Possible Future Directions for Longitudinal Surveys at Statistics Canada

Garnett Picot, Jean-Marie Berthelot, Maryanne Webber

Abstract

The 1990s were the decade of longitudinal surveys in Canada. In partnership with policy departments and other stakeholders, Statistics Canada launched several longitudinal surveys. This paper looks at some of the substantive findings from research based on these sources, and explores some of the lessons learned about such issues as design, complexity, research capacity and survey governance. In addition to these cross-cutting issues, some thoughts are offered regarding four specific surveys – the Survey of Labour and Income Dynamics, the Workplace and Employee Survey, the National Longitudinal Survey of Children and Youth and the National Population Health Survey. The intention is simply to stimulate discussion, as background for a possible "taking stock" of the longitudinal surveys conducted by Statistics Canada. The focus of the paper is on possible adjustments to the program of longitudinal surveys that would enhance the knowledge gained from them. *The authors wish to thank Dafna Kohen, Miles Corak, Sylvie Michaud, Peter Morrison, François Nault, Lorna Bailie and Marc Hamel for their helpful comments.*

Objectives of the Paper

The overall objective is to step back and take stock of where we stand regarding the longitudinal data program at Statistics Canada. In the interest of initiating discussion regarding the program's future direction, we identify a number of issues that we think warrant attention, and propose some possible options for consideration. To achieve this overall objective, we proceed as follows: (1) provide an overview of the longitudinal survey and administrative data program at Statistics Canada, (2) identify a number of issues we think are generic to most or all of the longitudinal surveys at Statistics Canada, and (3) in the last section, focus on four particular surveys, the Survey of Labour and Income Dynamics (SLID), the National Population Health Survey (NPHS), the National Longitudinal Survey of Children and Youth (NLSCY), and the Workplace and Employee Survey (WES). For these surveys we ask, if we were to move to a hypothetical version II of these surveys, what issues would we propose considering in such a transition, and suggest alternatives for discussion purposes. The issues discussed in the last two sections of the paper are:

Issues Generic to Most Longitudinal Surveys

- Governance
- International comparability

- Research capacity available to exploit the surveys
- Survey integration
- Simplifying the surveys

Survey Specific Issues

- The complexity, design and sample size
- The panel length
- The production of cross-sectional and longitudinal data simultaneously
- Increased used of administrative data

1. The Current Portfolio of Canadian Longitudinal Surveys

Canada has a large portfolio of longitudinal surveys implemented by Statistics Canada and designed to support academic and policy research. This portfolio was developed over a roughly fifteen year time span, as research needs were identified, funding located, and surveys put in place. This endeavour was very much a partnership between the national statistical agency and policy departments, with the academic community playing a strong advisory role. The result has been the creation of seven major longitudinal surveys in Statistics Canada, described in Appendix I. In addition, a number of specialized longitudinal data sources based on administrative data have been developed and used primarily within Statistics Canada. The longitudinal data sources are funded in some cases by Statistics Canada but many are dependent on external funding. Some are conducted on a cost-recovery basis, with funding coming directly from a policy department. Others have received funding through the Policy Research Initiative, established to fill data gaps at the federal level that are cross-cutting in nature.

As in many countries, the surveys were not developed based on an overall strategic plan. Rather, they were developed in response to the diverse research needs of a number of policy departments and academic disciplines, and as a result cover a very wide range of topics. The chronology and subject matter of the major longitudinal surveys are as follows:

- 1978: National Graduates Survey (NGS)
  * Labour market and educational outcomes for college and university graduates
- 1993: Survey of Labour and Income Dynamics (SLID)
  * Poverty dynamics, family formation/dissolution, labour market dynamics
- 1994: National Population Health Survey (NPHS)
  * Determinants of health, socio-economic correlates, health care utilization, dynamics of health and illness
- 1994: National Longitudinal Survey of Children and Youth (NLSCY)
  * Child health, development, and well-being and their correlates
- 1999: Workplace and Employee Survey (WES)

* Linked workplace and employee survey covering industrial relations and H.R practices, industrial organization (innovation, technology adoption, firm strategies, etc.) and workers labour market outcomes.

- 2000: Youth in Transition Survey (YITS)
  * Determinants of the major educational and work transitions for two cohorts: 15 year olds, and an 18-20 year old group.
- 2001: Longitudinal Survey of Immigrants (LSIC)
  * Social and economic integration of the 2001 arriving immigrant cohort

In addition to these seven major surveys, more specialized longitudinal data sources based primarily on administrative data have been created and used for mainly internal Statistics Canada research purposes. These include:

- Longitudinal Immigrant Data Base (IMDB)
  * Created for and used by Citizenship and Immigration Canada, based mainly on administrative (taxation) data on earnings and income of immigrant cohorts since 1980
- Longitudinal Administrative Data Base (LAD)
  * Administrative (taxation) data on earnings and income dynamics for all Canadian families
- Longitudinal Employment Analysis Program (LEAP)
  * A longitudinal company survey used for employment dynamics studies
- Census of Manufacturers Longitudinal Panel
  * A file longitudinally linking survey data on manufacturing establishments, and used for studies of employment dynamics, productivity, effects of technological change, etc.
- Worker Longitudinal Survey (WLF)
  * A longitudinal file of employed Canadians based on taxation and Record of Employment data, used for studies of worker turnover, worker displacement effects, etc.
- Intergenerational Income Data (IID)
  * Primarily taxation data used for the study of intergenerational income mobility
- Hospital Person-Oriented Information Database
  * Hospital morbidity data principally for health services research

The portfolio of longitudinal surveys available in Canada is now significant. The enormous amount of detail in these surveys, in part because of their longitudinal nature, means that producing public use files is problematic. To facilitate access for academic researchers, a suite of Research Data Centres has opened on thirteen university campuses through a partnership involving Statistics Canada, universities and SSHRC/CIHR (social science and health granting councils). In addition, remote access opportunities are available for many of the more complex surveys. Academic researchers have access to the confidential data for the seven major longitudinal surveys (and other cross-sectional surveys) through these means. Similar arrangements have

recently been made with policy department researchers. They too now have access to the data for research purposes, while protecting the confidentiality of the data under the Statistics Act, the legislation covering data confidentiality in Canada.

Although the Research Data Centres are quite new, opening over the past four years, the volume of use is impressive. Nearly 1,100 researchers are currently using data sources in the RDCs. This number has been rising rapidly over the four years of their existence, and is expected to continue to rise. About half of the research projects initiated since the launch of the RDCs have involved the use of longitudinal datasets. Many other researchers access the sources through remote access facilities or in partnership with Statistics Canada researchers.

Longitudinal surveys are, of course, long-term investments, and returns to the investment increase with time. Since many of the Canadian surveys are quite new, with only a few of the major surveys having a history of ten years or more, it may be too early to assess the benefits. Any "taking stock" process must focus on the value of what we have learned from these data sources.

