Catalogue No. 98-18

THE SEAM EFFECT IN THE SURVEY OF LABOUR AND INCOME DYNAMICS

Product registration Number 75F0002M

November 1998

Cathy Cotton, Income Statistics Division, Statistics Canada

Phil Giles, Income Statistics Division, Statistics Canada

The Income and Labour Dynamics Working Paper Series is intended to document detailed studies and important decisions for the Income and Labour Dynamics program. It is a continuation of the SLID Research Paper Series. These working papers are available in English and French. To obtain a summary description of available documents or to obtain a copy of any, please contact the Client Services Unit, at 7-B5 Jean-Talon Building, Statistics Canada, Ottawa, Ontario, CANADA K1A 0T6; by INTERNET: (DYNAMICS@STATCAN.CA); by telephone (613) 951-7355 or toll-free 1-888-297-7355; or by fax (613) 951-3012.

EXECUTIVE SUMMARY

Labour market transitions are one of the most important content areas in SLID. These data can be difficult to collect as they require precision in the reporting of dates and activities. Dependent interviewing has reduced seam problems. There were many more December job ends and January job starts during the reference period for which 30% of the sample did not receive dependent interviewing.

An awareness of the seam problem is crucial. Processing decisions can contribute to the number of changes at the seam, so indicators should be provided to enable researchers to avoid some types of contributions to changes at the seam.

Overall, the SLID transitions from one labour force status to another show reasonable patterns, though the number of December transitions is high in some cases. More research is needed to better understand the sources of these transitions at the seam and to reduce their impact.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Survey of Labour and Income Dynamics (SLID)

Objectives of SLID

The Survey of Labour and Income Dynamics (SLID) is a longitudinal household survey conducted by Statistics Canada. The objective of SLID is to capture changes in economic well-being of Canadian individuals and families over time and to understand the determinants of their well-being.

Survey Design

The target population of SLID is all persons living in Canada, excluding persons living in the territories, institutions, Indian reserves and military barracks. Individuals originally selected for the survey are interviewed for six years to collect information about their labour market experiences, income and family circumstances. In order to obtain complete information on families and to obtain cross-sectional data, persons who live with the original respondents at any time during the six years are also interviewed during the time of cohabitation.

The first reference year of the survey was 1993. A second six-year panel of respondents was introduced for 1996, halfway through the life span of the first panel. When the first panel ends, a third one will begin for reference year 1999. This pattern of rotating, overlapping panels will continue, with a new panel being selected every three years. The size of each panel is approximately 15,000 households, comprising about 31,000 adults aged 16 years and older.

Respondents are interviewed each January to collect labour information, educational activities and family relationships over the past calendar year. Conceptually, information on respondents' income could also be collected in January, but in fact collection of income data is deferred until May to take advantage of income tax time when respondents are more familiar with their financial records. An income interview is not required if the respondent gives Statistics Canada permission to use his or her tax information from Revenue Canada for purposes of this survey. All collection is done using computer-assisted telephone interviewing.

The labour and income questions are intended for people 16 years of age and over, but children who were present in the original households are also followed for the full six years. When they turn 16 they join the other longitudinal respondents in completing both labour and income interviews.

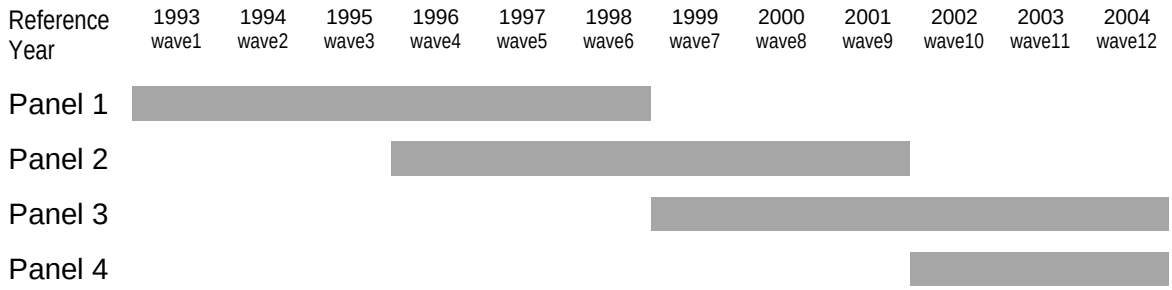

Figure 1: The sample design of the Survey of Labour and Income Dynamics (SLID)

