

**Catalogue No. 94-07**

**THE SURVEY OF LABOUR AND INCOME DYNAMICS:  
LESSONS LEARNED IN TESTING**

March 1994

Maryanne Webber, Household Surveys Division

The SLID Research Paper Series is intended to document detailed studies and important decisions for the Survey of Labour and Income Dynamics. These research papers are available in English and French. To obtain a summary description of available documents or to obtain a copy of any, please contact Philip Giles, Manager, SLID Research Paper Series, by mail at 11-D8 Jean Talon Building, Statistics Canada, Ottawa, Ontario, CANADA K1A 0T6, by INTERNET ([GILES@STATCAN.CA](mailto:GILES@STATCAN.CA)), by telephone (613) 951-2891, or by fax (613) 951-3253.



## **EXECUTIVE SUMMARY**

This report is based on a presentation made to the Annual Research Conference of the US Census Bureau in March 1994. It reviews some of the highlights of a year of testing for the Survey of Labour and Income Dynamics. Among the topics covered are the impact of user consultations on definition of the survey content; lessons learned in designing questionnaires for computer-assisted interviewing and in the use of dependent interviewing; testing the approach for assigning labour force status and for collecting income data; and the development of a prototype for data quality reporting.



## TABLE OF CONTENTS

	Page
1. BACKGROUND	1
2. USER CONSULTATIONS ON CONTENT: IMPACT ON TESTING	3
3. COMPUTER-ASSISTED INTERVIEWING	9
4. LABOUR MARKET ACTIVITIES: USING CAI TO COLLECT AND EDIT SPELL DATES	13
5. DEPENDENT INTERVIEWING	14
6. LABOUR FORCE CLASSIFICATION	20
7. INCOME REPORTING	23
8. DATA QUALITY REPORTING	26
9. CONCLUSION	27
BIBLIOGRAPHY	27



## 1. BACKGROUND

During 1993, Statistics Canada conducted field tests in preparation for the launch of a new panel household survey called the Survey of Labour and Income Dynamics (SLID). The purpose of this paper is to summarize the experience gained from these tests. To set the context, a brief description of the main features of the survey follows.

SLID is intended to complement cross-sectional labour and income data, and to support labour market and social policy research. Some of the research topics currently envisaged deal with employment and unemployment dynamics, transitions from school to work and from work to retirement, job quality or quality of working life, family economic mobility and the dynamics of low income.

Once selected, a respondent is followed for six years. In that time thirteen interviews are completed: a preliminary interview at the beginning of the first year, followed by six waves of annual, retrospective labour and income interviews. The purpose of the preliminary interview is to collect historical information (marital history, educational attainment, work experience, and so on) as well as information on current labour market activities. The labour interviews occur in January and the income interviews in May. In both cases, the previous calendar year is the reference period.

The survey will be conducted using overlapping panels. The first panel started in 1993. The second will be introduced in 1996, at which point we will be up to full sample (30,000 to 40,000 households). In 1999, the first panel will be rotated out and a third one introduced.

When the sample for a panel is chosen, all members of selected households are considered to be *longitudinal respondents* and are followed through time, even if

they move. People who move in with longitudinal respondents at some time after the sample is selected are called *cohabitants*. Cohabitants are interviewed as long as they continue to live with a longitudinal respondent.

In 1993, in addition to completing the preliminary interview on the first panel, a major two-part field test was conducted. This was a test of the content and collection procedures for the first wave of labour and income data collection in 1994. Computer-assisted interviewing (CAI) is being used; interviewers work on notebook computers, generally out of their own homes, importing and exporting their caseload through the phone lines. Thus learning to function in a CAI environment (and indeed contributing to the development of that environment) featured very prominently in our test objectives.

The test sample consisted of 1400 households in Newfoundland and the major cities of southern Ontario. These households were formerly in the Labour Force Survey (LFS), and information from the LFS was used to simulate a longitudinal survey setting -- for example it gave us some information to feed back as a test of our proposed use of dependent interviewing.

The test environment throughout 1993 was far from static. SLID interviewers are LFS interviewers and the LFS was itself gearing up for conversion to CAI. Along with the LFS, a whole host of household surveys are either being converted, or developed from scratch in CAI. Since these surveys are conducted by the same interviewers using the same notebook computers, design consistency from one questionnaire or application to the next was critical.

In addition to this major test -- the labour part was called Test 3A and the income part Test 3B -- we conducted a few mini-tests of our final questionnaire content late in the year. The results presented below are based on a range of evaluative



approaches. Where possible, we undertook comparisons to benchmark sources and micromatching. Observation of interviewing and systematic interviewer debriefing were often the most appropriate means, and these were frequently used. The tests were not in general designed to allow us to separate out various effects; rather a full package was being tested, a constraint imposed by a tight development schedule.

## **2. USER CONSULTATION: IMPACT ON TESTING**

When the SLID project was approved, the point of departure for defining content was the Labour Market Activity Survey (LMAS) and the Survey of Consumer Finance (SCF). LMAS was conducted five years -- one two-year and one three-year panel -- and SCF is the annual cross-sectional survey that produces estimates of family and household income, including official estimates of families and individuals living below the low income cutoff. Users were familiar with the content of these surveys, so rather than define content starting with clean slate, we sought to identify desirable or necessary changes relative to LMAS and SCF.

Consultations with data users resulted in decisions to test new sets of questions and new approaches including the following:

- detailed household relationships, similar to the matrix completed in the Survey of Income and Program Participation (SIPP);
- questions on workers' supervisory and managerial responsibilities (on an employer-by-employer basis);
- a module on the impact of unpaid care-giving and receiving on the respondent's ability to participate in the paid labour market;

- a limited set of questions on assets and debts, on the grounds that annual income on its own does not have enough explanatory power for some of the phenomena the survey is attempting to track.

