The LifePaths Microsimulation Model: An Overview

lifepaths@statcan.gc.ca

Compiled by Martin Spielauer

with contributions from

Chantal Hicks, Steve Gribble, Geoff Rowe, Xiaofen Lin, Kevin Moore, Laurie Plager, and Huan Nguyen

Statistics Canada – Modelling Division
R.H. Coats Building, 24-O
Ottawa, K1A 0T6

September 2013

Abstract

This document gives an overview of LifePaths, a dynamic microsimulation model of Canadian society developed and maintained in Statistics Canada's Modeling Division. By simulating detailed individual life histories for millions of virtual Canadian residents, LifePaths reproduces the socio-economic and demographic features of the past half-century and enables detailed projections into the future under alternative assumptions and scenarios. LifePaths is used to analyse government policies having an essentially longitudinal component, like the sustainability of social insurance systems or long-term effects of investments in education. Based on individuals in their family context, it supports distributional analysis in all its dimensions: cross-sectional, over individual lives, between cohorts and over generations.

1 Introduction

LifePaths is a dynamic microsimulation model of Canadian society. It is designed to simulate a large sample of detailed individual life courses that together represent the Canadian population in its diversity, both on the national and the provincial level. Building on the individual life histories of millions of virtual Canadian residents, it reproduces the socio-economic and demographic features of the past half-century and allows for detailed projections into the future under alternative assumptions and scenarios.

LifePaths is mostly based on statistical event-history models estimated from a wide variety of Canadian micro-data combined with detailed accounting models mirroring the Canadian tax-benefit and social insurance system. Individual actors attend school and make educational choices, form families, migrate, become parents, work, have earnings, take parental leave, lose and find jobs, acquire homes, pay taxes, contribute to mandatory and voluntary pension schemes, receive benefits and pensions, and eventually die.

LifePaths is used to analyse government policies having an essentially longitudinal component and whose nature requires evaluation at the individual or family level, such as postsecondary education costs and benefits or public pension sustainability. It can be used to explore a variety of societal issues of a longitudinal nature, such as intergenerational equity or time allocation over entire lifetimes. Based on individual histories, it supports distributional analysis in all its dimensions: cross-sectional, over individual lives, between cohorts and over generations.

LifePaths is implemented in the Modgen microsimulation language developed and maintained at Statistics Canada. Like all Modgen applications, it provides a graphical user interface and runs very efficiently on standard PCs or distributed on computer networks. LifePaths is highly modular; it is an evolving model with new components added on demand.

This document is organized in two main sections. The first provides an overall picture on model design, its technical implementation, data sources, the target audience and ways of model use. The second part—LifePaths components—sketches the content of LifePaths by describing each of the social processes and policies that are simulated and the underlying data.

2 LifePaths structure

2.1 Overview

At the very core of LifePaths lies the detailed simulation of individual life histories. LifePaths creates synthetic life histories from birth to death that are representative of the history of Canada's population. Individual actors are simulated in their family context, with spouses and children explicitly simulated and linked together to families. Individual life courses comprise a multitude of parallel aspects of life, which can be structured as follows:

- Basic demographics: Birth, mortality, internal and international migration.
- Family demographics: Partnership formation, generation of spouses, pregnancy and childbirth, children leaving home.
- Education: Educational transitions, levels and fields of study.
- Health: Health Utilities Index, disability and institutionalization.
- Employment: Type of employment, unemployment, hours worked, maternity leave and retirement.
- Income and earning histories: Students' and career earnings, retirement income.
- Owner-occupied housing: Buying, selling and upgrading and downgrading decisions; house prices.

A second component of LifePaths is a detailed representation of Canada's tax benefit and social insurance system as it evolves and changes over time. Simulated individuals thus live in a concrete and dynamic policy with context—they pay taxes, receive transfers, contribute to pension plans and receive benefits. These features, listed below, make LifePaths suitable for the detailed analysis of distributional aspects and long-term sustainability issues of government policies in the context of fast socio-demographic change:

- Employment/unemployment insurance contributions and benefits.
- Family, maternity and parental benefits.
- Federal taxes and transfers.
- Payroll taxes.
- Public pension plans, pension contributions and benefits.
- Senior benefits (Old Age Security, Allowance for the Survivor, Guaranteed Income Supplement).
- Registered Pension Plan, Registered Retirement Savings Plan.

Note: Implementing most economic components involves individual behavioural choices (e.g., saving decisions), incorporating past and a projection of future indices (e.g., interest rates), economic growth, and the 'mechanical' coding of the Canadian tax-benefit system.

2.2 Model design

LifePaths can be classified as a continuous time competing risk model following an open population approach. It is a case-based model with interactions between individuals limited to such between family members. All lives are simulated from birth, thus all lives are fully synthetic and do not correspond to real individuals from a single micro-data set.

Simulated life courses consist of various parallel and interacting careers, each defined by a set of possible states and events that lead to changes in those states. The life course of an individual is thus simulated as a series of events. These events are contingent on the previous history of the individual and usually have a stochastic component as well. The state of an individual changes when an event occurs. This in turn changes the likelihood of subsequent events.

By adding a stochastic component to the decision, the translation of a likelihood into a simulated event becomes a random process. For example, an individual who is currently employed is at some risk of a job loss event. The risk of job loss depends on factors such as the length of time that the individual has held their current job. In the event of job loss, the individual's state changes—they enter the state 'not employed.' This new state may influence other events; for example, it may diminish the likelihood that this individual will subsequently become married or increase the likelihood that an existing marriage will break-up.

LifePaths simulations take place in continuous time. Events can occur at any arbitrary moment and are not artificially restricted to annual intervals. Events that occur in continuous time are unlikely to occur simultaneously unless specifically defined as a joint event. Consequently, there is no need for the tie-breaking rules that are often necessary in discrete time simulations. For example, there is never ambiguity about whether a birth that is simulated as occurring in the same year as a marriage occurred before or after the marriage.

Over the course of a simulation, LifePaths keeps updating its list of pending events to ensure that the next scheduled event is the one that currently has the shortest waiting time. This provides a straightforward way of dealing with competing events. Waiting times provide a unifying framework for representing decision-making. Probabilistic decisions can be implemented so that the choice among alternatives is determined by comparing two or more waiting times. For example, a never-married person's decision to either marry or enter a common-law union can be implemented by generating waiting times for both events. Such decisions are generally influenced by other variables—for instance, current educational status and employment history—and the decision is always re-assessed each time the influential variables change. If the marital status waiting times are both long relative to the waiting times for changes in employment or education status, then neither marriage nor a common-law union will become the next event to occur. Otherwise, if the marriage waiting time is shorter than the common-law union waiting time, then the choice is to marry.

