 15% in certain provinces.

---

[1] Household Survey Methods Division, Statistics Canada, 16th floor R.H. Coats Building, Tunney's Pasture, Ottawa, Ontario, Canada, K1A 0T6, e-mail: sander.post@statcan.ca

Figure 1 presents the slippage rate for Canada since January 1976. The slippage rate is based on 1996 census projections. The series is characterized by a steadily increasing trend interrupted by sudden sharp drops between June 1976 and March 1977, June 1987 and February 1988, and July 1994 to March 1995.

The drops can all be attributed to operational effects:

- A mini redesign of self-representing unit strata in 1977 for areas of heavy growth.

- Several changes to the Interviewers Manual in April 1987.

- The phase in of the new sample design from October 1994 to March 1995.

Figure 1.

CANADA - Slippage Rate
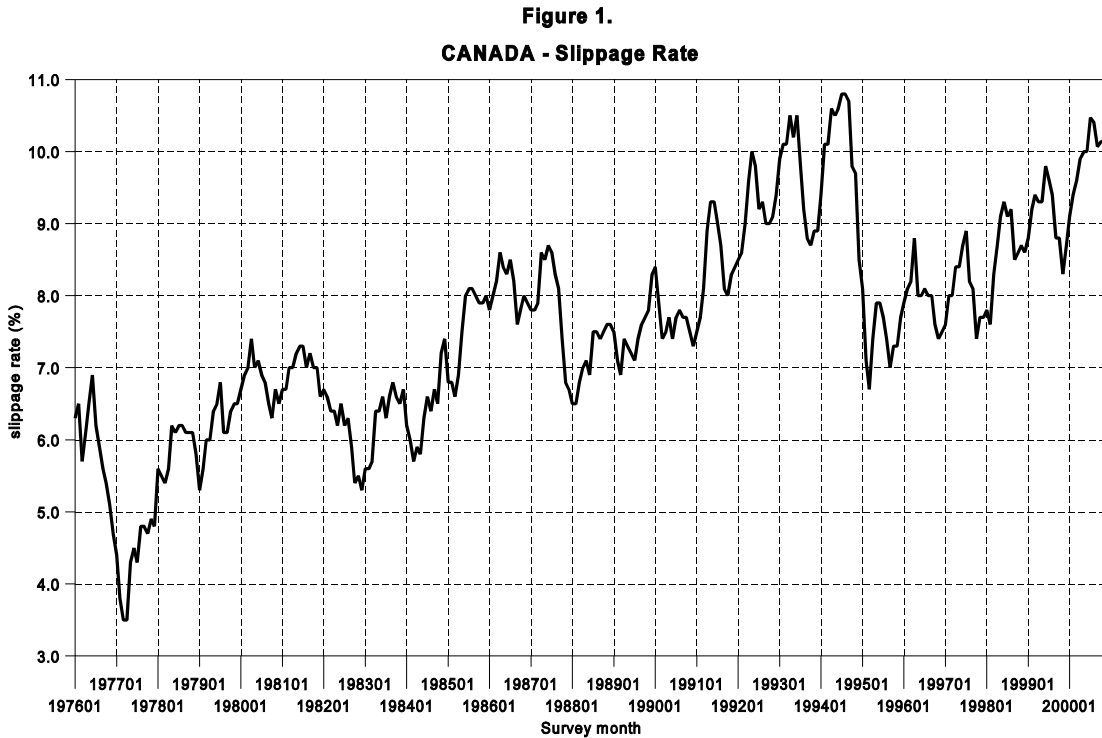
Two questions that need to be addressed are: "Why is slippage increasing?" and "What can be done to stop the increase and, if possible, decrease it?" To answer these questions, a two-part study was undertaken. The first stage was a historical review of the projects that have studied slippage in the LFS as well as the operational changes that have been implemented over time. The second stage was an analysis of various factors and their impact on the slippage rate.

2. HISTORICAL REVIEW

To avoid duplicating work, the first step in this project was a historical review of both previous studies of slippage and operational changes in the LFS over time. We restrict ourselves from the period of 1976 to the present.