2. Substantive Insights

The ultimate return to the investment in longitudinal surveys and administrative data files must be judged by the knowledge gained. As longitudinal surveys mature, and an increasing number of researchers become familiar with them, returns increase. Nonetheless, valuable findings have already been produced, both from the longitudinal surveys and administrative files. Through longitudinal surveys we now know that it is not only job loss that triggers a descent into poverty, but that family formation and dissolution play a major role in the movement into and out of low income. We have learned that persistently poor people, a major concern of much pubic policy, are concentrated in five groups, again focusing the attention of policy analysts. In the Canadian context, recent immigrants, aboriginal peoples, the disabled, and perhaps surprisingly, some groups of the unattached, along with single mothers, account for most persistent poverty. Together these groups account for only one-quarter of the population, but almost two-thirds of those with persistent low income. Such observations have changed the way that policy analysts think about poverty.

Longitudinal administrative data have also taught us that intergenerational transmission of poverty, while important, may not be as great as we once thought. Children from poorer families are more likely to be poor as adults than those from richer families, but this outcome is anything but certain. Furthermore, the likelihood of moving from poverty as a child to higher income levels as an adult is greater in Canada than in the U.S. or the U.K. In this regard, we more closely resemble the Scandinavian countries. Canada seems to have developed a set of institutions and practices (for example, the education system, labour market institutions affecting income inequality,

early childhood development practices) that are conducive to greater equality of opportunity.

Knowledge such as this is important to advance "evidence-based" policy designed to combat persistent low income and ensure that children who are in low-income families have an equal opportunity to lead productive lives. Institutions change, and the more longitudinal data teach us about poverty determinants and transmission, the better able we are as a nation to promote desirable outcomes.

Insights are of course not restricted to poverty dynamics. We have learned that specific parenting behaviours have direct consequences for child development. Children living in homes where physical punishment is used show more aggressive behaviour than those living in homes with no physical punishment. Preliminary largely descriptive analysis suggests that children raised in an environment of authoritative parenting (a warm and nurturing relationship but one that sets firms limits) are least likely to exhibit signs of vulnerability. This parenting style is positively related to better behaviour and school performance, and is also linked to a decrease in developmental problems. The parenting styles are grounded in theory and converted to empirical measures. The other parenting styles are *authoritarian* (highly controlling, with an absolute set of standards) and *permissive* (overly nurturing, with few behavioural standards and extreme tolerance of misbehaviour). Preliminary analysis suggests that they are associated with less positive outcomes. We await further analysis that more carefully assesses the causal magnitudes and directions to confirm these findings.

In the area of firm dynamics, people have asked why some firms grow faster than others. The important role of product and process innovation has been highlighted through the use of longitudinal surveys. Similarly, the role of "creative destruction" — the death of less productive firms, to be replaced by the more productive — in a country's productivity growth is now much better understood due to longitudinal studies. A nation's productivity growth is not only driven by existing firms "working smarter," a substantial share of the growth can be ascribed to creative destruction — an important insight for analysts concerned with the sources of economic growth. Firm and worker dynamics research has also taught us that job creation and destruction, along with its worker counterparts, hiring and layoffs, are driven largely by idiosyncratic events occurring to particular firms. These events (job creation/destruction and layoffs) are not primarily driven by economic change at the economy wide level (recessions) or at the industry level (e.g. trade effects), but rather issues associated with changes in market share of birth/death of individual firms within a market. The reallocation of economic activity among firms within markets is behind most job and worker dynamics.

Tracking health outcomes of Canadians has led to significant results as well. A recent study focused on the tendency of immigrants to be in better health than Canadians

when they arrive in Canada. Tracking the health of immigrants and Canadian-born individuals from 1994 to 2003, the study found that this "healthy immigrant effect" tends to diminish, as their health status converges with that of the general population. This more rapid deterioration in health was particularly strong among non-European immigrants, as they were twice as likely to report some deterioration in their health as Canadians. An increase in their body mass index (weight gain) was associated with this deterioration in health, which led to more doctor visits.

Further examples of substantive findings from the longitudinal surveys can be found in Appendix II (forthcoming)

Key to any possible review of longitudinal surveys is an exploration of the significant insights generated in a wide range of areas including health, workplace practices, low income and social assistance dynamics, firm dynamics and child development. As data sources improve and accommodate the testing and development of new theories, researchers are examining social and economic phenomena in a more complex manner. For example, both the causes and consequences of poverty are multi-faceted, involving health outcomes, labour market events, family formation and dissolution, access to education and training, early childhood development issues, and the design of the social transfer system. With the advancement of longitudinal data sources, researchers now contemplate empirically testing new and complex hypotheses regarding the causes and consequences of poverty. It is reasonable to ask whether the current generation of longitudinal surveys is up to the task, or is more integration of survey content needed to support such a new agenda. This focus on poverty issues is demonstrative only. The need for a discussion regarding knowledge gained, and our readiness for future advancements, applies to all domains touched by longitudinal surveys.

3. Taking Stock

This paper sets out a few ideas as a beginning to a "taking stock" process involving the longitudinal surveys conducted by Statistics Canada. Although no formal overall strategic evaluation of Statistics Canada's longitudinal survey program has been initiated, the experiences and lessons learned at the longitudinal conference will hopefully help formulate such a review. The ideas presented in this paper have not been discussed in any detail with the major policy research or academic user groups. These steps are obviously crucial to any review process. As noted earlier, many longitudinal surveys are funded from external sources. Any review process, were it to be created, would be a partnership between the funding agencies, Statistics Canada and the user groups.

A decade or more has passed since the launch of some of these surveys. The time is right to step back and consider their future. We restrict ourselves to comments on the existing surveys. The longitudinal conference has a number of proposals from other groups for new surveys, so for the purposes at hand we will not comment on potential

new data sources. The paper is divided into general comments on the overall survey program, and survey-specific comments. We restrict the latter to the more established surveys, notably SLID, NLSCY, and NPHS. Space, and the newness of the other surveys, suggests commentary on other data sets would be premature. The only exception made is for WES, because a decision has been made to wind down that survey in its current form, and it seems an opportune time to discuss lessons learned.

The issues addressed in this paper are driven primarily by one question: "What major adjustments could one make to the program of longitudinal surveys in order to increase the quality and quantity of the substantive insights derived from the research based on these surveys?" Such adjustments might mean significantly altering a host of characteristics of the surveys. As noted in the introduction, we feel the most important issues include:

Generic Issues

- Governance
- International comparability
- Research capacity available to exploit the surveys
- Survey integration
- Simplifying the surveys

Survey Specific Issues

- The complexity, design and sample size
- The panel length
- The production of cross-sectional and longitudinal data simultaneously
- Increase used of administrative data

4. Issues Generic to Many Longitudinal Surveys

4.1. Broadening Governance

All the longitudinal surveys under discussion had significant input from both the policy and academic research communities during the developmental stage. Expert teams were created to guide survey development and implementation, and considerable effort was made to ensure input from potential users. But following the launch of the surveys, communication with academic researchers regarding future survey content and direction was, in most cases, much reduced. This is unfortunate because the research community needs to truly influence decisions regarding content, design and processing. Most Canadian large longitudinal surveys struggle with this issue. Advisory committees in some cases have been established, meeting perhaps annually or semi-annually, but it is difficult for members to remain current regarding survey changes and their potential effects. Better means of obtaining continued input from committed members of the research community are needed.

Most Canadian longitudinal surveys have a somewhat unique development history. In Canada, policy departments have funded most of the major surveys and have therefore quite reasonably played a major role in the content decisions, along with significant input from the academic community at inception, as noted earlier. Given the funding structures in Canada, as compared to those of the Michigan, British and German panels, for example, where the funds flow through the scientific granting councils, it is perhaps not surprising that the partnerships have developed somewhat differently in Canada.