A LifePaths simulation consists of a set of mutually independent cases. Each case contains exactly one dominant individual in the first generation. The spouse and children of the dominant individual are simulated as part of the case. They are created to satisfy the union formation and fertility equations. Furthermore, their behaviour is determined by the same equations that govern the dominant individual's behaviour. Non-dominant individuals are excluded from most tables to avoid double counting. Nevertheless, non-dominant individuals contribute to tables by defining the family context in which the dominant individual lives. The family context may be as simple as children being present in the home or a complex function of the spouse's employment history.

LifePaths is an open model in that a non-dominant spouse will be created whenever it is decided that a union formation will occur. Moreover, LifePaths simulates the completion of one case before going on to another. These properties contrast with closed, discrete time models, which simulate all the members of a population in one time period before proceeding to the next time period. Those models maintain a list of currently single individuals and permit union formation only between individuals on the list. Closed models typically start from an observed cross-section—perhaps a demographic survey. Thus, closed model simulations can never be more accurate than that allowed by the sampling variability of the initial cross-section. In addition, they often begin with a limited history with which to inform the behavioural equations. In contrast, LifePaths generates a complete history for each individual starting from birth, which allows greater flexibility in specifying behavioural equations.

The oldest birth cohort that is represented in LifePaths was born in 1872. This year was chosen so that, in the year 1971, LifePaths could simulate a complete range of ages from newborn to elderly. This means that starting in 1971, LifePaths can produce cross-sectional annual tabulations that can be directly compared with historical cross-sectional data. The year 1971 is the first year for which high quality socio-demographic data was available from a census with a contemporary, self-enumeration design.

2.3 Technical implementation and model output

LifePaths is implemented in Modgen, a generic microsimulation programming language that supports creating, maintaining and documenting dynamic models. Modgen was created with the goal of automating as many aspects of the microsimulation model implementation as possible. Common to all Modgen applications is the graphical, bilingual user-interface with well-organized and fully labelled multi-dimensional input and output tables.

One of the most important concepts in the class of simulation models to which LifePaths belongs is Monte Carlo variation. A LifePaths run creates a sample of individuals and families using waiting times that incorporate a stochastic component. This stochastic component is a key element in the ability of LifePaths to generate life histories that match the diversity observed in actual populations. On the other hand, this variability affects the reliability of aggregate tabular results in the same way that a small sample drawn from a larger population has limited accuracy. In the context of simulation modeling, this statistical variability is called Monte Carlo variation. It can vary from cell to cell within the same table and depends on the number of observations in the cell, the model's equations and the nature of the aggregate quantity being tabulated.

Modgen automatically estimates the Monte Carlo variability of each cell of every table and reports it in the form of a standard error. This feature substantially increases the information content of all model output, as, before drawing conclusions based on tabular results, it enables examination of these measures of variation. If the standard error is too large, the number of cases in the simulation may need to be increased in order to achieve the desired level of accuracy. Alternatively, an increase in number of cases that contribute to a given table might also be accomplished by more carefully targeting the simulation. For instance, appropriate targeting might be achieved by simulating only people born in a certain narrow range of years. LifePaths allows the sample size to increase virtually without limit. (A typical LifePaths simulation run creates from 4 million to 30 million cases, which themselves consist of the dominant actor and a multitude of related family actors: spouses, children and grandchildren.)

Because it is based in individual life courses, LifePaths produces very rich output. On the aggregate level, model output is organized in a hierarchical list of tables. Because tables are assembled during a model run, LifePaths does not require massive output files—thus, that there is virtually no limit to the size of a simulated population. LifePaths makes use of the powerful table generating language provided by the Modgen programming technology. The Modgen table generating language is especially powerful for output of durations, in addition to cross-sectional ‘point in time’ output. Modgen also allows table outputs to be easily added, organized and customized.

The second form of output is a database that tracks the life histories of simulated individuals as represented through the values and timing of changes to user-selected state variables. A tracking database will usually contain a small number of carefully selected simulated individuals (hundreds rather than thousands) that can be inspected one-by-one. The tracking database serves as input to a secondary program named BioBrowser, which graphically displays values of specified variables for simulated individuals over their lifetimes—a BioBrowser biography.

2.4 Data

The content of simulation models is largely embodied in their behavioural equations. In LifePaths, these equations—together with their stochastic components—determine the distributions of waiting times to events. These equations represent behaviour by depending on a simulated individual’s past. For example, the timing of births is determined in relation to the timing of marriage and of previous births, rather than solely on a mother’s current age. (Age-specific fertility rates are common elements of demographic projection models that are not designed to generate individual life histories that seem realistic.)

The behavioural equations represent choices as well as contingent events such as death. Estimating these equations places particularly heavy demands on data because the required data must directly reflect individual events and the specific history that led to them. LifePaths makes use of a multitude of data sources with longitudinal components—for example, by constructing a database of Labour Force Survey data linked together over time and by using retrospective event history data collected in the General Social Survey, and in panel surveys, like the Youth in Transition Survey.

One of the biggest challenges of LifePaths is to be simultaneously consistent cross-sectionally and longitudinally—individual careers have to make sense and be representative for the real life-course diversity observed in Canadian society over time. It is one of the design goals of LifePaths that the distribution of characteristics in the simulated population closely reproduces data from censuses over all its waves. In this respect, the construction of LifePaths can be viewed as a data integration exercise, which, by detecting inconsistencies and data gaps, can improve data quality and collection in the long term.

2.5 Use and users

The largest user group consists of Canadian federal policy departments and provincial ministries involved in policy development. Other target audiences are non-governmental organizations that have a stake in long-term government policy, banks, demographers, economists and other researchers. Those to whom model results or analyses based on model results are presented comprise the second audience. This can include the press, political parties and, of course, decision makers in government.