Content

The survey content covers a wide range of topics, providing the scope necessary for longitudinal studies. The unique content mix also permits important cross-sectional analyses that were never before possible using Canadian data sets. All data are collected separately for every person in the household (age restrictions apply to certain question groups). Data on

household relationships are collected to allow the assignment of individuals into families.

The content can be described in a hierarchical fashion, under four broad themes: labour; income and wealth; education; and, personal characteristics. Each of these themes contains many topic areas as follows.

A. Labour

- Nature and pattern of labour market activities – summary information at the person level such as labour force status, spells of employment and unemployment, job changing and multiple job holding
- Work experience - number of years of lifetime work experience, in full-year full-time equivalents
- Characteristics of jobless spells - for periods of time without a job, information on job search, desire for work, and reasons for not looking
- Job characteristics - detailed information on each job held, with information updated annually. Content includes dates and durations; class of worker; work schedule; occupation group; supervisory responsibilities; wages and benefits.
- Characteristics of job absences - excluding paid vacation, for each job absence of at least seven days: dates and duration; reason for absence; and whether the person was paid by the employer during the absence.
- Employer attributes - characteristics of each job which are not particular to the employee: industry; firm size; public or private sector

B. Income and wealth

- Personal income - annual information on about 25 income sources; total income; income taxes paid

- Receipt of compensation - for each of three government transfer programs (Employment Insurance, Workers' Compensation, and Social Assistance), whether benefits were received, by month
- Assets and debts - no data are currently being collected, but we eventually expect information on about 20 categories

C. Education

- Educational activity - whether enrolled in a credit program, by month; type of institution; full-time or part-time student; degrees and certificates received. On-the-job training is not included.
- Educational attainment - total years of schooling; highest level of education; major field of study.

D. Personal characteristics

- Demographics - basic demographics (age, sex, marital status); some marital history; major activity
- Ethno-cultural - mother tongue; ethnic origin; country of birth; immigration status; parents' schooling and place of birth
- Activity limitations - annual information on activity limitations and their impact on working
- Information on person's children - number of children born and raised; age when first child born
- Geography and geographic mobility - detailed coding of place of residence; moves including reason
- Household, economic and census family information - size and type; dwelling type and tenure; annual summary information of personal labour and income characteristics; family events (marriage, separation, death, birth)

1.2 The Seam Effect

Definition

The seam effect refers to a disproportionately high occurrence of transitions at the "seam" between two reference periods. This is caused by response errors, either misplacing the beginning or end of a spell, or completely forgetting a spell. Clearly it is important to reduce these errors as much as possible so that the measurement of transitions from one state to another is not seriously distorted.

SLID collects data annually using the previous calendar year as the reference period, so the seam occurs between December 31 and January 1. As the survey attempts to link various spells from year to year, the seam effect is an important aspect of data quality.

Lemaître (1992) compared the effects of two feedback techniques and concluded that dependent interviewing is the only reliable method of collecting data that are consistent over time.

2. SLID AND THE SEAM EFFECT

2.1 Factors that Have an Impact on the Seam Effect

Dependent Interviewing

SLID uses computer assisted interviewing exclusively. This facilitates the use of dependent interviewing, i.e. providing respondents with their responses as reported one year earlier. For example in the January 1999 interview, an interviewer would say to a respondent, "Based on our interview of a year ago, you were working for company XYZ around the beginning of January 1998.