In addition, while not strictly new content, an attempt was made to improve the quality of income reporting by making the income categories as compatible as possible with the tax form. Our goal was to encourage respondents to transcribe amounts directly from their tax form. It was a gamble because the approach made it more difficult for respondents who did *not* refer to their tax records.

Innovative content consumed a great deal of time and energy in the lead-in to the test. We tried to adapt questions and procedures that had been used in other surveys, but changes were inevitable for substantive reasons or because we were adapting paper-based questionnaire items to CAI. A description of what we learned in testing this new content follows.

### **Detailed household relationships**

SIPP uses a matrix to describe the relationship of every person in the household to all other persons (in contrast to the more traditional approach of relating everyone to a single reference person). Although this matrix looks relatively straightforward on paper, it was not instantly converted to CAI. Even though substantial development time and effort were devoted to it, the version we tested in January 1993 was error-prone, with interviewers frequently selecting the incorrect response category. The version we are now using was retested in November and, based on observation and interviewer feedback, is working well. Other Statistics Canada household surveys operating in CAI have begun to incorporate this module as well -- the information collected is much richer than in the traditional approach.

### **Supervisory and managerial responsibilities**

These questions were partially based on a set used in the General Social Survey, and seemed to work well in the field. The questions included:

- the number of people directly supervised;
- the percentage of working time spent in supervisory activities;
- whether the person is best described as a manager, supervisor or something else;
- if manager, was it a top, upper, middle or lower management position.

Even though 55% of all interviews were completed by proxy, there was a very low occurrence of "don't knows", ranging from 0.1% to 1.6% of the number answering the question. The replies on supervision were evaluated in the light of the Standard Occupational Code (SOC) assigned using information on the kind of work performed by the person and his or her major duties.

**Table 1. SOC code by reply to supervisory question**  
**Reply to supervisory question**

	<b>Supervised</b>	<b>Did not supervise</b>
<b>SOC:</b>		
<b>Supervisor</b>	<b>15%</b>	<b>2%</b>
<b>Manager</b>	<b>37%</b>	<b>4%</b>
<b>Neither</b>	<b>48%</b>	<b>94%</b>
<b>Total</b>	<b>100%</b>	<b>100%</b>
<b>N=</b>	<b>255</b>	<b>734</b>

Close to half of the people who were reported as supervising were not classified as supervisors or managers in terms of their occupation description. However, the SOC reflects major job functions and it is quite possible that someone who is not

assigned a supervisor or manager occupational code is nevertheless supervising other employees. What is less plausible is the reverse: someone classified as a manager or a supervisor in SOC terms who claims no supervisory responsibilities whatsoever. This inconsistency occurred in 6% of the cases reporting no supervisory duties, which was tolerable. Anomalies in the other questions were also at a low level. Response patterns were roughly comparable for proxy and non-proxy respondents (Coish, 1993).

In short, this small set of questions worked reasonably well from the respondent's perspective. However, their analytical value was not that clear. We held a workshop with a group of researchers specialising in sociology of work. Their input clarified the analytical potential of the data, in studies on promotions and career paths; it also allowed some refinement of the questions. The end result is that we have retained the questions and are looking forward to seeing their uses.

### **Unpaid care-giving and receiving**

These questions, which all referred to the previous calendar year, covered the following topics:

- did respondent provide unpaid help to anyone needing care because of age, disability, physical or mental illness or any other reason
- relationship of respondent to person(s) for whom this care was provided
- months when care was provided
- number of hours per week
- did provision of unpaid help limit amount of paid work respondent could do
- did respondent spend any time providing unpaid care for the child of another person

- (if so, same questions as above on time spent and impact on ability to perform paid work)
- (if respondent has children under 18) did respondent receive any unpaid help taking care of own children (not counting help from other parent)
- months when help was received
- number of hours per week
- did help increase amount of paid work respondent could do
- (if no help received) did this limit amount of work respondent could do

These questions were intended to provide information on the impact of non-market activities (or one facet of them) on a person's ability to participate in the paid labour market. Once again, the General Social Survey had at one point asked a set of similar questions, and this was our point of departure.

Although there were not many "don't knows" and refusals to these questions, observation and interviewer debriefing indicated quite widespread respondent confusion about what "we were trying to get at"; The interviewers frequently had to explain and paraphrase (Saint-Pierre, 1993). As Table 2 shows, there were also differences in the response patterns of proxy and non-proxy respondents.

**Table 2. Percentage of respondents providing or receiving unpaid care, by proxy status**

	<b>Proxy</b>	<b>Non-proxy</b>
<b>Providing unpaid care to ill, disabled person</b>	<b>6%</b>	<b>10%</b>
<b>Providing care to other person's child</b>	<b>3%</b>	<b>8%</b>
<b>Receiving unpaid help for care of own children</b>	<b>8%</b>	<b>14%</b>

In these questions, there were problems delineating what activities were and were not to be counted, and there were no clear concepts, that we know of, to guide the way. Why should we not also include questions on care of one's own children? If we did, why not include other forms of unpaid help? As it was not possible to resolve these problems, this set of questions was dropped, despite its obvious relevance to a survey tracking flows into and out of the labour force.

### **Assets and debts**

User consultations made it clear that wealth information would be tremendously useful in the context of SLID. The last asset and debt survey conducted by Statistics Canada was in 1984. That was a long, detailed survey, not really appropriate for a telephone interview. Our point of departure was rather the wealth content of the Panel Study of Income Dynamics. Our objective, assuming reasonable test results, is to ask these questions perhaps twice in the six-year lifespan of a panel.

Test 3B included 17 categories of assets and 5 categories of debts. For assets, among respondents who chose to participate, the results showed a negligible number of refusals and "don't knows" with respect to whether or not a particular asset or debt applied. Estimating the value of an asset was predictably more difficult. Over all categories, the amount was unknown 4.6% of the time. However, the "don't knows" were not uniformly distributed throughout the asset and debt categories. In particular, the value of farms, of annuities and of Registered Retirement Income Funds were unknown in over 20% of the cases. In all other categories, the "don't knows" accounted for under 10% of all cases where the asset or debt was reported.