It is now time to encourage a greater role for the academic research community in on-going developmental decisions. In addition to hopefully ensuring that content and design directions taken are in concert with research directions, such increased input may lead to a heightened sense of ownership and responsibility for the surveys by the academic community, thus increasing use and dissemination of results. Academics are the principal users of the more complex longitudinal surveys, simply because most research capacity resides in the universities. There are a number of options one might consider to achieve these goals.

One possibility is to create a high-level on-going steering committee for each longitudinal survey, which would include members from the academic community, the relevant policy department and Statistics Canada, with an explicit and formal role to oversee changes in survey content and direction. An alternative model would be the creation of a research institute around individual or groups of longitudinal surveys. The form of such institutes could range from a "virtual" institute (a network of researchers conducting research with and providing input on particular surveys) to a number of small, perhaps university-based, institutes with research programs that draw heavily on the longitudinal surveys, and hence are heavily involved with developmental decisions. Changes in the funding process for academic research, moving towards funding of clusters or teams with research programs in thematic areas where longitudinal data are available, could also be of assistance. Increased involvement with and input into the longitudinal data sources could be one component of such a funding process. Yet another approach would be to have a Statistics Canada researcher with strong research credentials and strong ties to the academic research world be the manager, or co-manager, of the survey.

In short, we believe greater input to survey development from the academic community on an on-going basis is desirable. Combined with the on-going involvement of the policy departments, it would help maintain the policy and academic relevance of the surveys. It would also reinforce the strength in the current Canadian model, built around the involvement of the three communities…academic, policy and statistical…thereby contributing to the academic and policy-relevance of the surveys. The ideal is to strike a balance, engaging both the academic and policy communities with the statistical agency, and finding mechanisms that allow decisions on the future

of the survey to be made in a harmonious environment, without losing sight of the importance of stability in a longitudinal survey.

4.2. International comparability

Studies of social and economic processes can be enlightened by a capacity for international comparison. To better place Canada within the global spectrum, international studies are extremely useful. Indeed, the Cross-National Equivalence File (CNEF) is a beginning: it offers longitudinal income dynamics data for Canada, the United-States, the United Kingdom and Germany.

Similar opportunities await analyses based on other longitudinal surveys. There have been developments regarding internationally comparable content in health surveys, and the direct measurement of literacy and skills among adults and high school students. The Programme for International Student Assessment (PISA), measuring the reading, math and science skills of 15 year olds in over 30 countries, is linked in Canada to the Youth in Transition Survey. But most Canadian longitudinal surveys have been developed in isolation. Increased attention to this dimension of survey development would encourage researchers to conduct internationally comparative research, greatly increasing the value of the research output. However, the development of internationally comparable, culturally relevant items is time-consuming and costly, and realistic goals are required. Attempts to produce comparable data sets after the fact (such as CNEF) can help identify important issues and areas of focus in future design and development phases.

4.3. Research Capacity

If important and relevant insights are to be generated through longitudinal data, it will be done by skilled researchers. If insufficient research capacity is brought to bear on longitudinal surveys, a shortfall of relevant findings will result. This issue is of importance in Canada for at least three reasons. First, we are a small country with a relatively few empirical researchers in most disciplines compared to, for example, the U.S. Canadian researchers have more or less the same data infrastructure at their disposal as their U.S. counterparts. Hence intensity of use for any given data source will be less.

Second, analytical techniques are becoming more complex. This can limit the segment of the research community that chooses to embark on the use of the surveys, unless education and training is implemented to match the rise in methodological complexity.

Third, the data themselves are complex and require a considerable up-front investment for use. Not all qualified researchers are willing to make such an investment.

All of these factors affect the research capacity available to exploit the longitudinal data, and should be considered in any "stock-taking" discussion. In our view, Canada is

faced with a significant research capacity problem regarding the use of longitudinal surveys. It may vary from discipline to discipline, but in general the supply of data available to support research has increased much more quickly than the available research capacity. In such an environment, means of linking researchers more closely to the data sources, such as outlined above in the "governance" section, should be considered. Funding initiatives, promoting clusters and research teams in thematic areas where longitudinal data are available (e.g. child development, human capital development over the life-course, worker and firm dynamics) with a significant training component, could also increase capacity.

Finding means of making the data more accessible to researchers outside of Canada would also assist in this process. Concerted efforts to increase U.S and European researchers' knowledge of some of the more novel longitudinal data sets in Canada could bear fruit.

Finally, increased training could be beneficial in the longer run. In a partnership between Statistics Canada and SSHRC (social sciences granting council), some summer training programs have been created, but it is not clear that this level of activity if sufficient to introduce real change.

4.4. Scope for Survey Integration

In a few short years, several longitudinal surveys were launched by Statistics Canada, in partnership with policy departments and academic researchers. They were designed to meet specific needs and, while there was on-going communication between the various survey teams and stakeholder groups, there was no integrative, strategic plan.[1] If we take a step back and look at the collective research potential of these new surveys and datasets, we can ask ourselves whether they form a logical whole. Are there overlaps and significant data gaps?

There are several ideas for new longitudinal surveys currently at the discussion stage, in the fields of health, ageing, and retirement. As we discuss extensions to some longitudinal surveys and possible new ones, it is an opportune time to also consider whether some surveys could be integrated or become more focused. For example, where two surveys have the same or overlapping target populations could they be combined? Where two or more longitudinal surveys are focused on different stages of the life cycle, could respondents "graduate" from one survey to another, thereby increasing the period of observation and the richness of the available information? The following ideas are intended as illustrative and are offered to stimulate discussion on the scope for integration across surveys. As in Section 5, the issue of funding is held in abeyance in this discussion. Some of the integration ideas would only be relevant if a decision were taken to fund a new cohort.

[1] There was one attempt at integration: NLCSY and NPHS did share sample, but only for the first cycle. The practice was abandoned, partly because it created too much response burden.

4.4.1. National Longitudinal Survey of Children and Youth/Youth in Transition Survey/Survey of Labour and Income Dynamics

NLSCY is the only longitudinal survey that addresses issues of childhood development across ages including infancy, preschool, school age, and adolescents. It is designed as a longitudinal study tracking child development until adulthood, to provide information on factors impacting outcomes over the early part of the life course. The coverage of child outcomes is comprehensive including parent and teacher reported measures, standardized tests, and youth self-reports in areas such as health, language, cognition, affiliation with peers, social and emotional functioning, risk behaviours, school functioning, and specific abilities such as mathematical skills. Moreover, a tremendous amount of information is gathered about the child's environment including the community, school, family, peers, and participation in extra-curricular activities at multiple time points.

YITS in contrast ranges from the adolescent period to adulthood and focuses on transitions to school and work. In 2004, the YITS sample included individuals aged 19 and 22-24. In the same year, the NLSCY respondents (longitudinal component) were aged 12-22. The survey objectives are different but there is some content overlap between the two. Could NLSCY focus on the early years, and could respondents reaching a certain age "graduate" into YITS? What would the costs and benefits be of such a decision?