Is this correct?" At this point, the respondent can either confirm or deny the information. SLID feeds back information of six main items:

- employer name
- occupation
- an absence in progress at the end of the previous reference year
- a spell of job search in progress at the end of the previous reference year
- school attendance
- receipt of income from social programs (Employment Insurance, Workers' Compensation and Social Assistance or Welfare)
- previously reported wage rate is not fed back directly, but if the respondent reports a wage rate (with the same employer as the previous year) that has increased or decreased by more than 10%, then the respondent is prompted to ensure that this is a true change.

In the January 1997 and 1998 SLID interviews, the confirmation rates were about 98% for employer name, 97% for occupation and 94% for school attendance.

Processing

Certain aspects of the processing strategy lead to spells starting or ending at the seam. In many cases, the analyst can take this information into account. Essentially, the difficulty arises when persons who are not in their first year as SLID respondents report dates which are prior to the reference year. If this information were used to update earlier information, it would cause an inconsistency in the database that would have a "ripple effect" on derived variables for the person, family and household. It is not clear that this new date is correct, since the recall period for a date prior to the reference year is 12 months longer than that for the original information. Therefore, the SLID

processing strategy is to change the date to the end of the previous reference period.

An example may serve to explain the approach. In January 1997, the person reported working for ABC Company in a job that started in March 1996 and was still ongoing. The database reflects this job unended at the end of 1996. At the next labour interview in January 1998, the person reports that his job with ABC Company actually ended in August 1996. SLID processing does not revise the 1996 data to reflect this. Instead, the SLID database is set to show that the job ended in December 31, 1996. The same result would occur if the person said in January 1998 that he had never worked for ABC Company. Job ends are also set to the end of the previous reference year for persons who become out-of-scope (dead, institutionalized, emigrated, turn age 70) or for whom no data are collected at the next interview (principally refusals and those who cannot be traced). Similarly, if the person reported in January 1998 that he worked for DEF Company, starting in October 1996, the start date for that job would be set to January 1, 1997.

2.2 Start Dates

Figure 2 shows the pattern of job starts by month over the three years from January 1993 to December 1995. (A similar figure for weekly job starts shows so many spikes that it is difficult to see patterns.) These job starts have been reported by longitudinal respondents who were between 16 and 69 years of age during all three years.

The most striking feature is the high number of job starts in January 1994. This can be attributed to a collection problem that occurred during the interviews for reference year 1994. Due to difficulties in the computer-assisted interviewing application, about 30% of the SLID sample did not

receive labour information feedback during the January 1995 interview, though they did receive feedback on demographic variables. The interviewer did not have access to the employer and job as reported the previous year, so jobs that were described in different terms could have been interpreted as new, even though they were actually the same jobs as the previous year. For example, a respondent might have reported working at Statistics Canada one year and the Government of Canada the next year. With dependent interviewing, the processing system would have prompted the interviewer with the name of the employer as of the previous December and the respondent could have confirmed or denied that employer. Without dependent interviewing it might not be recognized that the two employers were actually the same.

Efforts were made to minimize the impact of this problem by manually reviewing the cases, but inevitably, some December 1993 jobs were not matched to the January 1994 jobs held by respondents. Since jobs cannot begin prior to the reference year for continuing respondents, these cases had to be treated as jobs that ended on December 31, 1993, with a new job starting on January 1, 1994.

Other than January 1994, the annual patterns are similar across the three years. May and June have the highest number of job