Early versions of the model focused on the young population and were used by federal and provincial departments to assess the viability of major reforms in postsecondary education funding. More recently, the LifePaths team has successfully developed longitudinal behavioural equations that accurately track major features in the evolution of Canadian society (demography, education, migration, labour market and earnings) over the 20th century. This establishes a high level of credibility for the model and provides a firm foundation for those in the policy community to perform projections with respect to current and proposed government policies. Most recently, it has been used by federal departments in a project to analyse various aspects of retirement.

LifePaths has also been used in a variety of research papers. Topics have included time use, generational accounting, active life expectancy, maternity leave, the returns to education, and changes in pension policy. LifePaths also provided the demographic foundation for the Population Health Model (POHEM) developed at Statistics Canada, which enables competing health intervention alternatives to be rationally compared in a framework that captures the effects of disease interactions.

LifePaths is a very complex model with a development history going back more than a decade. Getting familiar with LifePaths for any non-trivial analysis requires a considerable investment of time and hardware resources, and most users seek some kind of support from the LifePaths team. Statistics Canada is flexible in providing various forms of support ranging from carrying out the projects themselves as defined by clients to providing training to clients so that they can use the existing model, develop their own modules and perform their analysis independently outside Statistics Canada. The first option might be preferable for a single analysis of an issue using LifePaths, as it may be more cost-effective to pay to have Statistics Canada staff do the work than face a large upfront investment in learning and computer resources. The latter option, which centres on training, has proven successful, especially in cases of university-based research that uses the existing LifePaths model as a base and adds new modules without modifying the existing LifePaths base.

For large-scale projects that go beyond the use of LifePaths as a black-box model, the greatest difficulty in mastering LifePaths is understanding its structure and complexity. LifePaths has hundreds of variables that interact with each other in complex ways. Users as well as new members of the LifePaths team have found that it takes months to begin to understand the relationships, and can take years to become an expert in being able to use LifePaths productively in complex policy analysis. Thus, not surprisingly, most complex policy analysis using LifePaths is performed by or in close relationship with the LifePaths team. In addition, the LifePaths project has the unique advantage of being located in Statistics Canada. This allows the development team ready access to the many data sources required to estimate the distributions of characteristics and equations used to model the behaviour of individuals and families.

From the technical side, learning to use LifePaths involves learning the Modgen programming language in which LifePaths is implemented. When not intending to modify LifePaths itself, learning how to create output tables is typically the first step involving programming. While a new LifePaths user does not have to be a programmer, they need to understand the code in Modgen to make adjustments to the model.

Hardware resources for a large number of cases—typically from 4 million to 32 million—need to be run to get good results on the distribution of variables being studied in LifePaths. Simulating such large samples typically requires hardware investments. LifePaths makes use of multiple-cores in single and multiple networked computers; a computer or network with eight cores is recommended and sufficient for most purposes.

3 LifePaths components

3.1 Basic demographics: Birth, migration and death

The population in each province at any moment in time is determined by three factors: births, deaths and migration flows either within Canada or internationally. An adequate representation of Canadian population dynamics by its key demographic characteristics (age, sex and province of residence) is a core requirement for all simulation output.

LifePaths aims at exactly replicating historic and administrative population data, official population estimates and the three (low, medium and high) projection scenarios periodically published by Statistics Canada. In this sense, LifePaths does not produce its own demographic projections, but reproduces given macro projections. This makes the simulation results comparable with those obtained from other models using similar population assumptions.

The modeled demographic behaviours interact with other life course characteristics. For example, migration depends on the family type, while mortality depends on health (in addition to age, time period and sex). The calibration of the micro models in order to reproduce given macro data is one of the challenges to be faced in further developing LifePaths. An additional challenge consists of data integration issues arising from the enormous historical depth of the simulations that go back to 1872.

LifePaths has a synthetic starting population, meaning that all people are simulated from birth. Like in real life, LifePaths actors become residents either by birth or immigration and can leave Canada temporarily or permanently through emigration. Technically, we can distinguish between characteristics assigned at birth and such modeled over the life course. Those assigned at birth (the creation of an actor) are sex, year of birth and province of birth as well as the year and province of immigration for people born outside of Canada. All other individual characteristics and changes over the lives of the people are modeled; no simulated life has a direct real-life correspondent and thus avoids all confidentiality issues.

3.2 Becoming a Canadian resident

Compared with 1872, the population of today's Canada has increased more than seven-fold, with immigration being a key driver of this increase. Canada is a nation of immigrants. At the peak of immigration, in the decade from 1901 to 1911, the total number of immigrant arrivals amounted to about 25% of the average population. Immigration continues to make a crucial contribution to Canada's continued population growth.

The characteristics of immigrants differ from those of native-born Canadians. For example, recent immigrants are likely to be more highly educated than the Canadian average. For those reasons, special attention must be given to the immigrant population, beginning with a determination of the age and time at which they arrive.

Concerning births and immigration, LifePaths is based on population censuses going back to 1911, birth and immigration records, population estimates and population projections. Most administrative data are available since 1921. For the period before 1921, the 1921 Census of Population was 'reverse survived': the number of births that would have had to occur to produce the observed number of survivors in 1921 was determined using mortality probability estimates. By integrating these data sources, distributional tables were constructed that allow a set of key characteristics to be randomly assigned to each of the simulated actors: sex, year of birth, province of birth, and—for those foreign born—the year and province of immigration.

Immigrants become Canadian citizens after three years in LifePaths. The population represented does not include non-permanent residents, a large number of whom are foreign students.

Note: The way LifePaths creates a population as described above only refers to the so-called dominant actors of generation zero. These are the actors from which LifePaths constructs most table outputs and which together are representative of the Canadian population. Each of these actors can have a series of spouses and various generations of children and grandchildren—the generational depth is set by the user. The generational depth of LifePaths—besides creating context for the dominant actor—can be exploited for the study of intergenerational mechanisms and effects.

3.3 Emigration and back migration

All individuals in the model who are currently living in Canada have a probability that they might emigrate. Factors that influence emigration include age, sex, province of residence, calendar year, immigrant status and the year of immigration (for immigrants).

Canadian citizens who emigrate can also return to Canada. This means that immigrants who leave the country after having lived in Canada for three years or more as well as individuals born in Canada can emigrate and then return to Canada. However, immigrants who leave Canada before becoming Canadian citizens are not eligible to return to Canada. Factors that influence a return to Canada include age, sex, year, last province of residence, immigration status and year of immigration (for immigrants).

All migration is modeled at the family level: if one family member migrates, their spouse and dependent children also migrate.