Similarly, respondents could "graduate" from NLSCY through YITS into SLID. With some content modifications in SLID, aimed at covering social support and attitudes/values, the resulting data would have some commonalities with BHPS.

4.4.2. Youth in Transition Survey/National Graduates Survey/Enhanced Student Information System

The Enhanced Student Information System (ESIS) is an integrated database of all postsecondary student administrative records. ESIS is a new and, in due course, will offer the possibility of linking student records across institutions and jurisdictions, and through time. ESIS will provide detailed information on pathways through postsecondary, student persistence and time to completion for all postsecondary students. In contrast, the Youth in Transition Survey (YITS) offers information on transitions of all youth, whether they go to postsecondary or not, whether they complete high school or not. The National graduate Survey (NGS) covers the short-term labour market outcomes of postsecondary graduates, many of whom are older than the YITS target population. ESIS and NGS are linked, in that ESIS provides the institutional frame from which a sample of graduates is drawn to complete the NGS.

4.4.3. National Population Health Survey/National Longitudinal Survey of Children and Youth

NPHS included a sample of children from the outset; the youngest are now 12 years old. If a decision were taken to start new cohorts for these two longitudinal surveys, would it be worth considering starting NPHS at an older age, while increasing the health content in NLSCY? That would perhaps ensure that the health information on children is available for the same sample as other variables on child development.

4.5. Simplifying the Surveys

The notion that overly complex surveys limit the use of longitudinal surveys, the timeliness of the output, and possibly even data quality, is a recently developing theme among Canadian surveys. These surveys are inherently complex, so the real question is: can they be simplified somewhat without losing their analytical value? For example, in most cases in the Canadian context, longitudinal surveys are also required to meet cross-sectional objectives. This approach allows the surveys to meeting monitoring and trend reporting (cross-sectional) objectives, as well as longitudinal cohort analysis goals. This approach forces content compromises and adds complexity to weighting; the need for cross-sectional representativity creates competing pressures in the survey decision-making environment.

A second method of simplification centres on cohort selection. A panel with a very broad target population (obviously important for cross-sectional results) will entail complexity in the survey content, which in turn contributes to data processing delays and long learning curves for researchers. Conversely, a narrowly-defined cohort (for example, a birth cohort instead of a panel covering all children under age 12) will tend to reduce complexity, and focus resources on particular cohorts.

Finally, the variables themselves may be simple or complex. For example, longitudinal surveys may all be similar in that they contact respondents periodically – every year, every two years or so on. But they differ with respect to the type of picture they try to create. Some longitudinal surveys offer in essence a succession of snapshots on respondents. This is much simpler than a survey that attempts to create a "movie" of continuous measures, such as spells of job holding or school attendance. These two types of approaches can entail important differences in processing time and in accessibility for researchers.

The theme of reducing complexity is taken up in more detail in the next section, which addresses survey-specific issues.

5. Survey Specific Discussion

Three of the major longitudinal surveys have existed for a decade or more, the Survey of Labour and Income Dynamics (SLID), the National Population Health Survey (NPHS), and the National Longitudinal Survey of Children and Youth (NLSCY). In addition, because it was funded through a process requiring periodic budget reallocation, the current version of the Workplace and Employee Survey (WES) will be terminated following the release of the eighth year of data in 2008-09. This section will focus on these four surveys, and ask, if we were to develop a hypothetical version II of each of these surveys, what issues would be paramount in the revisions to the surveys. To facilitate discussion, some suggestions are put forward regarding these issues. It is well known that stability in longitudinal surveys is crucial, since the returns to the data increase with duration. The balance between necessary change and long-term stability would itself be a major issue if revisions were to be made to any of these surveys.

5.1. Issues to Address in a Hypothetical SLID II

We believe that the potential for important research results from the Survey of Labour and Income Dynamics (SLID) could be significantly enhanced through a number of changes. The major issues that we think require attention include:

5.1.1. Broadening the Content, including the Use of Modules

SLID likely has more focused content than most general purpose household panels, notably the PSID (Michigan panel study), SOEP (German household panel), and BHPS (British Household panel). SLID is very much an income survey, given its need to produce the official annual cross-sectional income estimates. The labour side of SLID focuses on dynamics, for example, employment and unemployment spells, flows into and out of the labour force, etc. We propose a broadening of the content, to bring it closer to other similar international surveys, such as those mentioned above. Analysis is strengthened tremendously when international comparisons are possible. Some content blocks that have been considered in the past and could be considered again to enhance the domestic and international analytical strength of the survey include health, wealth, housing, aging and retirement, and a strengthened immigration block of questions. Most of these topics are covered in the three surveys mentioned above. Additional health questions could assist in increasing the relevance of SLID for issues around work and health. A wealth module would broaden the "welfare" concepts employed in economic studies. Additional content on issues around aging and retirement would increase the relevance of SLID for current issues, as would a strengthened immigration set of questions (combined with over-sampling as proposed later).

Rotating content "modules" would allow the survey to remain current in its content, while maintaining a core of questions, including the "outcome" variables that one wants to track longitudinally. Whether one would maintain the current level of detail on the start/end dates, characteristics of jobs, social assistance and other spells lengths

could be considered. Given the relative lack of use of these data, and the difficulty in establishing accurate start and end dates, perhaps this content could be reduced.

5.1.2. Increasing the Panel Length to Support More Longitudinal Research

By international standards, SLID has one of the shortest panel lengths at six years. In a hypothetical SLID II, we would propose a much longer panel length. Surveys such as SLID are often used to assess adjustments over time to some event, such as job loss, divorce, welfare exit, retirement, birth of a child and so on. Let us take the example of adjustment to job loss. Data for a number of years prior to the loss is required to assess the difference in, say, earnings between the workers involved in job loss, and some reference or "control" group. Furthermore, it is well known that some adjustments, such as declining earnings, start well before the job loss takes place (e.g. through declining hours of work). Both short and long run earnings adjustment following the job loss are of interest, so 6 to 8 years of data following the event could be required. Taken together, this suggests that to reasonably study the issue of earnings changes stemming from job loss, 10 to 14 years of panel data are required. Many "adjustment" issues require such a panel length. It is interesting to note that many of the interesting insights gained from longitudinal analysis in Canada are based on administrative longitudinal data, whose advantage are very long panel lengths, as well as large samples. Their disadvantages are, of course, a lack of co-variates.

In addition, issues dealt with using SLID often have a cyclic characteristic, and hence assessing outcomes (e.g. earnings, income, employment, unemployment effects) related to particular events (divorce, change in health, job loss, child birth, retirement, etc.) requires data over one complete cycle at the very least, to ensure that the result is not unique to some particular phase of the business cycle. Recently cycles are typically 10 years in length and ideally one would have two cycles of data.

Finally, some longitudinal household surveys are now capable of addressing intergenerational transmission issues, an issue of very significant policy concern requiring perhaps 25 to 30 years of panel data. Panel length would be one of the most important issues in any discussion of a SLID II.

5.1.3. Separating the Production of Cross-Sectional and Longitudinal Data

While initially conceived as primarily a longitudinal survey, SLID became the official source of annual cross-sectional income statistics when the Survey of Consumer Finances was cancelled. Due to its official status, the production of the cross-sectional income data has likely become a more important consideration than the longitudinal component of SLID. If SLID II were to shift to longer panels, keeping the data source cross-sectionally representative every year would be a tremendous challenge. In such an environment, one might want to consider separating the two. Also, the initiation of a new longitudinal panel every few years (say every 5th year) would allow for the introduction of new cohorts of immigrants to be included in the panels, and would

allow for cross-cohort comparisons, an important issue if one wants to separate cohort, period and age effects.