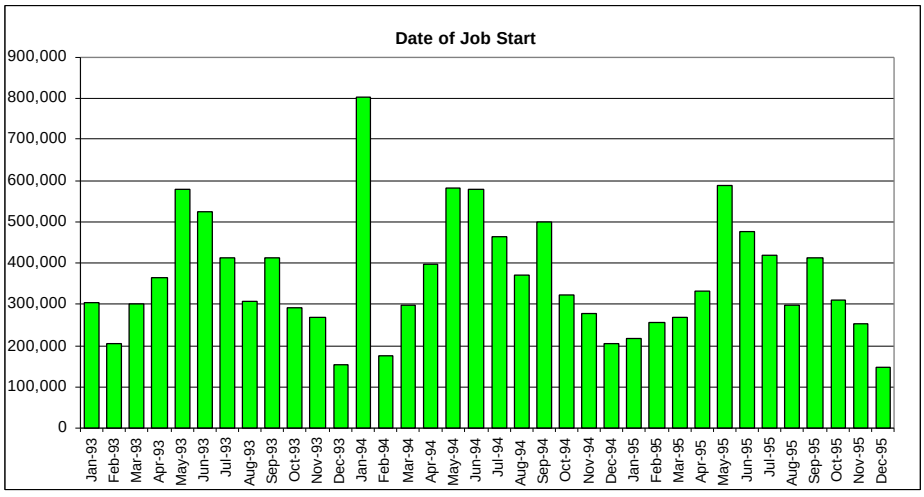

Figure 2: Number of job starts, by month

starts each year, due to student and other seasonal employment, and winter months have the lowest number of job starts.

2.3 End Dates

Figures 3 and 4 show the pattern of job ends by month over the three years from January 1993 to December 1995. As with the previous graph, these job ends have been reported by longitudinal respondents who were between 16 and 69 years of age during all three years. At the time of writing this paper, SLID data had not yet been released for reference year 1996. Therefore, data for December 1995 are not shown because the number of job ends depends on the status of jobs in January 1996.

Figure 3 shows the "normal" end dates of jobs, excluding jobs that were denied by respondents and jobs that were ended due to processing requirements, as described in sections 2.1 and 2.2. The annual patterns are regular over the three years. August, September and December show the largest number of end dates, reflecting the end of jobs as school semesters start in September and January and the end of Christmas jobs in December.

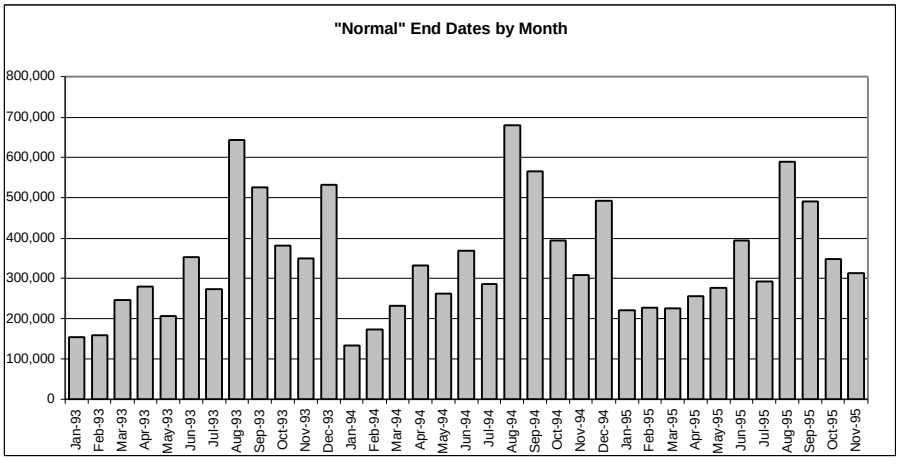

Figure 3: Number of "normal" job ends, by month

Figure 4 shows the "normal" end dates of figure 3, as well as "processing" end dates due to persons who become non-respondents, persons who become out of scope and persons who deny information fed back to them from the previous interview. These processing dates can be identified and removed before analytical investigations are carried out.