Respondents rarely refused to supply an amount once an asset or debt category was identified, except in the case of bank accounts (6% of all cases where this asset was reported), credit card balances (3%) and personal loans (3%).

The evaluation of the test results on wealth has not yet been completed. Indeed, the basis for evaluation is somewhat problematic, given the lack of benchmarks.

### **3. COMPUTER-ASSISTED INTERVIEWING**

The testing in 1993 provided many lessons on survival in the CAI world. In an on-going survey that is converting from paper-and-pencil to CAI, constraints are imposed by the need for continuity in the time series. In contrast, SLID was *designed* for CAI.

The first lesson we learned was that the questionnaire specification process needs a structure and procedures. We eventually developed procedures for communicating question wording and flows in a consistent way and in terms appropriate for programming. That procedure is described in a SLID research paper (Dibbs and Hale, 1993).

Next, it quickly became clear that the task of thinking through all possible response patterns is slow, painstaking work. In the paper world, the interviewer's manual serves as an escape route in questionnaire design impasses; a problem can be resolved by "dealing with it in the manual". Questionnaire design in CAI is excruciatingly explicit. Paper questionnaires are also constrained by space limitations, which is sometimes a blessing in disguise. In CAI, time rather than the space is the limiting factor.

CAI offers the potential to tailor the questionnaire to fit various respondent scenarios. This is a bonus, because it can reduce respondent confusion over questions that don't quite fit; the need for interviewer rephrasing of questions diminishes. However, the added complexity exacts a price. More complexity in the questionnaire paths means more testing of the application and a greater risk of error. Another pitfall of tailoring, that we discovered in Test 3A, is that it creates work after data collection. Inevitably, there is a need to pull the highly disaggregated information back together again through the creation of derived variables. After some experimentation, we eliminated some of the tailoring because the gains were more than offset by the added complexity. To keep this in perspective, our labour application is still far more extensive than a paper version could have been: 499 "display" questions that the interviewer may see on the screen and 731 "non-display" questions directing flow behind the scenes.

We began with a similar enthusiasm for interactive edits, and again pulled back somewhat from our initial thinking about what the optimal level of editing would be. The trick is to maximize accuracy and consistency without overburdening either the interviewer or the respondent. We found that edits had to be simple -- ideally a simple instruction to the interviewer to confirm the accuracy of an item. Also, we believe that interviewers need an "escape route" -- a way of bypassing the edit and allowing an inconsistency to remain unresolved -- unless the item is so critical that the interview cannot continue without it -- for example, in SLID, age is needed to determine whether or not labour and income questions should be asked.

In our income test, edits were triggered often, and we found interviewers were automatically overriding them. Except for the reporting of spell dates, we have kept the number of edits in SLID at a modest level, because testing showed that, if edits are triggered frequently, they can really impede the interview.



As specifiers, many of us began with an overly-structured questionnaire design approach and then retreated, leaving more discretion to the interviewer. One example is in the area of updating demographics to reflect changes in marital status. Our intention was to do "conditional updating": when there was a household joiner, when a person's marital status was common-law, when there was an adult leaver, and so on, we wanted to trigger an appropriate set of questions to ensure that the demographics were updated. However, there are so many possible situations that this amount of structuring became unwieldy and we have retreated to a simpler approach of training the interviewer to confirm or update the information in the household roster right on the screen. As a general rule, it may be easy to program a question and set of answer categories, but it is not at all easy to convert an interviewer procedure requiring judgement to a series of mechanical conditions and actions.

Finally, we found the task of applications testing to be a major undertaking. Our approach was rather unstructured at first; we developed a problem log form and distributed it to staff identified as testers. We found that testers would cause a program failure and then not be able to recreate the circumstances. We have since developed more systematic test methods. For example, scenarios are developed on paper, and particular testers follow through on pre-defined scenarios. If a crash occurs, the cause can easily be traced. This more disciplined approach also ensures that the testing effort covers all the bases. In a longitudinal survey, there are many links from one wave to the next and the testing must simulate scenarios where information collected in previous interviews is being fed back and updated. This means that test cases have to be pre-defined and loaded into the system. In short, the application testing is a complex, multi-phase operation and we have not yet got it down to a fine art. In addition, given that our interviewing is decentralised, there is an equally critical phase of transmission testing, which covers moving cases between head office to the regional office, and from there to

interviewers. Links are required between the interviewers, their senior interviewers and "tracers" in the regional offices. SLID is the first survey other than LFS to be conducted on this "household survey platform". As can be expected, we will be in development mode for some time to come.

With respect to software, we began with CHRR, the CAI software developed at Ohio State University, because it was suitable for a decentralized operation and had many attractive features for the collection of spell dates. However, we found some of the design features impeded our collection of demographic data. (Our approach is to start by building or updating a household roster, then do demographics on the household, then labour or income on eligible individuals. ) For this part of the interview, we found it preferable to build our own application using FOXPRO, a database software. However, since the labour questions are programmed in CHRR, we had to ensure that the demographic questions did not look alien to the interviewers. For example, the function keys had to serve the same purpose, whatever the software. This was quite time-consuming, but it worked. The initial development effort aside, the main issue from the interviewer's perspective is the time delays due to switches between software.

For the income interview we also elected to build an application using FOXPRO. We wanted to show the full form on the screen and allow the interviewer to scroll through, much as in a paper form. This feature was the deciding factor. In addition to a mix of CHRR and FOXPRO, the SLID application is embodied in a case management system written largely in Clipper. So the full system is quite complex.