5.1.4. Exploiting Administrative Data to an Even Greater Extent

Administrative taxation data already play a significant role in SLID, as the annual income data for the majority of the records is drawn from the T1 tax file, with very high levels of respondent consent. With similar respondent consent, taxation data could be used to extend the panels back in time. This could yield good longitudinal earnings and family income data. Such an approach could provide useful analytical insights based on income data extending back at least 20 years from the beginning of the panel. Essentially, it would combine a longer longitudinal panel of income data with the covariates and outcome data that one can only obtain from survey data (e.g. education, details on jobs/retirement, health covariates, etc.).

The production of cross-sectional income statistics primarily from the administrative taxation data could also be considered. One possibility would be to use the Labour force Survey to seek respondent permission to access tax information, as is currently done in SLID, as well as to provide important characteristics for the production of cross-sectional income data (e.g. family composition, education, full/part-time status, family status, etc.).

5.1.5. Over-Sampling Some Populations

Even though SLID has a very large sample by international standards (30,000 households), sample size prevents important longitudinal research from being conducted. The economic assimilation of recent immigrants (and the visible minority population) has been, and is likely to remain, one of the most important policy research issues in Canada. However, the current sample of immigrants in SLID allows for only the most basic analysis. Over-sampling immigrants would address this important issue. Economic and labour market issues surrounding the off-reserve aboriginal population is another area that could benefit from over-sampling. The use and dynamics of social assistance and other government transfers is another potentially important research area for SLID, but the sample of social assistance leavers, for example, prevents significant analysis. Finally, aging, retirement and the associated work and health issues, an obviously important future research area, might require some over-sampling of the older population. In short, SLID II could over-sample particular populations to enhance the research output of the survey.

5.2. Issues to Address in a Hypothetical WES II

The Workplace and Employee Survey (WES) does not fit the mold of the traditional household survey panel survey, since it contains both firm and worker data. However, it is used to address issues common to household panels, such as the effects of training and education, immigrant economic assimilation, flexible work patterns, and family friendly human resource practices, Also, it was designed to allow the determinants of

worker labour market outcomes to be extended beyond those covered in the traditional human capital model, in particular focusing on firm characteristics. Because of such overlap with the issues covered by household panels, it was decide to include the WES in this discussion. All of these topics have received attention, and will continue to do so, as the WES remains quite immature (with only four years of data currently available). In practice, the firm component of the WES has also received considerable attention from the industrial relations, human resource management and business research communities.

In a budget reallocation process, the current Workplace and Employee Survey will be cancelled after eight years of data are available (2008-09). This survey represents the only large scale longitudinal survey of firms, and we believe that it is important to maintain the infrastructure that has been developed. Without such a survey, knowledge of how firms adjust to the changing domestic and international competitive environment, and its effect on the firms' workers, would be unavailable. We believe the following issues should be considered if a WES II were to be developed.

5.2.1. Narrow the Content to Focus on Issues Related to Firm Adaptation and its Effect on Worker and Firm Success

WES was originally created to provide information on both outcomes of workers and firms measured longitudinally. We suggest concentrating on measures of change in the firm, such as business strategies, competitive environment, product and process innovation, training strategies, other H.R. practices, all variables that are currently in WES. Measures of change at the individual worker level (e.g. changes in wages as one moves from union to non union environment, or from small to large firms) would be dropped. Workers would not be tracked longitudinally, reducing the complexity and cost of the survey. The focus would be squarely on change in the firm and its effect on firm outcomes (profitability, employment growth, market share, etc.) and aggregate worker outcomes measured cross-sectionally (e.g. wages, training, etc. controlling for human capital characteristics). Only firm fixed effects could be introduced in the analysis; introducing worker fixed effects would not be possible, since the longitudinal worker data would not be available.

5.2.2. Move to the Company Rather than Establishment Level

Currently WES tracks establishments (locations) longitudinally on the firm side. This was done because locations are easier to track longitudinally than are companies (that can merge and split), and because the focus was in part on the activities of establishments that influence worker outcomes, such as technology use, H.R programs and collective bargaining arrangements. These can vary from location to location within a company. However, the major drawback of the current design is that financial outcomes are difficult to measure at the location level, and really must be measured at the company level. If the emphasis is shifted to firm adaptation and outcomes as

proposed, this becomes critical. Hence, one might consider a model where the company, rather than the location, is the unit of analysis.

5.2.3. Increasing the Panel Length

The same arguments that were introduced above to support an increase in the panel length in SLID also apply here. Adjustments to any major events occurring to firms or their environment may also take a number of years, and business cycle effects are often important. One would therefore want to have outcome measures for different segments of the cycle, to ensure that a substantial proportion of any effect measured was not just a cyclical effect. Extending the panel beyond the eight years planned in the current WES would be desirable.

5.2.4. Increased Integration with Administrative Data

In a WES II, changing the unit of analysis to the company rather than the location would allow effective use of administrative financial data. The sample of companies could be selected so that T2 corporate tax data could be used, thereby reducing response burden and cost, and improving the financial data available. This would allow not only current financial variables to be added to the file, but a history of employment and financial data could be added by integrating any new sample of companies with the existing LEAP-T2 data base, which provides longitudinal data at the company level back to the early 1980s. Information on worker separations (quits and layoffs) are also available at the company level, and could be linked to the new sample. Taking advantage of the administrative data available at the company level would reduce response burden and increase research opportunities in any version II of the WES.

5.3. Issues to Address in a Hypothetical NLSCY II

The NLSCY filled a major data gap on children and is supporting a very robust research program. Still, after ten years in the field, ideas have emerged on ways to make the survey even stronger. There is work underway on possible future directions for NLSCY. The ideas presented below are intended as illustrative and not at all a complete reflection of the thinking to date.

5.3.1. Adjusting the Content

The survey content on child care could be strengthened. The requirement of child care and the impact of changing family policies are a reality for Canadian families yet no national level child care data has been collected in Canada since 1992. The topic is very much in the public eye. Additional content in the NLSCY would support research on the impact of various family policies on both children and families as well as the impact of different types and qualities of child care arrangements on child development and longer-term outcomes.

Physical health measures constitute another gap. The biological mechanisms and processes of development are at the forefront of research and are being implemented in

many health surveys internationally. These measures are not only being collected in adult but in child studies as well and many clever and non-intrusive measures have been developed to collect such measures from children (e.g. blood from baby teeth that have fallen out, measures of salivary cortisol, etc.). Physical and biological measures are important indicators of the biology-environment interactions that can influence development.

Other possible new indicators are: unmet health service needs for children, observer ratings of the child's home environment, additional youth outcome measures (not only mathematical ability).

Requesting permission to link to various administrative sources could add significantly to survey content without increasing burden. For example, linking to tax data would provide income after tax, which is more appropriate for low income studies. A link to school transcripts would provide direct measures of school performance.

Finally, some existing indicators are not based on established methodologies or scales. An NLSCY II would afford an opportunity to adjust content based on experience, and to scale back on content in some areas to allow for the addition of the types of measures noted above.