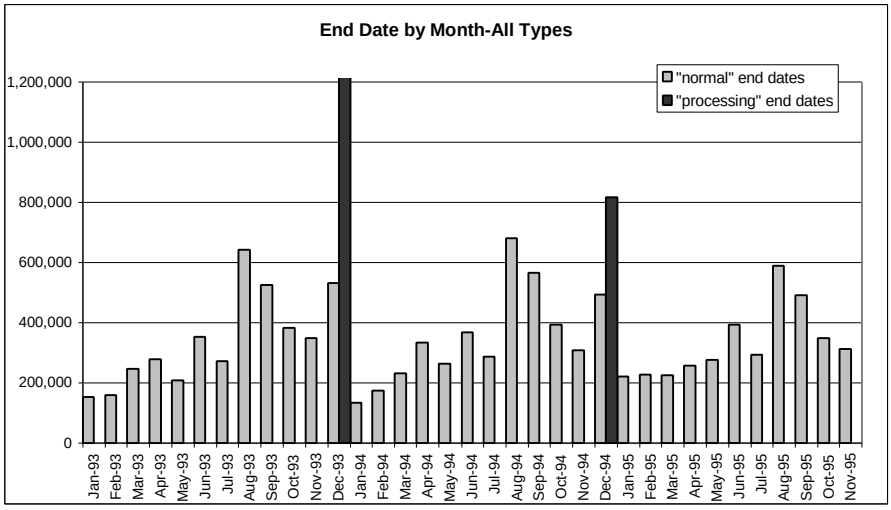

Figure 4: Number of "normal" and "processing" job ends, by month

2.4 Labour Force Status Transitions

Definition

In any given period, a respondent's labour force status can have one of three mutually exclusive values: employed (E), unemployed (U) or not in the labour force (N). If more than one state applies to a reference period, then priority is given to E over U and to U over N. There are six possible transitions between pairs of states.

Weekly or Monthly Transitions?

As with start and end dates, transitions from one labour force status to another can be tabulated by week or by month. In contrast to job starts and job ends, monthly labour force status transitions are not simply the sum of the relevant weekly transitions because short spells of unemployment or job search may

not appear as monthly transitions. The following two examples illustrate how a situation may result in different sets of monthly and weekly transitions.

Figure 5 shows a situation in which a short spell of unemployment does not enter into the monthly labour force status transitions. This case would be interpreted as two weekly transitions: one E → U (week 1 is classified as E and week 2 is classified as U) and one U → E (week 2 is classified as U and week 3 is classified as E). There would be no monthly transitions because the labour force status is E in March, April and May.

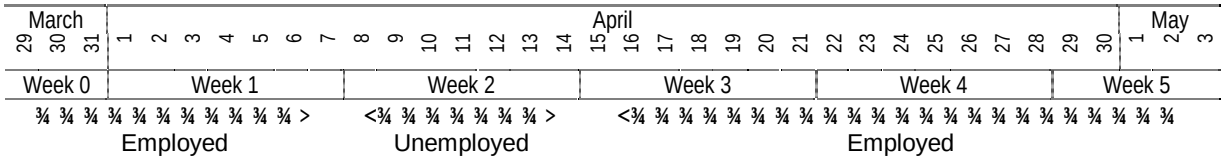

Figure 5: Two weekly transitions (U ® E and E ® U), but no monthly transitions

Over the six possible labour force transitions, there were about 25% more weekly transitions detected than monthly. In the case of N → E there were fewer weekly transitions than monthly. This can be explained by the situation shown in figure 6. In this case there are two weekly transitions: N → U (week 1 is classified as N and week 2 is U) and U → E (week 2 is classified as U and week 3 as E) and one monthly transition: N → E (N throughout March and E as the overall status in April).

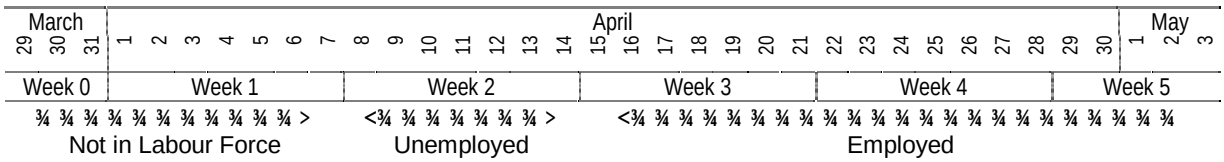

Figure 6: Two weekly transitions (N ® U and U ® E), one monthly transition (N ® E)

Graphs of Transitions

The data in figures 7 to 12 have been restricted to longitudinal respondents between the ages of 24 and 55 in order to reduce the transitions due to school attendance for younger respondents and due to retirement for older ages. Only monthly transitions are presented because weekly transitions show so many spikes that it is difficult to see patterns. As with job ends, transitions in December 1995 are not shown because they depend on the labour force status in January 1996.