2.1 Historical Studies

Many studies have examined the slippage issue in the past. They can be grouped into several broad categories: listing checks, vacancy checks, studies into roster coverage, examination of manuals and procedures, evaluation of Census projections and analytical studies.

While many studies have been done into slippage historically, there is no room in this document to review them all. Some of the more relevant studies are listed below.

- Several studies have concluded that LFS nonrespondent households were generally smaller than respondent households. Several studies concluded that LFS nonrespondent households were 15% smaller than the average size for all reinterview households (e.g. Earwaker 1981 and Ghangurde *et al* 1979). Weighting therefore creates a positive bias in the average household size (estimated as 0.7%). It was concluded that the within household undercoverage rate was twice as high as the dwelling undercoverage rate. A micromatch between LFS records and the 1991 census showed average respondent household size to be 2.6 persons, and average nonrespondent household size to be 2.1 persons, or about 19% smaller (Allard *et al* 1998).

- In 1986 and 1987 a large scale slippage study was conducted involving analytical studies, listing checks, vacancy checks and examination of both interviewer procedures and Census estimates. The first component of the study examined final and sub-weighted estimates in areas with high slippage. Several operational studies were conducted: a special listing check program was undertaken in fringe areas of large cities between October 1986 and March 1987, a vacancy check program was undertaken in July 1987 for one quarter of all assignments (this resulted in a 7.4% conversion rate of vacants to occupied households) and a rural listing check was done in July 1987. The post-Censal estimates were examined at various geographic levels and interviewer procedures and the interviewer's manual was examined in detail. The studies resulted in many changes to concepts and procedures, many of which were aimed at harmonization with Census concepts (Switzer 1987).

- A vacancy check is a visual check of selected dwellings coded as vacant in the previous month. An analysis was done of data from monthly vacancy checks to assess the extent of vacant misclassification. It was found that 2.7% of the dwellings checked had been incorrectly coded as vacant (on average over the months checked) and of these, 81% were given a response code when checked. As a part of the same study, the relationships between vacancy, nonresponse and slippage rates at the national level were examined. A relatively strong inverse linear relationship was found between nonresponse and slippage, likely because nonresponding households are smaller than responding households (Wysocki 1993).

- In 1993, the error in the post-Censal estimates of the LFS target population was estimated to be between -0.4% and 1.0% resulting in a slippage error between -8% (underestimation) and 18% (overestimation). It was also noted that following the decision in January 1993 to adjust population estimates to account for Census coverage error LFS estimates would be adjusted accordingly. The authors concluded that "Slippage will increase as a consequence of this change particularly for those individuals like young males who are undercovered the most by the Census" (Clark *et al* 1993).

- In 1996, a vacancy check of ALL dwellings coded as vacant in the previous month was conducted in the prairies. A listing quality check was conducted at the same time. In the three provinces, 86% of the vacant dwellings were confirmed vacant while 6% were occupied and 8% could not be confirmed. As a result of the cluster check, 45 of the clusters had changes made with an average of 2.9 changes, additions and deletions per checked cluster.

- A micromatch was conducted between May 1991 LFS and Census data to study the quality of LFS data. A key observation was that nonrespondent households in the LFS tend to be smaller than respondents. Of 1393 matched dwellings coded vacant in the LFS, 813 were occupied in the Census. Of 753 vacants in the Census,

173 were occupied in the LFS. This indicated that LFS vacancy rates were higher than the true vacancy rate, and this contributes to slippage (Allard et al 1998). Similarly, after the 1996 Census, another micromatch was conducted. Most results were similar to the previous study, and it confirmed key findings, such as nonrespondent households being smaller than respondent households (Tolusso 2001).

- A commonly held belief is that students are the cause for the traditionally high slippage rates of 20-24 and 25-29 year-olds. Therefore, in 1998 the slippage rates for students and nonstudents were compared. When comparing slippage rates between students and nonstudents while controlling for other variables, student status was not significant. High slippage rate respondents were found to live in low-rise apartments, in urban areas, and in large cities (Gardner *et al* 1998).