#### **4. LABOUR MARKET ACTIVITIES: USING CAI TO COLLECT AND EDIT SPELL DATES**

As mentioned above, one of SLID's predecessors is the Labour Market Activity Survey. LMAS collected information on spells of employment, joblessness, job search, absence from work and so on. LMAS used a traditional paper questionnaire, and one of the most time-consuming activities in processing was checking for and correcting obvious errors in the reporting of dates. The errors may be due to respondent recall to interviewer error. (Interviewers had to match up dates recorded in various parts of the questionnaire to see whether certain blocks of questions should be asked. This procedure was error-prone.) In SLID, one of the greatest advantages of CAI is the capacity to avoid interviewer error by automating these date comparisons, and to reduce respondent error by detecting and querying inconsistencies in the dates reported by respondents.

Our approach, like LMAS, is to first establish the dates of employer spells during the year. Employer spells refer to the start and end dates of a person's attachment to a particular employer. At the end of the year, an employer spell may be in progress. If so, the identity of the employer will be fed back the following year. Employer spells may be concurrent, overlapping, intermittent, and so on. If the respondent is uncertain about exact dates, the interviewer will ask for an approximate date; as a fallback, the interviewer records "don't know", in which case employer start dates are set at the first of the month and end dates are set at the end of the month.

This maximization of employer spells is admittedly biased, but it ensures that jobless spells -- which are automatically derived from the employer spells -- are not created where they do not exist. In Test 3, we went through a process of confirming the jobless spell dates with respondents; virtually 100% of the spells

were confirmed. Observation and interviewer debriefing indicated that respondents viewed this confirmation process as a statement of the obvious. Therefore we dropped it.

Dates of absences from work are also collected, for absences lasting one week or longer. In this case, the start and end date of the absence must be within the employer spell or an interactive edit is triggered. These edits do not guarantee that the data are fully consistent but they do greatly minimize inconsistencies.

## **5. DEPENDENT INTERVIEWING**

We have always been concerned about seam problems and variable recall quality over the course of the reference year. The information from SLID is heavily oriented towards spells -- for example, employment and unemployment spells, work absence spells, spells of social assistance and unemployment insurance receipt. Without dependent interviewing, we believe that we would have far too many spells starting and ending at the interface of the two reference years.

Evaluation of data from LMAS illustrates that transitions at the seam, between December and January, can far exceed those between any other pair of months. Two charts taken from Lemaître (1992a) illustrate the point. These charts are based on month-to-month transitions, over a two-year period, 1988 and 1989. They reflect annual data collected retrospectively in January 1989 and January 1990. Chart 1 shows monthly transitions from the state "self-employed" to "not self-employed" and the converse. In both cases, the number of transitions at the seam between December and January far exceeds the level recorded in any other month.

CHART 1 Monthly flows into and out of self-employment, LMAS, 1988-1989.

— Not self-employed → Self-employed  
 - - - Self-employed → Not self-employed

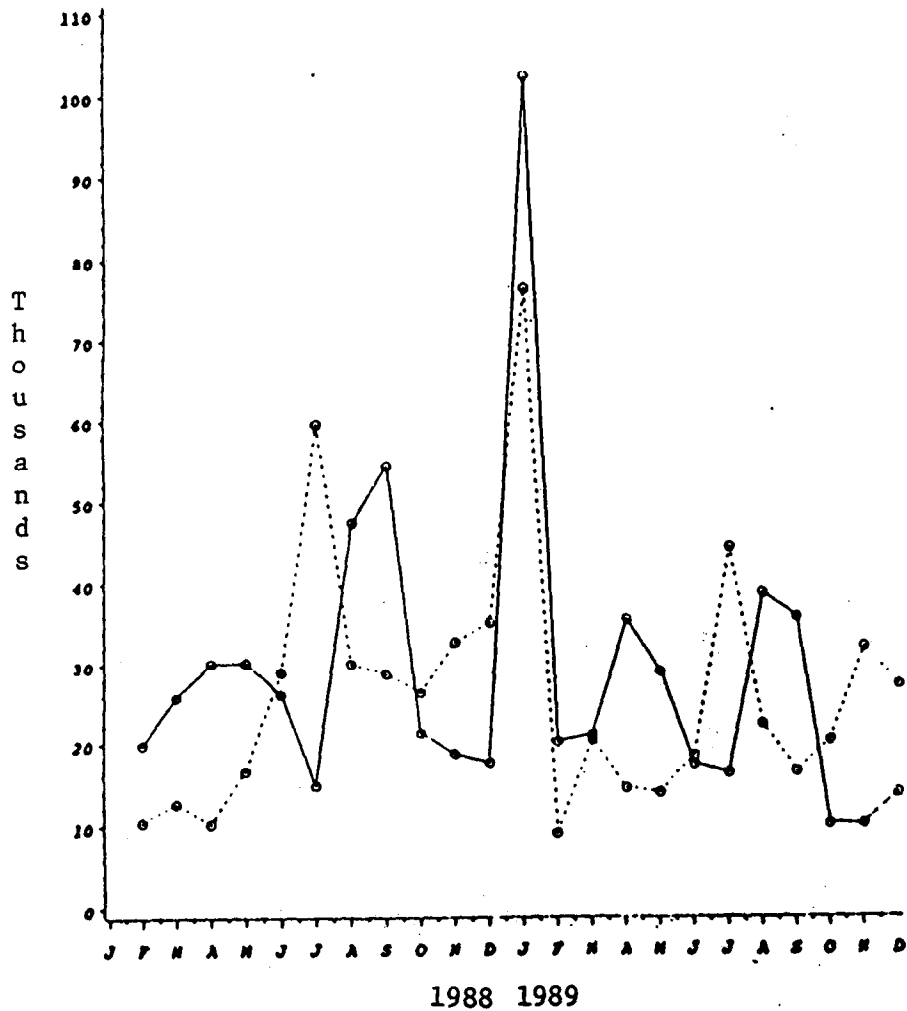


CHART 2 Monthly flows between "looking" and "not looking" for work, LMAS, 1988-1989.