3.4 Internal migration

As well as moving in to and out of Canada, people in the model and their families can move from one province to another. A base probability of moving between each pair of provinces serves as the benchmark for estimating family migration probabilities for each of five family types: husband and wife couples, common-law couples, lone parents, other single males and other single females. These probabilities incorporate the influence of factors such as age, education and whether or not a person would be moving from or back to their province of birth. Two additional factors are the current province of residence and the province of birth. Return migration to the province of birth proved to be among the strongest of the migration patterns uncovered in the analysis of migration data. Within a simulation, the base probability of moving between each pair of provinces varies from year to year.

Data for the earliest years are limited; in some cases, all that are available are census data comparing province of birth with province of residence, which is indicative of lifetime migration. Currently, more direct and timelier data are available from administrative files. Historical lifetime internal migration data was based on decennial census data from 1911 to 1971. Base probabilities of individuals moving between each pair of provinces were derived using the family allowance and child tax credit data (1972 to 1996). Estimates of family-level migration probabilities were obtained (using the base probabilities as benchmarks) from census data on place of residence one year prior to the census.

3.5 Mortality

Mortality in the LifePaths model is determined by re-assessing the chance and timing of death at each birthday. The process of re-assessing mortality continues until death or until the individual reaches the maximum allowed age of 119 years. The chance of dying is based on the age-specific mortality rate of Canadian residents sharing the sex and year of birth of the simulated individual as well as (user-selectable) sets of relative mortality risks. The cohort mortality data were derived from death registration statistics. Concerning future mortality, model users can parameterize—and choose between—a low-, medium- and high-mortality scenario.

While the model reproduces the life-tables based on administrative data and official population projections on the aggregated level, it also implements various types of relative risks, of which different sets can be selected by the user:

- institutionalization only or health status (including institutionalization)
- Health Utilities Index (on/off)
- education and marital status (on/off).

Relative risks by marital status and education level were estimated from the census mortality follow-up study. Accounting for the difference of life expectancy by education and marital status can be of importance—e.g., in pension modeling, the remaining life expectancy of a 60 year old varies by two years between high school non-completers and university graduates and the combined effect with marital status can lead to a difference in life expectancy of up to three years for males 60 years of age. Similarly, accounting for mortality differentials by health status and institutionalization can be of importance when modeling—e.g., formal and informal care needs. Health status and the Health Utilities Index are explained in separate chapters on the according behaviours.

Individuals who are destined to immigrate to Canada are not exposed to a risk of death until they arrive. This avoids simulating individuals who then die before reaching Canada and who would, therefore, make no contribution to the simulation reports.

3.6 Family demographics: Fertility, partnership formation and dissolution, leaving home

3.6.1 Fertility

Women born in the 1880s had an average of nearly four children in their lifetimes—a level never reached again even in the baby boom of the 1960s. Currently, like in most developed countries, the fertility level is well below the replacement rate, thus population growth is driven by immigration and decreasing mortality only.

LifePaths assigns women a fertility status that depends on age and whether any fertility events have occurred in their immediate past. At 15 years of age, women are at risk of pregnancy for the first time. At age 50, women have no further risk of pregnancy. Between those two ages, births are simulated as a sequence of fertility decisions. Each fertility decision has two parts: first, the decision whether to have a child is made and only then is a waiting time to conception generated. All pregnancies result in a live birth unless the mother dies first. Twin births are simulated using an appropriate incidence rate.

The main factor influencing the decision to have an additional child is the number of children already born in the family. Secondary factors include the mother's year of birth, province of birth, marital status, age at marriage and educational attainment. For most women, fertility decisions are delayed until schooling is completed. Women in common-law unions sometimes have reduced fertility compared with their married counterparts. The latter is important because marriage and fertility are linked to each other in LifePaths. If a child is conceived, the waiting times to marriage or common-law union are re-assessed and if a marriage or common-law union occurs, the chance of an additional birth increases.

The normal sequence and timing of fertility related events are: (1) each birth occurs nine months after conception, (2) a spell of infertility lasting three months follows each birth and (3) a new fertile spell begins immediately after that. The careful attention given to the timing of these events makes possible straightforward models of maternity leave from employment and marriage following pregnancy.

The decision probabilities were derived from pooled census data from 1971 onward using the distribution of children ever born from women with completed or nearly completed fertility. Intervals between births are based on census responses of women in their thirties for whom all children were still at home. Birth intervals were then calculated using the age differences among the children.

3.6.2 Marriage, common-law unions and divorce

In the past, only a small proportion of people never married in their lifetimes and people rarely divorced or entered a common-law union. Currently, one-third of marriages is expected to end in divorce and non-marital cohabitation became a wide-spread phenomenon both as alternative to marriage and as an option preceding marriage. Institutional change reflects societal changes—e.g., the first changes to the *Divorce Acts* of 1968 and 1985 to simplify divorce procedures led to peaks in observed divorce rates.

The growing acceptability of common-law unions is a more recent phenomenon than increased divorce incidence. In the 1981 Census, just over 6% of couples indicated that they were in a common-law relationship; meanwhile, marriage is no longer the universal norm, with Quebec leading this de-institutionalization trend.

Societal change is also reflected in the changing age of marriage and childbirth. Historically, the trends of steadily declining marriage and childbirth ages were reversed following the baby boom. Since then, life course patterns have started to diversify again, with higher degrees of childlessness and a steady increase of age at marriage and childbirth. Careful modeling of these trends is important, as partnership status and parenthood have a considerable influence on other events—e.g., they play an important role in migration and labour market events. Increasing inequality concerning family size is also a factor in the general increase of economic inequality.

Marital status transitions can be very complex. One example could involve the events leading to an individual who is separated, but still legally married, being simultaneously in a common-law union. LifePaths deliberately limits the complexity of marital status transitions. The earliest age at which an individual can enter a common-law union or marriage is 15. For the never married, common-law unions and marriage are competing events—either one may occur. Factors that influence the timing of common-law unions and marriage include sex, age, pregnancy, employment status, education, year of birth, place of birth (born in Quebec, born elsewhere in Canada or born outside of Canada), age at first job and time employed since first job.

When a marriage or common-law union occurs, LifePaths creates a suitable spouse. Such spouses have an age and education following observed statistical patterns between spouses' characteristics. There are two methods in LifePaths for creating a spouse. For large runs, a spouse market is created before the simulation of the cases. When a spouse is needed, he or she is selected from this spouse market with an age and education close to target levels. The second method involves creating a series of individuals with appropriate age and sex until one with the appropriate level of education is found. This individual then becomes the spouse or common-law partner.