- In August 1999, a field validation was conducted by the Edmonton Regional Office. Primary results showed that on the whole, the interviewers do an excellent job, both in the actual interviews and in checking vacancy - only three dwellings coded vacant, out of 1763, were discovered to have been occupied during the reference week. Results from a cluster check showed dwelling undercoverage of varying degrees in all areas. In fact, based on this cluster check, the dwelling undercoverage in the prairie region was about 1.5%. The slippage in the Edmonton R.O. in July 1999 was 4.6%, which indicates that about a third of the slippage was due to missed dwellings (Prairie Region 1999).

Common conclusions emerge from the many historical studies:

- Quality control programs provide an important benefit in improving coverage. Reinterview programs give a more accurate measure of interviewer quality than observation programs. Alert field work and verification checks can produce substantial reductions in slippage. Persons missed within a listed dwelling contribute more to the slippage rate than missed dwellings.

- Miscoding dwellings as vacant significantly influences slippage rates. Studies conducted at different times and in different regions have shown misclassification rates as high as six percent. The vacancy rate for the apartment frame is well above the rate for the nonapartment frame. The number of incorrect vacants in a rotation group generally increases during the six months that the group is in sample.

- Nonrespondent households are generally smaller than respondent households. This is not an encouraging finding, as this means that with a one hundred percent response rate, we would have lower (precalibrated) estimates and higher slippage.

However, while all these factors contribute to slippage, even the combination of all of them cannot explain the slippage rate and its long term increase.

2.2 Operational Changes

Operational changes, particularly those involving quality control procedures, can have a large impact on slippage. Minor changes are an ongoing affair in a survey like the LFS, so not all changes will be mentioned in the list below.

- There was a dramatic increase in telephone interviewing in the mid 1980s. This probably had a large impact on listing maintenance since it was no longer necessary for the interviewers to do a large volume of interviews in the field. The increase in telephone interviewing may also contribute to slippage since interviewers lose the opportunity to see individuals who reside in the household.

- In 1987, several changes were made to the manual to harmonize concepts between the LFS and the Census, to ensure consistency with the rules on the use of noninterview codes and to eliminate ambiguities in the noninterview instructions. Specifically, persons with two dwellings in Canada were now counted at the one

they occupied for the greater part of the year. Seasonal dwellings were now listed and the S code (seasonal dwelling) was dropped. The effect of these changes and instructions was a decrease of approximately two percentage points in the slippage rate between July 1987 and December 1987.

- In January 1993, it was decided to adjust population estimates to account for Census coverage error. The post-Censal estimates used by the LFS were adjusted accordingly. This was expected to increase slippage particularly for individuals who are undercovered most by the Census, e.g. young males.

- A major sample redesign was done in the mid 1990s. Large changes were made to the sample, the method for the nonresponse adjustment was changed and a new questionnaire was phased in late in 1996, one goal of which was to improve within household coverage (Bench *et al* 1994). This redesign resulted in an initial decrease in the slippage rate. However, the slippage started increasing shortly after implementation.

- Quality assurance programs, such as the reinterview program and interviewer observation, have been scaled back or cut over the years. Interviewer observation was decreased from once a year to once every two years in 1993/1994. Group interviewer training used to be done twice a year (five days for seniors, three days for interviewers). However, in 1992/93 the frequency was reduced to one group training session per year (three days for seniors, two days for interviewers).

- Interviewer workload has increased over time with the addition of more and more supplementary surveys and larger clusters. In addition, demands increase or stay constant with decreasing budgets. In fact, in 94/95 seniors were given a small assignment of cases for the first time.

- Computer-assisted interviewing was implemented in November 1993. This resulted in a vast number of changes to field and head office procedures. A new CAI application was phased in from September to November 1999. There were significant technical problems during the phase in, resulting in markedly higher nonresponse. In November 1999, the first month of full implementation, the total nonresponse was 13% including nonresponse due to technical problems of 6.6%. These problems have decreased over time. Preliminary analysis shows that interview times significantly increased under the new system but have since decreased due to increased interviewer familiarity and more powerful laptops.