— Did not look for work → Looked for work  
 - - - Looked for work → Did not look for work

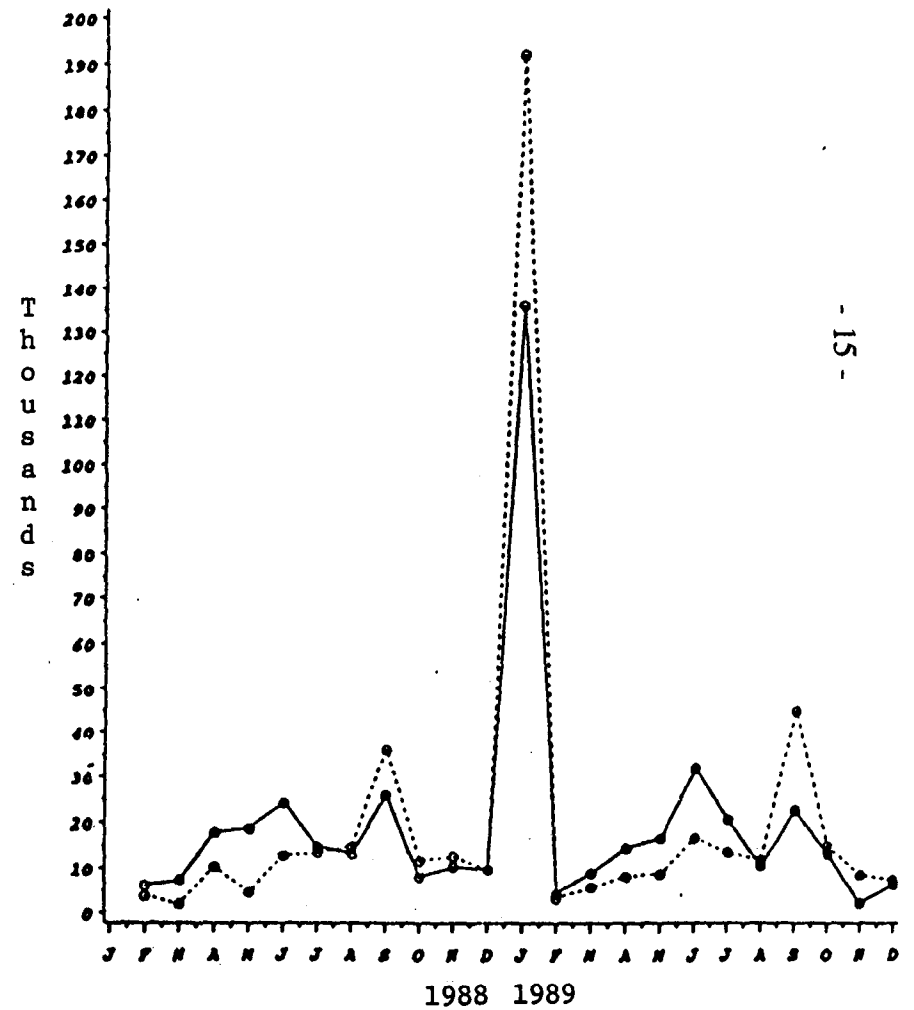


Chart 2 shows the monthly flows from the state "looking for work" to "not looking for work" and vice versa. The peak in this case is even more extreme.

LMAS experimented with dependent interviewing. If a job was in progress at the end of one reference year, the name of the employer would be fed back at the beginning of the interview the following year. This technique was found to reduce the number of jobs starting and ending at the seam quite substantially.

We wanted to extend this practice and use dependent interviewing for several key variables. Thus a major objective of Test 3A was to gain some experience in the use of dependent interviewing. The variables where dependent interviewing was considered were:

- names of employers at beginning of reference year;
- absence spell of four weeks or more in progress at the beginning of the reference year;
- occupation;
- class of worker, in the case of self-employed or unpaid family workers;
- search status, that is, whether the person was looking or not looking for work at the beginning of the reference year.

In addition we planned to use dependent interviewing for wage rate information, and for school attendance, but our test design did not allow us to test these.

Dependent interviewing can take many forms. One approach is to feed back information in an active way, which was largely the approach we adopted. In this approach, one must seek confirmation that the information being fed back is indeed correct and allow for denial. The following illustrates how the approach works in feeding back occupation.

"Our records show that in January 1992, the kind of work Ray Jones was doing was carpet laying. Is this correct?" *If the respondent agrees:* "Did the kind of work Ray Jones was doing for Carpetmania change during 1992?" *If the respondent disagrees:* "What kind of work was Ray Jones doing in January 1992?" *And then:* "Did the kind of work Ray Jones was doing for Carpetmania change during 1992?"

This line of questioning is necessary to separate corrections from substantive change. Here are a few other examples of questions used in dependent interviewing:

"Based on our interview of a year ago, Roy Jones was working for Carpetmania around the beginning of January 1992. Is this correct?".

*If the respondent was jobless at the beginning of the year:* "Based on our interview of a year ago, Ray Jones was looking for work around the beginning of January 1992. Is this correct?"

Every time the interviewer asks for confirmation, there is a possibility that the respondent will deny the accuracy of the statement. However, as shown in Table 3, confirmation rates in the test were quite high.

**Table 3. Confirmation Rates Obtained in Test 3 on Information Fed Back Regarding Activities One Year Ago**

<b>Employer (respondent at work)</b>	<b>99%</b>
<b>Employer (respondent absent 4+ weeks)</b>	<b>96%</b>
<b>Absent from work (4+ weeks)</b>	<b>85%</b>
<b>Occupation</b>	<b>97%</b>
<b>Class of worker</b>	<b>94%</b>
<b>Search status: looking</b>	<b>96%</b>
<b>Search status: not looking</b>	<b>93%</b>

The only instance where the confirmation rate was under 90% concerned absences in progress at the beginning of the year. We therefore learned that it was operationally feasible to build dependent interviewing into our instrument. We also learned other things:

*Beware of overstructuring.* It is very tempting to try to sort out denials in the field, get some understanding of what the problem is. We succumbed to this temptation in testing, but it became clear that it was too complex, and would probably deter the respondent from ever again denying anything.