5.3.2. Simplify the design by selecting a birth cohort

The original NLSCY sample covered a twelve-year age spectrum. The content varies substantially by age so, in any one collection cycle, there are in fact several surveys going on in parallel. In addition to the content changes needed for different age groups, there is the added complexity of evolving content to meet changing policy requirements.

NLSCY's wide age band made sense in the mid-1990s, when there was a dearth of information on children but, with ten years of NLSCY data now in hand, some thought could be given to a simpler approach.

One possible approach that would simplify the design would be to select a birth cohort and follow it through to adulthood. The content would evolve as the children age, but in any one cycle, the focus would be on a single age group. One drawback of this idea is that policy interests evolve and can shift over time to focus on different age groups. To ensure that data are available for age groups of emerging policy interest, new panels – perhaps every five years or so – would be important. Overlapping panels are more complex than a single cohort followed through time but still simpler than the existing NLSCY model: in effect, a compromise.

5.3.3. Panel length

The intention at the outset of the survey was for respondents to exit the survey at age 25. As noted earlier, some thought could be given to actually ending the survey at an earlier age, say 18, on the understanding that respondents would "graduate" to the Youth in Transition Survey. Alternatively, NLSCY could continue, without merging with another survey. The life of the panel could perhaps be extended but with less frequent interviews. In this event, the objectives would need to be clarified. If sample "graduated" from NLSCY to YITS, the objectives would by definition be mainly around work and school transitions. Would this make best use of the explanatory data collected over the years in NLSCY? What about other youth outcomes?

5.3.4. Interview every year

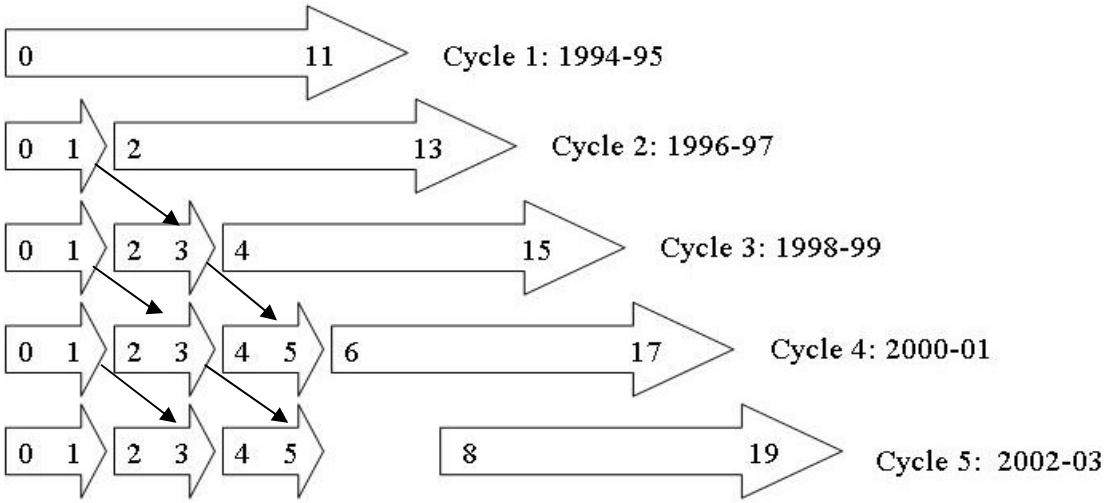
Currently, NLSCY interviews are conducted every two years. There may be some advantage in interviewing every year in the case of very young children, when development is occurring very rapidly. In fact, the frequency of the interviews could be linked to age or developmental milestones. Interviews in early adulthood could perhaps occur less frequently than every two years.

5.3.4. Sample Size

The survey design for the NLSCY, although nationally representative, does not allow for the clustering of subjects to examine nested effects. For example, the effects of living in a certain city or community or effects of different school factors on child, youth and adolescent outcomes cannot be determined in a methodologically rigorous manner.

5.3.5. Separating of longitudinal and cross-sectional objectives

As in SLID, NLSCY is not only a data source for research on determinants and outcomes, it is also being used for monitoring purposes. In particular, the federal and provincial governments have agreed on a suite of indicators regarding child well-being, and a number of these depend on NLSCY data for the 0-5 age group. As a result, the survey also includes a new cohort of children ages 0-1 each cycle, but these children are followed only until age 5. This is an important application of the data but adds complexity to the survey; it is quite a challenge just to grasp what the coverage and target population of NLSCY really are. The diagram below illustrates design. Cycle 1 covered the initial 0-11 cohort. In Cycle 2, a new cohort aged 0-1 was added. Ultimately, this new cohort was dropped after Cycle 4, at age 4-5. The design complexity translates into multiple weights. Any new researcher tackling the survey faces a steep learning curve. One idea would be to separate the long-run component from the short-run component – separate survey teams and dissemination strategies could potentially increase use.



5.4 Issues to Address in a Hypothetical NPHS II

5.4.1. Background on NPHS

The National Population Health Survey (NPHS) is a biennial survey which collects information on the health of the Canadian population since 1994. It was designed to provide cross-sectional estimates for population health monitoring but also as a longitudinal survey to study the path or processes that lead to good health. The NPHS had both, a household and an institutional component with a longitudinal sample size of around 20,000 individuals. It was planned that NPHS would include direct measures, such as height, weight, blood pressure and some blood sample, but the budget never allowed for that component to be implemented.

As with the NLSCY and the WES, the NPHS was an innovative survey program, addressing new subject matter topics and breaking new ground. It filled a significant data gap, since the last national health survey was the 1978/79 Canada health survey. It was also the first National survey to cover the institutionalized population about which little was known. NPHS also included new and innovative content based on the population health framework, such as social support, sense of coherence, self-esteem, mastery, the Health Utilities Index, stress and mental health questions. NPHS was also the first national health survey to request an informed consent from respondents to link their survey information to their administrative records for population health studies. Statistics Canada has access to hospital morbidity database (i.e.: all hospital discharges) and vital statistics. Linking survey responses to these data sources allows for the study of the determinants of hospitalisation and of mortality. Statistics Canada is currently discussing with provincial ministries of health and the Canadian Institute for Health Information to gain access to physician billing and prescription drug data. Gaining access to these data holdings will significant improve our capacity to conduct health services research with NPHS.

NPHS successes created the impetus to establish an ambitious population health monitoring program, the Canadian Community Health Survey. This survey was launched in 2000 with a sample size of 130,000 individuals for broad population health monitoring at the health region level in the first year of a 2-year cycle and a sample of around 30,000 for an in-depth focused content on the year 2 of the survey cycle. As a result, NPHS became solely longitudinal in the year 2000.

There are currently 5 cycles (8 years of follow-up) of data available from NPHS. The longitudinal component of NPHS has been used in many research projects to study the impact of a variety of determinants of health and to look at health trajectories. A simulation model was developed, using NPHS trajectories and CCHS cross-sectional data, to project obesity rate increase, its impact on acute myocardial infarction and the needs for cardiac procedures. This simulation model is a unique tool that can be used by policy makers to estimate the potential impact of public health interventions through a series of scenarios.