- Centralized CATI was implemented in 2000. The first rotation group was brought into centralized CATI in June 2000, with full implementation reached in September 2000. Initial results show that nonresponse from centralized CATI is higher than under the previous system. There is some concern that with fewer interviewers in the field, less new growth will be discovered.

It seems that quality control procedures are one of the first activities to be reduced or eliminated when budgets are tight. However, the decision to cut quality control procedures has most likely had a detrimental effect on the slippage.

3. ANALYSIS

This analysis covers the time period from January 1976 to June 2000. Issues covered are changes in demographics over time, the relationships between slippage and other key variables measured in the LFS, such as nonresponse and vacancy, and the effects of operational changes. In addition, issues related to the design and weighting of the LFS and census projections are covered.

3.1   Vacancy Rates

Analysis of vacancy and slippage shows a strong correlation between the two variables. One concern that has been noted is that misclassified vacants may contribute to slippage. To verify whether the vacancy rate in the LFS is accurate, independent sources of the Canadian vacancy rate need to be found.

The Canadian Community Health Survey (CCHS) is a new survey using a combination of an area frame survey and RDD. The first two months of data collection were September and October 2000. For the LFS, the vacancy rates at the national level were about 2.5% lower than those for the CCHS. At the provincial level, differences between vacancy rates ranged from 5% lower in the LFS than the CCHS in Québec, to 1% higher in the LFS in British Columbia.

Vacancy data from the Canadian Mortgage and Housing Corporation (CMHC) were examined, but differences in definitions mean the rates are not useful for comparison.

At the national level, 13% of vacant dwellings (918 dwellings, or about 1.5% of the sample in October 1999) are occupied by people not to be interviewed. Although weights are not available for these dwellings, they are distributed across Canada, so by using the average household weight of approximately 200, we obtain approximately 180 000 dwellings, or 450 000 people. Further examination of these dwellings revealed that a number of them are occupied by students, which were properly excluded by current rules. It should also be noted that the 918 such dwellings found in October 1999 appears to be higher than average, and that 600 to 800 is the typical range. Nonetheless, even at the low end of the range, this potentially represents 300 000 people.

3.2 Nonresponse Rates

It has been shown several times that the household size is smaller for nonresponding households. This is likely because the smaller the household, the less likely it is that there is someone at home for the interviewer to contact. So when the nonresponse rate increases, the average household size of respondents also increases. Since the weight of respondents is adjusted for nonresponse, a higher nonresponse rate will decrease slippage because the actual number of persons in the LFS sample is overestimated.

This effect is noticeable, although hard to see without removing the seasonality intrinsic to nonresponse. When 12 month moving averages are used, the slippage rate shows an inverse relationship to the nonresponse rate. An important result about the relationship of nonresponse and slippage concerning subgroups smaller than nonresponse adjustment level groups was developed by Gambino (2000). It states that "The slippage for Group G is overstated if and only if the nonresponse rate in G is higher than the overall nonresponse rate." Nonresponse adjustment is done at a certain level. For smaller groups, such as age/sex crossings, we calculate slippage, but the results are biased if the nonresponse in these groups does not match the nonresponse rate in the adjustment level group. One example given in this paper is, in a group with a true slippage rate of 10%, a nonresponse level of 10%, and a nonresponse adjustment of 5%, we will publish a slippage rate of approximately 14.7%.

3.3 Household Size

Multivariate analyses of slippage reveal that the best explanatory variable for slippage is household size. There is a strong negative correlation between slippage and household size. There may be a causal effect, since, if on average the same number of people are missed per household, then they will have a larger impact on the slippage rate if the household size is smaller.

The average household size of LFS respondents has been steadily decreasing since 1976. This trend has been confirmed with auxiliary data from the Census and from the CANSIM database. For unweighted data, the average household size (including children) has fallen from 3.24 to 2.53. The final weighted household sizes have slightly smaller values, falling from 3.08 to 2.51.