*Beware of feeding back the absence of activity.* In theory, if an activity is fed back (for example, attending school) so should its converse (that is, not attending school). If this is not done, there is a greater risk that the start date of a new activity will be erroneously pushed back to the beginning of the reference year, creating an abnormally large flow *into* the state in question at the seam. In practice, feeding back non-activities does not work well. For example, some older respondents were perplexed or even offended when we fed back the fact that they were *not* seeking work a year ago. We have therefore opted for an asymmetrical approach in feeding back information on job search and educational activity.



*Beware of feeding back complex information.* For example, we fed back the fact that a person was looking or not looking for work one year ago. But what is "looking"? Our definition centres on active job search, but there is no guarantee that respondents share our view of the world. Discouraged workers were a particular concern; from their perspective, "looking for work" could be the same as "interested in working". Because of this uncertainty, we deliberately broadened the definition of "looking". We included:

- ! anyone who was looking in January 1992 (that is, "one year ago")
- ! anyone who last looked in December 1991
- ! anyone who looked at some time since August 1991 but who stopped looking for reasons related to the labour market. In addition to discouraged workers, this includes people waiting for a recall to a former job and people waiting for replies from employers.

Although this incorporated a large group that was not actively looking in January 1992, the confirmation rates we received were very high. This would logically (and did) result in higher unemployment levels relative to our LFS benchmark. However, the pattern did not maintain itself: our feedback process apparently induced a high level of "looking", but *only* early in the year. We still believe that feeding back "looking" will result in better data, but only if we use a simple activity-based definition of "looking".

On the whole, we concluded from the test that this active type of dependent interviewing is viable when the information being fed back was straightforward and reasonably reliable. Feeding back relatively soft information is hazardous. For this reason, we are not feeding wage information back actively. Instead, an edit has been built in: if the wage reported for an employer in the current reference year is either higher by 10%, or lower, than the wage reported last year, the interviewer probes to ensure the amounts have been correctly reported. The capacity exists to

correct either this year or last year's wage. This type of edit cannot be used too frequently as is it potentially offensive to respondents.

## **6. LABOUR FORCE CLASSIFICATION**

One of SLID's content requirement is to produce spell data on employment, unemployment and inactivity, where the spell may be as short as one week. To the extent possible, the concepts and definitions of the monthly LFS are used or approximated. We tested an algorithm for the assignment of this weekly labour force status variable, and compared the results with LFS monthly data for the 1992 reference year. Also, SLID respondents were former LFS respondents, who had been rotated out of the LFS in May 1992. The overlap allowed us to perform a micromatch for the first five months.

LMAS also produced a weekly labour force status value. Experience from LMAS (and its predecessor, the Annual Work Patterns Survey) suggests that the main quality problems in creating such data from an annual retrospective survey are the under-reporting of short spells of unemployment immediately preceding the start of a new job (Lemaître, 1992b) and the telescoping of job search activity. The result is a tendency to shift the annual pattern of unemployment, under-estimating early in the year and over-estimating at year end, with the net annual effect being under-estimation.

SLID is relying on dependent interviewing to correct for some of the under-reporting and telescoping problems. In addition, we specifically asked questions on how new jobs were obtained, hoping to stimulate recall of job search activities preceding the start of the job.

Employed as a percentage of population aged 16-69, LFS and SLID, 1992.

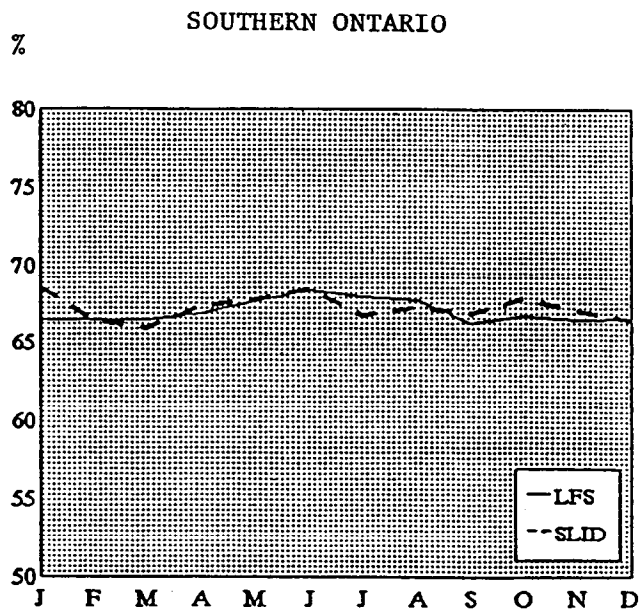
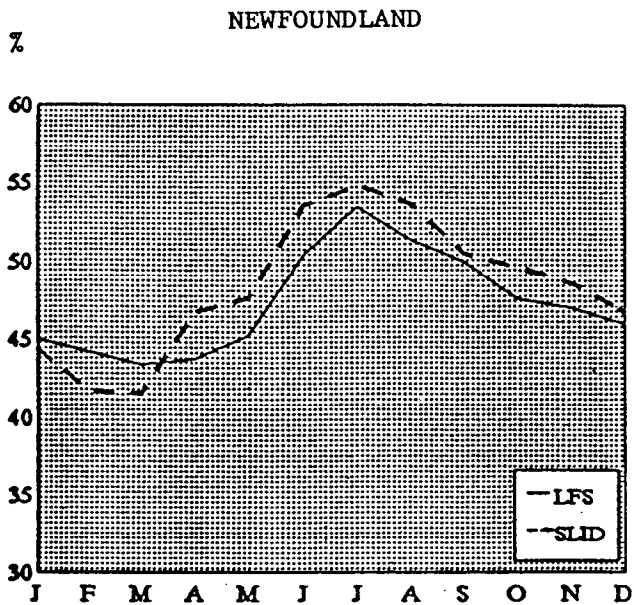


CHART 4

Unemployed as a percentage aged 16-69, LFS and SLID, 1992.