Once an individual enters a common-law union, there are three possible outcomes: death of one of the common-law partners, dissolution of the common-law union and marriage of the common-law partners. Factors influencing the outcomes include number of previous common-law unions, duration of the union, age, sex, pregnancy, year of birth, place of birth, and measures of employment experience.

Once an individual marries for the first time, there are two possible outcomes: death of one of the spouses or permanent separation. Factors influencing these outcomes include year of birth, age at marriage, duration of the common-law union that became the marriage (if applicable), duration of the marriage without children, duration of the marriage with children, age of the youngest child, education, measures of employment history and historical period. The historical periods represent, respectively, the period before the divorce legislation reform (before 1968), the period before the second change to divorce legislation (from 1968 to 1984) and the present (after 1984).

When a permanent separation occurs, a waiting time until a possible divorce is calculated. The factors that influence the waiting time include child status at separation, duration of the separation, year of birth, place of birth, having a common-law union prior to marriage, education and the historical periods related to divorce legislation reform.

Following the end of a first marriage by divorce or the death of a partner, a waiting time for a possible second marriage is calculated. Age, sex, province of residence and previous marital history are factors that influence this waiting time.

The primary sources of data used to calculate the probabilities of union formation and dissolution—prior to the first marriage—were the 1984 Family History Survey and the 1995 General Social Survey. Census data from 1981 onward were used to augment survey data. Probabilities of second marriages were based on marriage registration data. Data representing the distributions of age differences between spouses came from marriage registrations. The census provided data on the relationship between spouse education levels.

3.6.3 Leaving the parental home

There can be some ambiguity regarding what leaving home means. Postsecondary students, for example, may study in a different province for most of the year, yet maintain a legal residence in their parents' home. In fact, several movements in and out of the parental home might not be unusual over the course of an educational career. LifePaths focuses on reproducing distributions of non-students living with their parents by age, sex and year, in order to approximate the final move out of the parental home. Even in this non-student population, recent trends indicate that young people may be leaving home later in life.

In the LifePaths model, leaving home is determined by an annual re-assessment of the chance and timing of the event. This event happens at most once in a lifetime. An individual maintains their permanent residence with their parents while attending school. If an individual is still living at home at age 35, then they stay at home for the remainder of their life. However, if a marriage occurs, then all other considerations are over-ridden and the individual leaves their parental home immediately.

One outcome of leaving home is that the individual who left will no longer migrate with their parents.

Probabilities of leaving home were derived that would reproduce observed proportions of non-student children living with their parents, grandparents or foster-parents by child's age, sex and province of residence and by year. The annual average proportions of non-student children living at home were taken from waves of the Labour Force Survey from 1976 onward.

The mobility of postsecondary students who have not left home yet is treated separately by assigning a province of study, which might be different from the parent's province of residence.

3.7 Education: Primary, secondary and postsecondary education choices

Soon after Confederation, school attendance in Canada became compulsory and free schooling in tax-supported public schools, at least for young children, became widely available. Since then, Canada, like all developed societies, has experienced a tremendous educational expansion: while before the First World War half the population did not go beyond Grade 8, more than two-thirds of 30-year-olds today attained postsecondary education.

Educational attainment, while interesting by itself, also plays an important role in many processes modeled in LifePaths. It is a key variable for demographic events influencing fertility, family formation, migration decisions and mortality. Regarding the timing of life events, most other life careers (i.e., household formation, marriage and parenting) usually start after leaving school, while in economic modeling, education is a key determinant of human capital and therefore of income and job careers.

LifePaths models school attendance distinguishing elementary school, secondary school, community colleges, trade and vocational training, and the various levels of university education up to PhD studies. Study careers can consist of a series of full-time, part-time and out-of-school spells mirroring the variety of observed study patterns. LifePaths also distinguishes 100 fields of study and calculates tuition, thus supporting a wide range of policy analysis.

3.7.1 Primary and secondary education

The modeling of school entry and primary education is quite simple—individuals enter in September of the year of their 6th birthday. Students are assumed to stay in school full-time at least until June of the year of their 15th birthday. At this date, it is determined if Grade 8 was passed and at least some secondary education was started. If Grade 8 was not passed, students are assumed to drop out permanently; if it was, they move on to the secondary education module. This date also marks the earliest entry date into the formal labour market. For more recent cohorts, progression to secondary education has become virtually universal.

Starting from the year of the 15th birthday onward, secondary education patterns are modeled in more detail, allowing for temporary or permanent dropout and part-time studies. LifePaths models graduation at any age, but restricts the detailed modeling of education careers to graduations up to the 25th birthday. Later adult graduation is implemented by a graduation event without tracking study enrolment statuses.

3.7.2 Postsecondary education

Universities provide the clearest picture of the growth of postsecondary education in Canada, where university graduation rates (a bachelor's degree or higher) have quadrupled from below 8% to 30% over the 50 year birth cohorts from 1926 to 1976 as observed in the 2006 Census. For other components of the postsecondary system, the picture is not so clear-cut. This is because provincially organized systems of community colleges did not come into existence until the 1960s and because the boundaries between university and other postsecondary institutions have changed over the years, two examples being nursing and education.

In LifePaths, a person's postsecondary education fate is determined immediately upon graduation from secondary school. These fates represent the person's lifetime postsecondary attainment and include the following possibilities: remaining a secondary school graduate, various combinations of trade vocational school certificates, other types of non-university certificates and university degrees. The fates also include dropping out of either university or college. In all, 30 distinct fates are represented.

Once a person's postsecondary fate is decided, a variety of decisions remain to be made. People fated to attend more than one type of institution choose the type of institution they will attend first. Then, for people who are going on to university, the province of study is chosen. The sex of the individual, as well as their province of residence, influences the choice of province of study. Individuals studying in non-university institutions always study in their province of residence.

Once the province of study is known, the waiting time to the start of the next program is calculated. This waiting time depends on the province of study, the postsecondary fate and the sex of the individual. People going to college or university choose one of 100 fields of study appropriate to their institution. They also choose, based on the field of study, an appropriate pattern of study. This pattern of study represents the total length of the program, as well as the split between full time, part time and out of school years.