One factor contributing to the slippage is the difference in size between respondent and nonrespondent households. Several studies have shown that nonrespondent households are smaller than respondent households. The conclusion to be reached is that nonresponse lowers the slippage rate, as it leads to overestimation of the number of people living in sampled households. The two studies show that there is not a significant difference between the size of respondent and nonrespondent households over time, as both seem to be getting smaller at the same time.

Therefore, an analysis that includes the relationships among the unweighted, the subweighted and the final weighted series of average household size over time is more worthwhile. The differences between the unweighted and both of the weighted series have decreased with time. This is mostly a reflection of changes in both society and the sample design. There is a larger difference early in the series as household sizes in the smaller provinces tended to be significantly larger than those in larger provinces. Since smaller provinces had higher sampling rates, the unweighted sizes were significantly larger than the weighted sizes. Over time, the household size has been decreasing in all provinces, but has been decreasing faster in the smaller provinces. As a result, higher sampling rates in the small provinces are having less of an effect on the unweighted results.

The decrease in household size can have definite results on slippage, even if all else remains the same. Since the difference between the final weight and the subweight is not changing very much, we are missing, on average, the same fraction of a person per household. Yet, as the households are getting smaller, this fraction of a person makes up a larger and larger percentage of the average household. Therefore, there is a connection between decreasing household size and the slippage rate.

3.4 Demography

Since 1976, there have been changes in the demographics of Canadian society. Due to increased life spans and decreased birth rates, the population is getting older. Although there are slight differences in the actual percentages, the patterns are the same for males and females. It should be noted that the three age groups between 15 and 30 have not only decreased as proportions, they have actually decreased in absolute size between 1976 and 2000.

Since slippage varies by age group, changes in the makeup of society affect the slippage rate. Of course, slippage rates by age have also changed over time, so it is worthwhile to look at the contribution of each age/sex group to the slippage rate over time to see which groups have the largest impact.

For 15-19 and 20-24 olds, the contribution to slippage has decreased over time. For 25-29 year old males, the contribution has been fairly constant over time. For 25-29 year old females, the contribution has increased slightly due to a major increase in slippage rates for that group (from below 5% prior to 1978 to an average of more than 15% in 2000). The increase in slippage among females under the age of 30 is likely due to increasing presence in the labour force and a corresponding increase in young females living in smaller households, which are most often missed. For males and females, the 30-39 and 40-54 groups show definite increases in both their actual slippage rates and their percentage contribution to the total slippage rate since 1976. Males 30-39 have gone from about 10% of the total to about 15%, while for females the rate has gone from below 5% percent to just under 10% in 2000. Results for the 40-54 group are similar. For the 55+ age group, although both the actual population in this age group and the slippage rate have increased, the contribution has not changed significantly.

Changing demographics have a notable influence on the slippage rate. While the slippage rate among those under 30 is increasing, the contribution to the overall slippage rate is decreasing. The decreasing proportion of younger people in society means that more and more of the slippage is due to people over 30 years of age. This trend is likely to continue for some time, with the aging of the population. Efforts to decrease the slippage among the older age groups would be of more use in decreasing the overall slippage rate than would focussing efforts on younger people. However, this ignores the bias caused by the higher slippage rates among the younger age groups.

3.5 Interviewer Workload

Average interviewer workload has been increasing constantly since 1976. All else being equal, increasing workload would likely have a deleterious effect on slippage, as dwellings or persons would be more likely to be missed as interviewers would spend less time on each case.

However, increasing use of technology, both the increasing use of telephone interviews, and the use of laptop computers, has made the interviewer more productive, particularly as much less time is spent travelling between cases. However, the increasing use of telephone interviews means the interviewer spends less time in the field, and is therefore less likely to discover new dwellings.

One last concern regarding interviewer workload is the increasing number of supplementary surveys. There have been an increasing number of supplementary surveys to the LFS during the study period, and this has increased workloads more than just the increase in average number of dwellings surveyed would indicate.