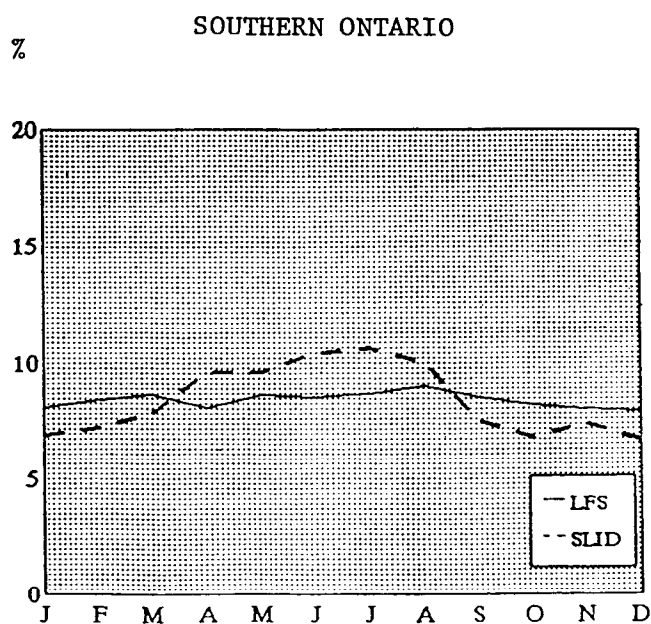
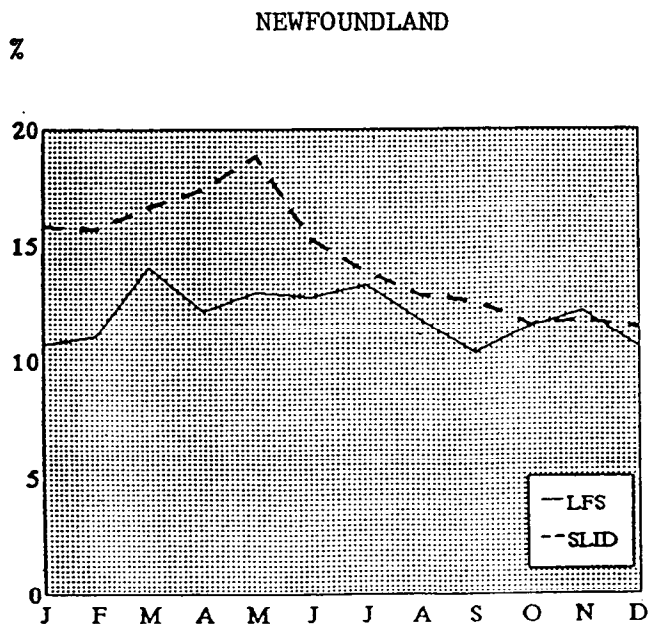


Chart 3 (from Hiemstra, Lavigne and Webber, 1993) shows the percentage of the population classified as employed in SLID and in the LFS, for LFS reference weeks. Results for the two test sites are shown separately because the labour market conditions are very different. The maximum monthly difference is about 3 percentage points and the SLID estimates reflect the seasonal employment pattern quite well. Chart 4 shows the unemployed as a percentage of population. The results for Newfoundland show the impact of the overly-broad definition of "looking" that we used in dependent interviewing: SLID overestimated unemployment early in the year. (Had the pattern maintained itself for the year, we would probably have opted for a broader definition of "looking" for SLID.) Some results from the micromatch are shown in Table 4. This indicates the proportion of cases where the same labour force status was assigned in SLID as in the LFS. For Ontario, by month, at least 89% of the cases were classified the same way. This is remarkably high because there are several unavoidable differences between SLID and the LFS in the labour force classification process. The agreement rate was highest in January (94%), presumably because recall error was offset by the effect of dependent interviewing.

**Table 4. Agreement Rate (Proportion of Cases Where the Same Labour Force Status Was Assigned in the LFS and SLID), Newfoundland and Southern Ontario**

	Newfoundland		Southern Ontario	
	Total cases	Agreement rate	Total cases	Agreement rate
<b>January</b>	<b>839</b>	<b>89%</b>	<b>1175</b>	<b>94%</b>
<b>February</b>	<b>838</b>	<b>83%</b>	<b>1166</b>	<b>91%</b>
<b>March</b>	<b>837</b>	<b>83%</b>	<b>1162</b>	<b>89%</b>
<b>April</b>	<b>834</b>	<b>81%</b>	<b>1157</b>	<b>90%</b>
<b>May</b>	<b>832</b>	<b>78%</b>	<b>1150</b>	<b>91%</b>

The Newfoundland situation is less encouraging. The classification errors were examined, and they appear to be fairly random, except in the case of errors between unemployment and inactivity. Each month, between 1.0% and 2.5% of all cases were classified as inactive by SLID but as unemployed by LFS. However, the reverse -- cases that were unemployed in SLID but inactive in the LFS -- ranged from 6.2% to 8.4% of the total population.

In addition to the changes we plan to make in the use of dependent interviewing, we will be revisiting some of our decisions on how to classify particular subgroups, for example on-call workers who did not work at any time during the month (counted as inactive in the test; we think it may be more appropriate to classify them as unemployed, as we do temporary layoffs). These changes are likely to influence the behaviour of the labour force status data.

## **7. INCOME REPORTING**

The May 1993 income test included a mailout of a questionnaire booklet to respondents prior to the telephone interview, with a request to complete the booklet before the interviewer's call. This approach is similar to the one used by the SCF, although the SCF restricts its content to a one-page form (about 25 income sources). The income test had several specific content objectives, which resulted in deviations from the SCF:

- Make it as easy as possible for respondents to consult their tax records, and provide encouragement to do so. This involved making the reporting categories as comparable as possible to the tax return to facilitate transcription. Also, providing tax line numbers right on the questionnaire.