Employed ® Unemployed

As might be expected, E → U and U → E are the most frequent transitions, accounting for over half of both weekly and monthly transitions.

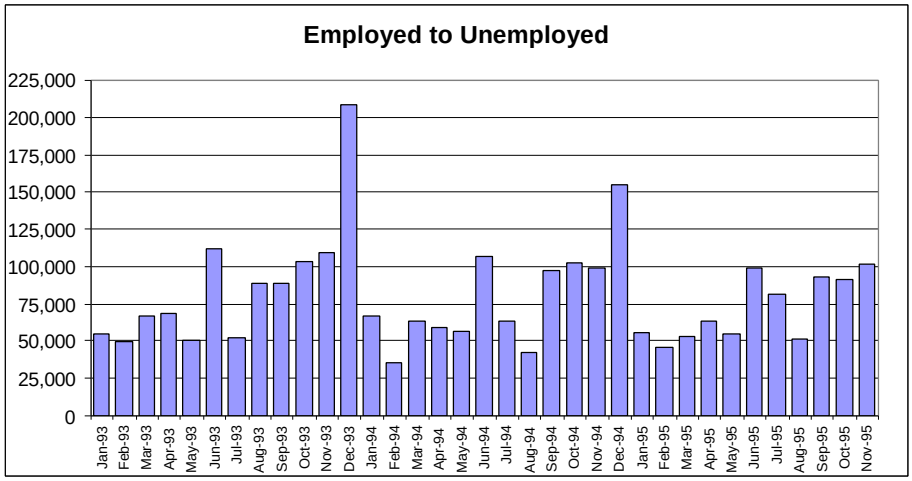

Figure 7: Employed to Unemployed transitions

December 1993 and 1994 have the highest number of E → U transitions. In general, there are more transitions closer to the end of the year, although June is second highest. The higher number of transitions near the end of the year may be due to the tendency of respondents to recall short periods of employment or unemployment that occurred relatively close to the collection period in January.

Unemployed ® Employed

The three years have similar patterns: peaks in April/May and in August. December 1993 has a large number of transitions, though it is not the highest of the three years. Some December transitions should be due to seasonal employment.

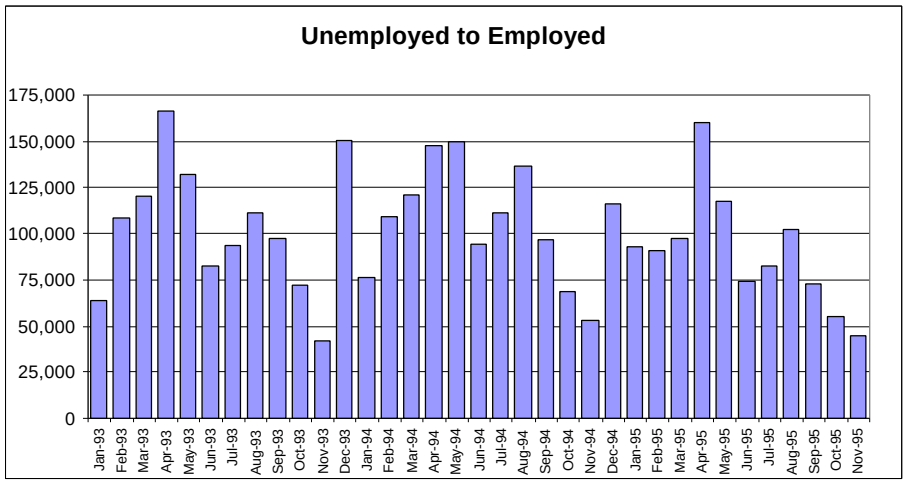

Figure 8: Unemployed to Employed transitions

Many of the monthly transfers in figures 9 to 12 are relatively small. An estimate of 20,000 has an estimated coefficient of variation of 30%. This means that a 95% confidence interval on an estimate of 20,000 has a lower bound of 8,000 and an upper bound of 32,000.