5.4.2. The need for a NPHS –II

There are significant age, period and cohort effects in health. For example, communicable disease which was the main causes of death in the early 20th century has been replaced by chronic disease in the 21st century. The teenage generation of males of the Second World War had the highest rate of smoking ever observed in Canada, while women pick-up smoking in the 50's. Obesity, which was more prevalent in women in 1994 is now more prevalent in men. Are the dynamics underlying the increase of obesity observed between 1994 and 2002 still at play? Are they accelerating or slowing down?

With the aging of the population, the increase in ethnic diversity, the improvement in health technology, the increased pressures on the health care system, the Baby Bust and other social trends, there is a need to study the determinants of health on an on-going basis. Taken into account the expertise gained with NPHS, the current context and thinking about forthcoming issues, we are commenting on what NPHS II could look like.

5.4.3. Adjusting the Content: a Focus on Specific Issues and Collecting Physical Measures

The NPHS has mainly been an omnibus health survey covering all age groups. In addition to the core content, modules are used throughout the survey's life to collect in-depth data on specific topics. This allows supporting a mixture of an omnibus health survey with results available every 2 years with a more in-depth survey content but on a least frequent basis. This is particularly relevant for health-related phenomena that vary slowly at the population level, such as health behaviours. The approach for NPHS-II should be similar, that is, the survey should have a basic core content that will be measured at each wave of the panel and a focus content that may vary from one cycle to the next through the use of modules.

Since NPHS was initially designed as both a cross-sectional and longitudinal survey, it covers all age groups and does not over-sample specific population subgroups. Even though the sample size can be seen as relatively large for a general longitudinal health survey, it becomes rapidly very limited if we want to study a specific age group or a specific subpopulation group. One of the first steps in a discussion regarding a NPHS-II relates to the basic objective of the survey. Should it be an omnibus health survey or should it be focusing on one or few emerging issues? We are suggesting that it would be more appropriate to focus on emerging health and social issues. Otherwise, the sample size for NPHS-II would have to be increased quite significantly. Focusing on specific issues will have a significant impact on the age groups covered by the survey and its sampling design. The survey content should be planned from the beginning and linked to specific research questions. The challenge here is of course that research and policy issues change; how does one anticipate future "hot topics"?

The following are examples of emerging health and social issues that could be considered:

The aging of the baby boom
The baby boom is aging and will be living longer than any previous generation. How is the baby boom generation aging? What are the factors that are related to healthy aging of the baby boom? Are they modifiable? There have been significant changes in living arrangements of seniors over the last few decades. How will this impact on the need for long-term health care services of the baby boom? In order to answer these questions, over-sampling of the baby boom generation and of seniors would be required. A content focusing on aging would also have to be developed.

The reliance on immigrants for population growth and the increase in ethnic diversity
Canada is relying more and more on immigration for population growth and a significant change in the source country of immigrants has happened over the last 20 years. NPHS results indicate that the health of the recent generation of immigrants may be deteriorating faster than the health of the Canadian born. A better understanding of the adaptation of immigrants in terms of health is needed. In this case, NPHS II would have to have specific content relating to the health of immigrants, the adjustment process and an over-sampling of this population group.

The health gap between Aboriginals and other Canadians
Aboriginal Peoples of Canada significantly lag behind other Canadians in terms of health and social well-being. There is a need to better understand the path to good health for Aboriginal Peoples, while being respectful of their identity and culture. NPHS II could include a significant aboriginal component, including

First Nations on reserve. Another option could be to develop a longitudinal Aboriginal Peoples health survey in close collaboration with Aboriginal Leaders.

These are only few examples of emerging health and social issues that could be used to start the discussion on NPHS II substantive content.

The need for physical measures

One of the weaknesses of NPHS is the lack of physical measures. For example, if we want to understand properly the path to obesity, its link with diabetes, the role of physical activity in preventing disability, we need to have direct measures. Even though physical measures are expensive to collect, they are critical in the context of a longitudinal health survey. The objective should not be to replicate the Canadian Health Measures Survey, but to collect only basic and easy to gather physical measures. The exact measures to be included in NPHS II should be discussed with researchers and should take into account the logistical, privacy and ethical issues associated with direct measures.

5.4.4. Special attention should be given to the survey design of NPHS II.

The survey design, including stratification, population covered and sample size should be based on a set of research questions and topics to be covered. NPHS was designed to cover all age groups due to the need to support the production of cross-sectional estimates for population health monitoring at the provincial level. We think that this would not be an appropriate approach for NPHS II, since the Canadian Community Health Survey provides cross-sectional data for the Canadian population health monitoring program. Therefore, NPHS II should solely be a longitudinal panel survey.

5.4.5. The same data collection method should be used throughout the duration of the panel

There was a change in the collection of NPHS from personal interview in the first cycle to telephone interview in subsequent cycles. For some variables, such as smoking and obesity, there is a significant mode effect. Therefore, to maximize the capacity of the survey to study transitions, the same mode of collection should be used throughout the life of NPHS II.

5.4.6. Linking survey and administrative data

As was done in NPHS, and is being proposed for the SLID, NPHS II should be requesting, from respondents, permission to link their survey data to the relevant administrative records, in this case health records. Statistics Canada currently has access to cancer registry, vital statistics and hospital morbidity data at the national level. Current discussions are taking place between Statistics Canada and provincial ministries of health to gain access to prescription drug and physician billing data. Linking NPHS II data -for respondents who provided consent- to these administrative databases would greatly improve NPHS II analytical potential. It would allow to study the impact of socioeconomic status and risk factors on the use of health services. Does

the rate of hip and knee replacement vary by socioeconomic status? Is there a socioeconomic gradient in the use of drug for diabetes? Is the rate of disability reduced for individuals who properly control their diabetes? To mention a few.

5.4.7. NPHS II should be a 20-to 35-year panel survey

There are long latent effects in health. For example, the lag between smoking initiation and the development of smoking related diseases is quite long, ranging from 10 to 35 years, depending on the condition. For obesity, little is known about the lag between becoming obese and the development of diabetes or circulatory diseases. There are also lag effects in the association between some social phenomena and health outcomes. Does the health status of immigrants converge with the health status of Canadian born? If, so, after how many years after immigration? A 20-to 35-year panel length would allow for the evaluation of these effects.

5.4.8. NPHS II should interview respondents every 2nd year, with physical measures every 4th year

A 2-year cycle has been adequate for NPHS. However, if physical measures are included within NPHS II, it may be that a biennial survey with physical measures is simply too expensive. Physical measures can be collected every second cycle. Further discussions are required before formally making a proposal on wave length. These discussions should take into account the proposed research program, survey content and budget.

As noted in the "generic issues" section, to make NPHS II an even greater success, we need to involve the scientific community in the identification of emerging health and social issues and in establishing the survey content.

Appendix I

An overview of Statistics Canada's Longitudinal Data Sources

The following list is not exhaustive, but it provides a thumbnail sketch of some of the most important longitudinal surveys in progress.

National Population Health Survey
NPHS started in 1994-1995, with funding from the first Data Gaps initiative. NPHS is conducted every two years and has a longitudinal sample of 17,000 persons of all ages. These same persons will be interviewed every two years over a period of 18 years (10 cycles).