People attending a trade vocational institution choose one of six program types. Those who choose the registered apprenticeship program are employed full-time as part of their program and their program lasts longer than other trade vocational programs. Once the type of trade vocational program is chosen or if they are entering a private institution, then the length of the program is chosen immediately.

The 1996 Census was used to derive the probabilities of the 30 possible postsecondary education fates as well as probabilities of the 100 possible fields of study for each sex, year of birth and place of birth. The university and college student administrative information system was used to derive probabilities of province of study as well as the probabilities of patterns of study within each field of study. The Labour Force Survey and National Graduates Survey were used to derive waiting times to the next study program. The 1995 Census of Students—conducted by the National Association of Career Colleges—was used to derive the distribution of lengths of private institution programs.

3.8 Employment and employment/unemployment insurance

In LifePaths, employment activity is partitioned into various modules. Employment for full-time students is dealt with separately because its transitions are determined, in part, by the school year cycle. Career employment refers to the mainstream labour market that has its own autonomous dynamics. Once employment status is determined, LifePaths models the actual hours worked per week.

The current release contains an expanded UI/EI subsystem that calculates premiums and benefits under all the UI/EI Acts up to 1997. The LifePaths model focuses on regular UI/EI benefits as well as maternity and parental benefits. There are no provisions for other special benefits such as sickness or compassion. The UI/EI model also has no provisions for individuals who may also work while on UI/EI. UI/EI premiums are calculated for those eligible to make such contributions. The UI/EI model keeps work and earnings histories in order to calculate benefits.

3.8.1 Student employment

Given that an increasing proportion of young people go on to postsecondary education, it has become important to identify their possible sources of financial support. In that regard, the sensitivity of student employment to the business cycle is significant. Trends in student employment are, in part, explained by growth in postsecondary education. Nevertheless, there has been an increasing trend toward part-time work, even among secondary school students. Most employment undertaken by full-time students occurs during the summer break. As a result, the distinguishing features of the employment patterns of full-time students are short employment spells and the concentration of events at times determined by phases of the school year.

LifePaths classifies full-time student employment status as not employed, self-employed or as a paid employee. Self-employment is distinguished from paid-employment for students, even though only 10% or less of employed students report themselves as self-employed. However, only two transitions—not employed to employed and employed to not employed—are modeled. When a transition from not employed to employed occurs, LifePaths then makes the decision whether to assign the self-employed or paid employee status.

The chances and timing of either transition are assessed monthly. Variables influencing transition probabilities include sex, province, status in the parental home, variables representing seasonal employment patterns and time trends. Given that the timing and length of summer breaks vary by school type, the transition probabilities were estimated for each of four types: secondary, college, university and other schools. The transition from not employed to employed was estimated separately for students with and without prior work experience.

The modeling is based on the results from analyses of student work, using longitudinal data constructed from hundreds of waves of the LFS from 1976 onward.

3.8.2 Career employment

As the nature of the labour force has changed over time, so too have the concepts used in employment data—for example, until 1941, the census used the concept 'gainfully occupied.' The concept was suitable for the time period, when farmers and farm workers represented a significant proportion of all workers and when unpaid family workers, often children younger than 15 years of age, were a significant proportion of agricultural workers. The importance of agriculture gradually declined from representing about 33% of all workers in 1921 to 16% in 1951. Starting from that point, it becomes possible to track annual changes in the labour market. Among the most striking trends of the whole period were the declines in the participation of males aged 65 and older and the increased participation of females aged 25 to 44.

Annual participation rates, whether displayed as trends over time or as age profiles, may give a misleading sense of stability. Underlying these patterns are high rates of job turnover. From the mid-1970s to the mid-1990s, an annual average of 21% of jobs terminated with a permanent lay-off and another 15% were interrupted with a temporary lay-off. Correspondingly, 22% of jobs were newly acquired each year. Some individuals hang on to jobs for decades, whereas others move between jobs with great frequency. In some cases, the jobs involved are seasonal or part time, but many individuals change jobs frequently for other reasons. An adequate model of labour market activity needs to reproduce features of the real market that combine gradual changes in behaviour associated with age, cyclic changes associated with annual variations in business conditions and sub-annual seasonality or volatility.

Certain demographic groups, such as children younger than 15 years of age and institutionalized people, do not participate in the labour market. Consequently, the first step that LifePaths takes in modeling employment is to determine whether an individual is eligible at all for an employment status transition. There are two special cases of ineligibility: retirement, which is treated as a permanent state, and certain people who never work throughout their lifetime. Women with primary-only education or who were born before 1930 make up a large proportion of this latter group, as do people with a lifelong disability. Women on maternity leave experience no risk of job loss for the duration of their leave.

If LifePaths determines that a person can participate in the labour market, their employment status can change from day to day. LifePaths uses a three-category classification of employment status: not employed, employed as a paid employee and self-employed. Based on this classification, seven transitions are distinguished:

- a person who is not employed becomes a paid employee
- a person who is not employed becomes self-employed
- a paid employee becomes not employed
- a paid employee becomes self-employed
- a paid employee changes jobs with no interruption
- a self-employed person becomes a paid employee
- a self-employed person becomes not employed.

The chances and timing of each transition are re-assessed monthly or whenever a transition occurs. The variables influencing employment transition probabilities include duration of current job, duration not working, age, sex, province of residence, education level, the presence and ages of children, the presence of a spouse and the spouse's employment status. Also included are variables representing time trends and monthly seasonal patterns of employment.

Each year after an individual reaches 60 years of age, an independent process determines whether the individual will permanently retire. The chances and timing of retirement are influenced by age, sex, year, region of residence and education level. A forced retirement also occurs when an individual is institutionalized. After retirement, no further labour market transitions can occur.

In the same manner as student employment, employment transition probabilities were estimated from Labour Force Survey data collected from 1976 onward.

3.8.3 Hours worked per week

Once employment status is determined, LifePaths models the actual hours worked per week. Hours worked are dependent on a variety of individual characteristics, including sex, level of education, age, earnings rank, province of residence, family characteristics, student status, job tenure and the month of the year.