Employed ® Not in the Labour Force

August and September are high in all three years, though June is highest in 1995. Months in the latter part of each year seem to have more transitions than months in the earlier part.

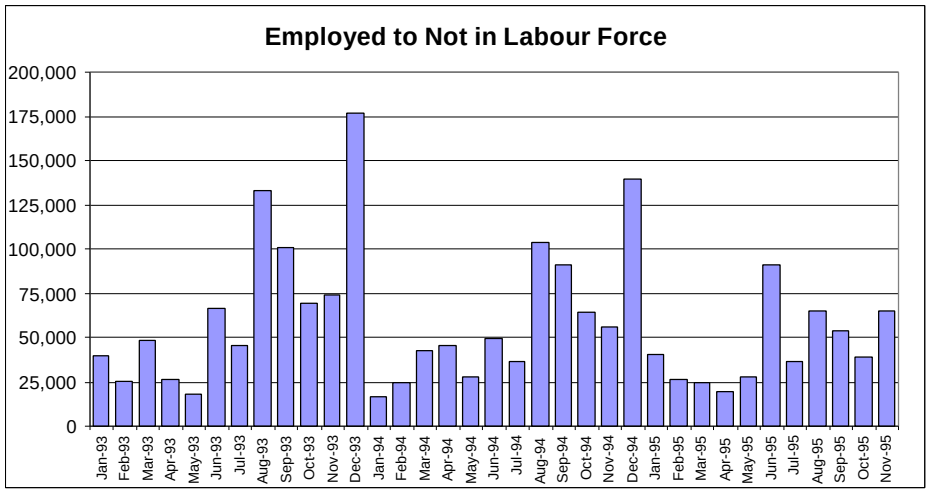

Figure 9: Employed to Not in Labour Force transitions

Not in the Labour Force ® Employed

December transitions are high, both in 1993 and 1994. The pattern over all three years shows high values in March/April/ May and in August. This is similar to the U → E pattern.

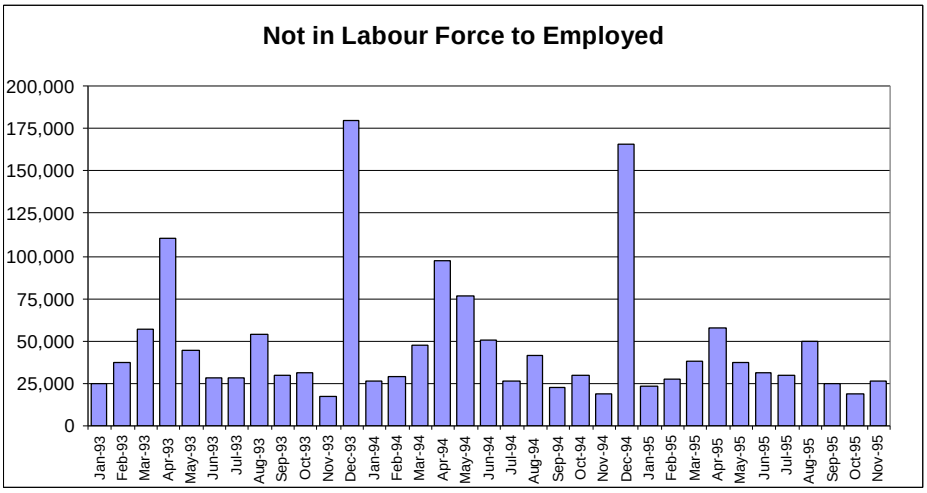

Figure 10: Not in Labour Force to Employed transitions

The N → E flows may reflect situations in which there is a high incidence of on-call, or a lack of job search because of a perceived lack of opportunity in one-industry towns.