The objectives of the NPHS are to examine:

- the level, trend and distribution of the health status of the population;
- the determinants of health;
- the economic, social, demographic, occupational and environmental correlates of health;
- the relationship between health status and health care utilization;
- the dynamic process of health and illness.

NPHS was also designed to serve as a platform for supplementary content or sample, and to be linked to routinely collected administrative data such as vital statistics, environmental measures, community variables, and health services utilization.

Survey of Labour and Income Dynamics
Also funded from the first Data Gaps initiative, SLID examines changes experienced by individuals over time in terms of their labour market activities and income. At the heart of the survey's objectives is the understanding of the economic well-being of Canadians: what economic shifts do individuals and families live through, and how does it vary with changes in their paid work, family make-up, receipt of government transfers or other factors?

SLID is the first Canadian household survey to provide national data on the fluctuations in income that a typical family or individual experiences over time which gives greater insight on the nature and extent of poverty in Canada. Added to the longitudinal aspect are the "traditional" cross-sectional data: the primary Canadian source for income data and providing additional content to data collected by the Labour Force Survey.

The SLID sample is composed of two panels. Each panel includes roughly 15,000 households. A panel is surveyed for a period of six years. A new panel is introduced every three years. Thus two panels are always overlapping. Annual interviews are

conducted for all household members aged 15 and over; respondents have the option of authorizing access to tax data instead of completing income questions

National Longitudinal Survey of Children and Youth
NLSCY is a study of Canadian children that follows their development and well-being from birth to early adulthood. The NLSCY began in 1994 and is jointly conducted by Statistics Canada and Social Development Canada (SDC).

The study collects information about factors influencing a child's social, emotional and behavioural development and monitors the impact of these factors on the child's development over time. The survey covers a comprehensive range of topics including the health of children, information on their physical development, learning and behaviour as well as data on their social environment (family, friends, schools and communities).

The NLSCY surveys the non-institutionalized population (aged 0 to 11 at the time of their selection) in Canada's 10 provinces. Interviews are conducted every two years, so five cycles of data have now been collected.

The longitudinal sample at Cycle 5 consists of three cohorts. The first cohort consists of children aged 0 to 11 at the time of their selection at Cycle 1 in 1994, who are 8-19 at Cycle 5. They will remain in the survey until they reach the age of 25. The second cohort is made up of children aged 0 to 1 at the time of their selection at Cycle 3 in 1998, who are 4-5 at Cycle 5. It is their final cycle in NLSCY. The third cohort consists of children aged 0 to 1 at the time of their selection at Cycle 4 in 2000, who are 2-3 at Cycle 5. These children will be interviewed one more time in Cycle 6.

Workplace and Employee Survey
WES is a GAPS-funded survey designed to explore a broad range of issues relating to employers and their employees. The survey aims to shed light on the relationships among competitiveness, innovation, technology use and human resource management on the employer side and technology use, training, job stability and earnings on the employee side.

The survey is unique in that employers and employees are linked at the micro data level; employees are selected from within sampled workplaces. Thus, information from both the supply and demand sides of the labour market is available to enrich studies on either side of the market.

WES uses two reference periods. Questions concerning employment breakdown use the last pay period of March for the reference year while other questions refer to the last 12-month period ending in March of the reference year.

Some 6,000 business locations are surveyed. The initial sample selected in 1999 is followed over time and is supplemented at two-year intervals with a sample of births selected from units added to the Business Register since the last survey occasion. Business locations are in the WES sample for eight years. A sample of about 20,000 employees in these firms is followed for two years.

Youth in Transition Survey
YITS is designed to examine major transitions in young people's lives. HRSDC-funded, YITS includes measurement of major transitions in young people's lives including virtually all formal educational experiences and most labour-market experiences, achievement, aspirations and expectations, and employment experiences. The survey covers two cohorts: youth aged 15 and 18-20 in 2000. Interviews are conducted every two years.

The 15 year old cohort was selected from schools. The sample of 30,000 young people also completed the Programme for International Student Assessment (PISA), which offers direct measures of skill in reading, mathematics and science. PISA was conducted in over 30 countries.

National Graduate Survey & Survey of Earned Doctorates
NGS examines the labour market outcomes of postsecondary graduates two and five years after graduation. The sample is drawn from postsecondary institutions and includes an oversample of Masters' graduates and a census of PhDs. NGS is a long-standing survey, originally funded by HRDC. It is currently funded under GAPS II.

The survey content covers graduates' job and career satisfaction; the rates of under-employment and unemployment; the type of employment obtained related to career expectations and qualification requirements; and the influence of postsecondary education on occupational achievement.

The survey is conducted about every 5 years, the last cohort being the Class of 2000.

Recently, a Survey of Earned Doctorates has been added to the program. SED collects information on the plans of PhD at the point of graduation, including plans for further study, migration and work.

Longitudinal Administrative Dataset
LAD is a longitudinal file designed as a research tool on income and demographics. It comprises a 20% sample of the annual T1 Family File and the Longitudinal Immigration Data Base. Variables have been harmonized where possible and individuals can be linked year to year starting with 1982 data. The file is augmented annually with new data.

The longitudinal file contains many annual demographic variables about the individuals represented and annual income information for both the individual and their census family in that year. For immigrants landed since 1980, the file also contains certain key characteristics observed at landing.

The longitudinal nature of the LAD permits custom-tailored research into dynamic phenomena, as well as representative cross-sectional patterns. Data are used to evaluate government programs and support policy recommendations, and for analyses of socio-economic conditions.

Longitudinal Immigration Database
IMDB is a database combining linked immigration and taxation records. It covers the immigration landing years since 1980 and is updated with tax information annually for 16 years. The IMDB offers data on the economic behaviour of immigrant tax filers and is the only source that provides a direct link between immigration policy levers and the economic performance of immigrants. The database is managed by Statistics Canada on behalf of a federal-provincial consortium led by Citizenship & Immigration Canada. The database covers persons who obtained their landed immigrant status since 1980 and filed at least one tax return after becoming a landed immigrant.

The IMDB supports analysis of labour market outcomes of different categories of immigrants, along with immigrant characteristics, such as education and knowledge of French or English. It also supports research on the role of social assistance as well as secondary inter-provincial and inter-urban migration.

Longitudinal Survey of Immigrants to Canada
LSIC was launched in 2001 to meet a growing need for information on recent immigrants. While integration may take many years, LSIC is designed to examine the first four years of settlement, a time when newcomers establish economic, social and cultural ties.

The survey objectives are two-fold: to study how new immigrants adjust to life in Canada over time; and, to provide information on the factors that can facilitate or hinder this adjustment.

Topics covered in the survey include language proficiency, housing, education, foreign credential recognition, employment, health, values and attitudes, the development and use of social networks, income, and perceptions of settlement in Canada.

The target population for the survey consists of immigrants who meet all of the following criteria: arrived in Canada between October 2000 and September 2001; aged 15 years or older at the time of arrival; landed from abroad as permanent residents,

therefore, must have applied for admission to Canada through a Canadian Mission Abroad. All individuals who applied within Canada have been excluded from the survey as these people may have been in Canada for a considerable length of time before being granted permanent resident status and would likely demonstrate different adaptation characteristics from those recently arrived in Canada.

Hospital Person-Oriented Information Database

The HPOI is a file of all hospital separations in Canada longitudinally linked at the individual level used to conduct health services research.