3.8.4 Maternity leave and benefits

Income support during maternity leave is a special benefit of the Employment Insurance Program that was first introduced in the 1972 Unemployment Insurance Reform. More recently, the benefit was extended to apply equally to adoption and to cover fathers as well as mothers. Nevertheless, conventional maternity leave remains the most commonly used form. When the maternity benefits program was first enacted, a mother could receive a maximum of 15 weeks of benefits. The introduction of parental leave in 1990 provided an additional 10 weeks, which could be shared between mother and spouse. Examining how maternity leave durations are distributed shows that the peak frequency shifted abruptly from four months duration during the years from 1976 to 1989, to six months duration during the years from 1991 to 1995. This strongly suggests a behavioural change, with some women who would formerly have returned to their job after 15 weeks of maternity leave taking an additional 10 weeks of parental leave. In contrast, a substantial proportion of women—15% to 20%—take essentially no leave regardless of legislation.

In LifePaths, a pregnant woman makes decisions concerning her labour market activities four months before she gives birth. Different decision-making mechanisms apply depending on employment status. Women who are not employed withdraw from the labour market until after the birth. At that point, the transitions to work (becoming a paid employee or self-employed) are determined by the career employment component. An employed woman makes three decisions in sequence: whether to take maternity leave or quit her job, when to start maternity leave or when to quit, and when to return from maternity leave. Nevertheless, unless she has chosen to quit, she remains employed throughout the duration of her maternity leave.

The probabilities determining maternity leave choices and leave durations were based on an analysis of data from Labour Force Surveys (LFS) and the Survey of Labour and Income Dynamics (SLID). SLID data were used to obtain initial patterns of maternity leave, while the LFS data were used to benchmark these initial patterns to the LFS levels for years not covered by SLID—especially the years before 1993.

LifePaths has not yet implemented the changes to parental leave introduced in 2001 (increasing parental leave to 35 weeks) nor has it introduced the 2006 change in Quebec where employment insurance maternity and parental benefits were replaced by the Quebec Parental Insurance Plan.

3.9 Earnings

Annual earnings can be decomposed into weekly earnings rates and weeks worked per year. There is a distinct age profile to annual earnings. They grow fastest at young ages, plateau and then start declining. The peak age of average earnings has been increasing in the past few decades. In 1970, the peak was reached at the age of 40 whereas in 1995 it was reached at 51. During the same decades, the growth in average annual earnings for people younger than 25 years of age decreased. In 1995, people aged 24 and younger had real average earnings that were smaller than those of their counterparts in 1990. This may be because of a variety of factors, such as people staying in school longer.

In LifePaths, each person is assigned a potential weekly earnings rate whether or not they are employed. This potential weekly earnings rate is then combined with employment status and hours worked information to produce actual earnings. Once a person starts a job, whether in paid work or in self-employment, LifePaths assigns weekly earnings in current dollars. Earnings are derived differently for student work as opposed to career work.

Weekly earnings for non-students are derived using three separate sets of equations: wages and salaries, positive self-employment earnings and negative self-employment earnings. For the self-employed, the first step is to decide whether or not this person is going to have positive or negative earnings. Variables influencing earnings include, age, education level, field of study, sex, province of residence and immigration status. These variables can each independently influence the general level, dispersion and degree of asymmetry in simulated earnings distributions.

Weekly earnings are recalculated at the start of a new job as well as on job anniversaries. They will also be recalculated whenever the variables influencing earnings change. A correlation was calculated between a person's rank in the earnings distribution from one year to the next.

Student earnings are treated differently than career earnings. The earnings equations are much simpler and depend on level of schooling and the time of year (summer or school year). They are recalculated each September and at the beginning of the summer. For now, only postsecondary students are eligible to receive student earnings in LifePaths.

Weekly earnings rates and the logit equation that decides whether a self-employed person will receive positive or negative earnings were derived from census data. For the years before 1980, the earnings rate was deflated using a factor derived from a consistent historical series of average weekly wage rates (industrial aggregate). The correlation in the person's year to year ranking in the earnings distribution was derived using the Longitudinal Administrative Databank. Student earnings were derived from the Labour Market Activity Survey and the Labour Force Survey.

3.10 Taxes and family benefits

LifePaths implements the Canadian tax-benefit system and its changes over time at a high level of detail. Personal income taxes in Canada began in 1917. LifePaths models a number of historical tax programs that have since been eliminated (e.g., Old Age Security Tax, National Defence Tax, Graduated Income Tax and various surtaxes throughout the ensuing years) and the various deduction/credit programs that have come and gone since the inception of income tax.

At the end of each year, the federal tax form is completed. This involves the calculation of total income, net income, child care expenses, personal exemptions, taxable income, non-refundable tax credits, the child tax credit, the federal sales tax credit and federal income tax as well as any applicable surtax. Also at year-end and when applicable, the Goods and Services Tax credit is calculated. On a quarterly basis, the Goods and Services Tax credit is assigned to eligible recipients. The federal taxes are also calculated at the time of death.

Along with personal income taxes calculated at year-end, each month payroll taxes such as employment and unemployment insurance premiums, as well as pension plan contributions, are calculated.

On the benefits side, family allowances are calculated each month. These include the federal family allowance, including the Alberta and Quebec configurations, as well as Quebec’s Allowance for Newborn Children. Other benefits calculated on a monthly basis are the Canada Child Tax Benefit, which is additionally recalculated at any changes in family structure, and the Universal Child Care Benefit.

3.11 Pensions and retirement savings

LifePaths simulates the Canadian retirement income system with its various public and private components and senior benefit programs in a high degree of detail, reflecting its historic evolvement and changes in legislation, plan designs and plan participation. Public components are the Canada and Quebec Pension Plans, Old Age Security, the Spouse's Allowance and the Guaranteed Income Supplement. The corresponding LifePaths modules are largely rule-based for the determination of eligibility and benefit calculation, and full take-up is assumed. Private pension programs modeled in LifePaths consist of two tax-assisted retirement savings plans: the employer-sponsored Registered Pension Plans (RPPs) and the individual's Registered Retirement Savings Plans (RRSPs).

3.11.1 The Canada and Quebec Pension Plans retirement pensions

The Canada Pension Plan (CPP) is a contributory, earnings-related social insurance program that ensures a measure of income protection to a contributor and their family against the loss of income because of retirement, disability or death. The plan operates throughout Canada except in the province of Quebec, where a similar program—the Quebec Pension Plan (QPP)—is in effect.

The Canada Pension Plan was enacted in 1965 and came into force on January 1, 1966. Contribution collection began on January 1, 1966, retirement pensions became payable in January 1967, survivors' benefits in February 1968 and disability benefits in February 1970.