Unemployed ® Not in the Labour Force

Many of the U → N transitions are low and therefore have a high variance. The high values in August reflect the start of school and the end of some opportunities for seasonal employment. These flows can also indicate discouragement.

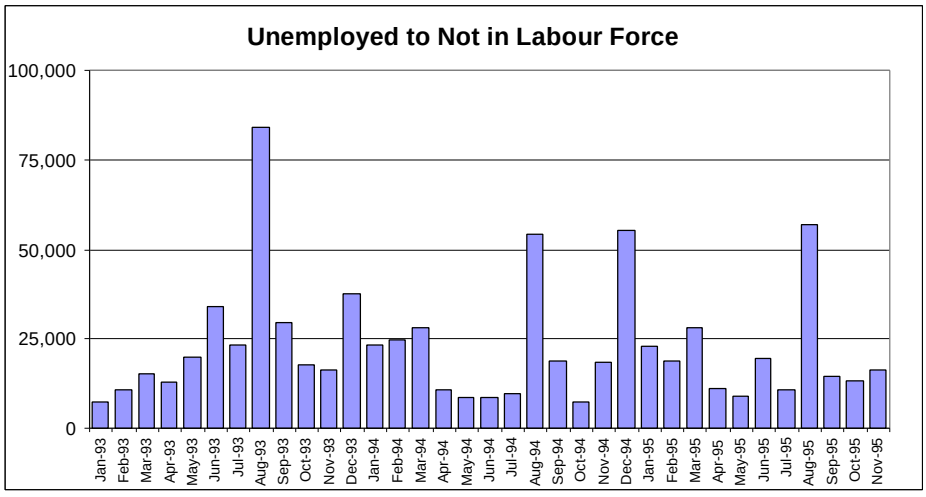

Figure 11: Unemployed to Not in Labour Force transitions

Not in the Labour Force ® Unemployed

The high values of $N \rightarrow U$ transitions in December 1993 and 1994 are due to a change in reported job searches.

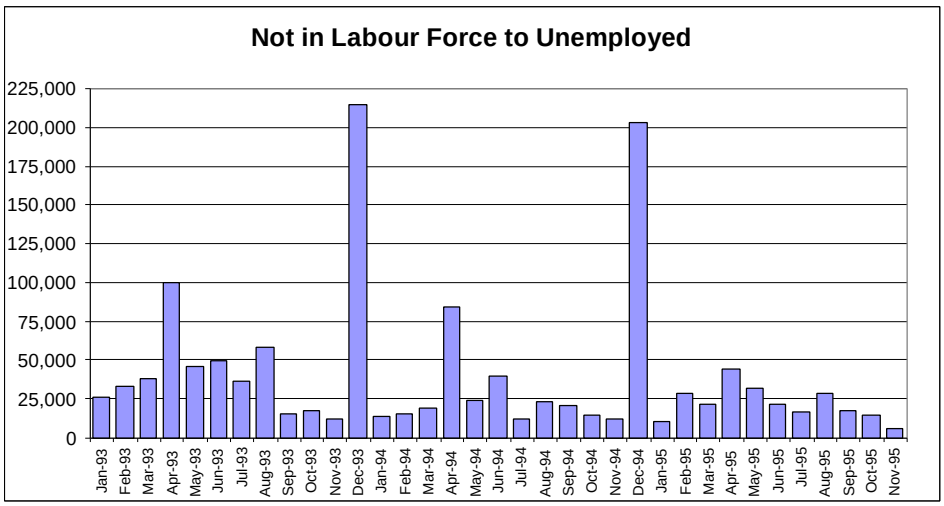

Figure 12: Not in Labour Force to Unemployed transitions

3. REFERENCES

Lemaître, G. (1992). "Dealing with the Seam Problem for the Survey of Labour and Income Dynamics". Statistics Canada, SLID Research Paper Series 92-05.

Statistics Canada. (1998). "Survey Overview: Survey of Labour and Income Dynamics". Ottawa.