- Add specific questions to separate out income sources not usually identified in income surveys but which emerged as important items in user consultations (for example, spousal and child support; household transfers).
  
- Test a set of questions on assets and debts.
  
- Get a reading on the consistency of data collected in the labour and income interviews, and test a few interactive edits designed to improve consistency.
  
- Evaluate the impact of a nicely-designed and "wordsmithed" respondent booklet.
  
- Ask some respondent debriefing questions, including their opinion on authorizing a link to tax data as an alternative to providing the information via interview.
  
- Gain experience in the CAI design of an application that looks like a form on the screen -- very different from the labour application, which was one question at a time, with complex flows.

There was neither the time nor the funds to design the test in a way that would allow us to separate out various effects. The strategy was rather to test the full package and use observation and interviewer debriefing to interpret the results.

What did we learn? Our response rate was only 67% (compared with 75% for labour).<sup>1</sup> We concluded that it would be wiser to limit the number of questions, even at the cost of losing the one-to-one relationship with the tax form. We also concluded that it would be wiser to postpone the wealth questions until the second wave.

Among those who did respond, only 36% completed the questionnaire booklet before the interview, 17% consulted their tax records during the telephone interview and 46% did the interview on the fly. Given the complexities of income reporting, these results have encouraged us to look seriously at seeking respondent authorization for a direct link to their tax records. The earliest we could conceivably do this is 1995, so this year's income interview will be reduced in scope, to encourage a higher participation rate. Respondents' general reaction to our carefully crafted booklet was that they liked the illustrations -- but not the words. In reply to our hypothetical question on accessing tax records directly, 43% said they would agree and another 22% said "maybe". (The question asked if they would be willing to give us their social insurance number in order to perform the link. We have since concluded that this number is not technically necessary. A simple request for permission would likely result in higher acceptance rates. )

One interesting finding concerns the reporting of Unemployment Insurance (UI) income. UI tends to be under-reported in income surveys. In SLID the labour interview includes questions on the receipt of UI; the income interview captures amounts. This "double-barrelled" approach may improve the quality of reporting. The test data have been matched to the 1992 tax file, and preliminary results

---

<sup>1</sup> We expect better performance in our 1994 interviews for several reasons. The interviewers who participated in the test had no prior experience with CAI; the system was new and clearly in test mode; the respondents had been rotated out of the LFS seven months ago and were not expecting another contact.

indicate that respondents who report UI inconsistently (that is, they report UI in the labour interview but not in the income interview, or vice versa) tend to be in receipt of this income source, based on tax data. Further study will show whether the same is true for Social Assistance and Workers' Compensation, which are handled in the same way.

Regarding the CAI design, our income application had two separate streams. If the respondent had completed the booklet beforehand, the interviewer accessed a form-like questionnaire on the screen and asked the respondent to report items that had been completed. Thus, the interview was respondent-driven. We called this the *notebook approach*. If the booklet had not been completed beforehand, we could not ask the questions one-by-one, as there were 61 distinct items (39 income and 17 wealth questions) to go through, most of which would apply only to a small subset of respondents. So we grouped similar items into blocks and asked a general question to determine if any items in the block applied and if so, the interviewer would call up a screen with the distinct items on it. This was called the *block approach*. This turned out to be rather complicated and our ability to interpret results have been somewhat hampered by problems the interviewers encountered in using the application. The application for 1994 has been reduced in scope -- the number of income categories is down to 27 -- which has made it possible to use the same form on the screen, whether the respondent has prepared for the interview or not.

## **8. DATA QUALITY REPORTING**

The tests also provided us with an opportunity to develop prototypes for the data quality information we will provide on SLID when data are released. We are proposing to prepare regular reports, summarizing response information on each wave and cumulative response rates; describing tracing results; examining the



results of interactive edits; and so on (Michaud, Lavigne and Pottle; Michaud, Le Petit and Lavigne). The development of these prototypes drove home the difficulties in measuring response in longitudinal surveys, difficulties that are increased by the survey's "two phase" design.

## **9. CONCLUSION**

Test results from 1993 provided a basis for making several content decision. In addition, it was a period of learning how to function in a computer-assisted interviewing environment. Perhaps most importantly, it was a period of adaptation to "longitudinal thinking". Since the first wave of SLID labour interviews was in effect updating information collected during the preliminary interview, the development work that took place throughout 1993 was inherently longitudinal. While much fine-tuning remains to be done, the year of testing was crucial for SLID.

## **BIBLIOGRAPHY**

David Coish. **SLID Content Evaluation-The Authority Series: Supervision and Management**, SLID RP No. 93-09, Statistics Canada, 1993

Ruth Dibbs, Alison Hale. **Questionnaire Design in a Paperless Society**, SLID RP No. 93-08, Statistics Canada, 1993.

Doug Hiemstra, Mylène Lavigne, Maryanne Webber. **Labour Force Classification in SLID: Evaluation of Test 3A Results**, SLID RP No. 93-14, Statistics Canada, 1993

Georges Lemaître. **Dealing with the Seam Problem for the Survey of Labour and Income Dynamics**, SLID RP No. 92-05, Statistics Canada, 1992a.

Georges Lemaître. **The Measurement of Job Search and Unemployment in a Retrospective Setting**, SLID RP No. 92-04, Statistics Canada, 1992b.

Sylvie Michaud, Christel Le Petit, Mylène Lavigne. **Qualitative Aspects of SLID Test 3A Collection**, SLID RP No. 93-07, Statistics Canada, 1993.

Sylvie Michaud, Mylène Lavigne, Jacqueline Pottle. **Qualitative Aspects of Test 3B Data Collection**, SLID RP No. 93-11, Statistics Canada, 1993.

Yves Saint-Pierre. **Questions Relating to Social Support: Results from the SLID January 1993 Test**, SLID RP No. 93-13, Statistics Canada, 1993.