3.6 Post-Censal Population Projections

Slippage is measured as the difference between census projections and LFS estimates. Generally, when analysing slippage, the assumption is made that the census projections are correct and discrepancies are indicative of problems in the LFS. However, as can be expected, there are problems with both surveys. The LFS has recently changed from 1991 to 1996 Census projections. This resulted in lower slippage rates at the national level and in every province except B.C.

Also of interest are the Census' coverage studies. The net undercoverage in the 1996 Census is 2.45%(±0.10%) at the Canada level. Studies show that slippage is typically highest in British Columbia, and Census undercoverage is also highest in British Columbia, at 3.68%(±0.25%). This also applies to age groups: males 20-24 and 25-29 have the highest slippage rates in the LFS and have net undercoverage in the Census of 7.14%(±0.39%) and 7.08%(±0.40%) respectively.

4. CONCLUSIONS

Common conclusions emerge from the many historical studies:

- Quality control programs (reinterview, interviewer observation, vacancy and cluster checks) provide an important benefit in improving coverage.

- Miscoding dwellings as vacant significantly influences slippage rates. The vacancy rate for the apartment frame is well above the rate for the nonapartment frame. The number of incorrect vacants in a rotation group generally increases during the six months that the group is in sample.

- Nonrespondent households are generally smaller than respondent households.

- Despite their positive effects on coverage, quality control procedures are one of the first activities to be reduced or eliminated when budgets are tight.

- The inclusion of seasonal dwellings had a strong impact on the slippage rate, decreasing the observed slippage by more than a percentage point.

- The sample appears to degrade over time, as the slippage rate is positively correlated with the length of time since the last redesign.

Analysis of slippage and various factors has revealed:

- Slippage and nonresponse behave inversely.

- The vacancy rate is closely related to slippage. Misclassified vacants contribute to slippage, but in the absence of quality control programs it is not known whether this is a serious problem.

- It is difficult to measure the change in interviewer workload during the study period. Interviewers are surveying more households, and asking more questions as the LFS questionnaire has increased in length over time, and there are more supplementary surveys. It is not known how much the use of computers or telephones has compensated for this increase in workload, but higher workloads could contribute to more missed individuals and dwellings.

- There is a strong negative correlation between slippage and household size. There may be a causal effect, since, if on average the same number of people are missed per household, then they will have a larger impact on the slippage rate if the household size is smaller.

REFERENCES

Allard, B, I. Lévesque and J. Dufour (1998), *Evaluating Labour Force Survey Data Using the 1991 Census of Population*, unpublished report, Ottawa, Canada: Statistics Canada

Bench, J., C. Clark, G. Diotte, J. Dufour, R. Généreux, P. Rappak, M. Wysocki (1994), *LFS Questionnaire Redesign Household Membership Questions*, unpublished report, Ottawa, Canada: Statistics Canada

Clark, C., B. Kennedy, M. Wysocki (1993), *Coverage error in the Canadian Labour Force Survey*, unpublished report, Ottawa, Canada: Statistics Canada

Earwaker, S. (1981), *Slippage Estimates and Coverage Analysis*, unpublished report, Ottawa, Canada: Statistics Canada

Gambino, J. (2000), *The Effect of Nonresponse Adjustment on Slippage Calculations*, unpublished report, Ottawa, Canada: Statistics Canada

Gardner, E. and C. Walker (1998), *Excerpt from The Effect of Student Status on the Labour Force Survey Slippage Rate*, unpublished report, Ottawa, Canada: Statistics Canada

Ghangurde, P and A. Gower (1979), *Reinterviews as a Supervisory Tool in the Field*, unpublished report, Ottawa, Canada: Statistics Canada

Prairie Region LFS Field Validation - August 1999, unpublished report, Ottawa, Canada: Statistics Canada

Switzer, K. (1987), *Reducing and Monitoring Slippage,* unpublished report, Ottawa, Canada: Statistics Canada

Tolusso, S. (2001), *Evaluating the Quality of the May 1996 Labour Force Survey Data Using the 1996 Census of Population*, unpublished report, Ottawa, Canada: Statistics Canada

Wysocki, M. (1993), *Misclassification of Vacants,*, unpublished report, Ottawa, Canada: Statistics Canada