The CPP covers virtually all employed and self-employed people in Canada (excluding Quebec, which operates the QPP) who are from 18 to 70 years of age and who earn more than a prescribed minimum level of earnings in a calendar year. This minimum level, which changes annually, is known as the Year's Basic Exemption (YBE).

CPP is financed through contributions from employees, employers and self-employed people. Contributions are required on the portion of an individual's annual earnings above the YBE and up to a prescribed ceiling known as the Year's Maximum Pensionable Earnings, which is linked directly with the average wage. The QPP is closely associated with the CPP and both plans are operated and coordinated through a series of agreements between the federal and Quebec governments. Benefits from either plan are based on pension credits accumulated under both, as if only one plan exists.

Retirement pensions are available to individuals any time from the month following their 60th birthday, providing they have made at least one valid contribution to the CPP. To receive a retirement pension before age 65, individuals must have wholly or substantially ceased pensionable employment at the time the benefit commences. Starting in September 1987, CPP flexible retirement is available at age 60 (70% of pension) through age 70 (130% of pension). Starting in September 1984, QPP flexible retirement is available at ages 60 to 70. Once an age adjustment is established, it is permanent and applied to each month benefits are received, regardless of the age of the beneficiary. The amount of the retirement pension before 1976 reflects the legislation, which states that any retirement pension is calculated on the basis of a minimum of 10 years of contributions. The pension is designed to replace about 25% of the earnings on which contributions have been paid. Once qualified for a pension, CPP benefits can continue to be collected even if the pensioner has returned to work; however, no further contributions to CPP are deducted.

LifePaths keeps track of the full working history and contribution career, thus being able to implement the complex system rules and to support analysis of future policy options. Besides the mechanics of calculating contributions and benefits, LifePaths also models the timing of pension take-up, as—beginning in 1986—individuals can opt to collect benefits before age 65—i.e., starting at age 60. The detailed modeling of family demographics enables LifePaths to implement survivor, combined and death benefits.

3.11.2 The Canada and Quebec Pension Plans survivor benefits

Spouses of deceased contributors may be eligible for a survivor's pension if the deceased contributor made contributions for a sufficient number of years. Surviving spouses with dependent children, surviving spouses with a disability and those aged 45 and older are eligible for survivor benefits. For surviving spouses who are younger than age 45 at the time of the contributor's death and who are neither disabled nor caring for the contributor's dependent children, benefits are reduced by 1/120th for each month the survivor is younger than age 45.

Since January 1987, the CPP survivor's pension continues to be payable on remarriage and benefits are reinstated for people who remarried prior to this provision coming into effect. However, no retroactive payments are paid for any time before January 1987. Since January 1984, recipients of a surviving spouse's QPP pension continue to receive their pension if they remarry. There are various conditions that determine when the survivor's pension will cease. If the survivor is older than 35 years of age at the time of the contributor's death, the pension will end at the survivor's death. If the survivor was younger than 35 at the time of the contributor's death, the pension will end when the survivor is no longer disabled, no longer raising the deceased contributor's dependent child or upon the survivor's death.

A person who is receiving a surviving spouse's pension may also have contributed to the CPP on the basis of their own earnings and may thus be entitled to a retirement or disability pension. There is a limit on the combined benefit based on the maximum retirement pension, which one would receive if they were 65. Depending on the retirement pension, the survivor benefit could be reduced.

3.11.3 Canada Pension Plan pension options

Various CPP pension options have been incorporated into the LifePaths model. These options allow users to model various changes to the CPP program such as changing the starting point of earnings at which contributions begin, the upper level of pensionable earnings, the contribution rates for income ranges, benefit rates for pensionable earnings ranges, and options for drop-out provisions in the calculation of pensionable years of earnings.

3.11.4 Old Age Security

An Old Age Security (OAS) pension is a flat-rate benefit payable to people who fulfill certain residence requirements. The age of eligibility was originally 70, but it was lowered to 69 in 1966 and progressively by one year in each succeeding year to 65 in 1970. The pension amount is based on length of residence in Canada currently subject to a minimum of 10 years of residence. The pension amount is reduced for people with higher incomes—LifePaths implements this clawback through the income tax system. OAS pensions are considered as taxable income.

From 1952 to 1967, benefit rates have been periodically adjusted. From 1968 to 1970, the OAS pension amount was adjusted annually in January in line with changes in the Pension Index, with its 2% ceiling, which had been developed for the CPP. (The Pension Index is defined for any given year as the average of the Consumer Price Index for a 12-month period ending June 30 of the preceding year, with an upper limit of 2%.)

Presently, benefits paid under the *Old Age Security Act* increase in January, April, July and October of each year, when there is an increase in the cost of living as measured by the Consumer Price Index. If the Consumer Price Index declines, there is no reduction in payment.

3.11.5 Guaranteed Income Supplement

A Guaranteed Income Supplement (GIS) may be added to the OAS pension of a pensioner who has no or limited other income. Entitlement is normally based on income in the preceding calendar year, calculated in accordance with the *Income Tax Act* but excluding OAS payments and certain other sources of income. In determining entitlement for a married applicant, the incomes of both spouses are taken into account.

If a person is receiving a partial OAS pension (because of residency requirements), the maximum GIS is increased by the difference between the partial pension and the full OAS pension. GISs are not considered as taxable income.

3.11.6 Allowance for the Survivor

Spouse's Allowance dates back to 1975. The program was designed to provide benefits equivalent to the OAS pension and the GIS. The initial target group was people in need aged 60 to 64 who were married to GIS recipients.

Individuals who received Spouse's Allowance used to lose all their benefits when their pensioner spouses died; however, that inequity was eliminated in 1979 under legislation that allows recipients to continue receiving benefits to age 65, when they become eligible for the OAS pension and GIS.

A Spouse's Allowance may be paid to the spouse of an OAS pensioner if the spouse is from 60 to 64 years of age. A spouse may be entitled to receive the allowance if he or she has a minimum of 10 years of residence in Canada after the age of 18. The spouse must be a Canadian citizen or legal resident in Canada when the application is approved, and the combined income of the couple must not exceed the maximum permitted by the law. Widows and widowers 60 to 64 years of age are also eligible to receive Spouse's Allowance benefits provided they meet the residence requirements and have limited income. The Widowed Spouse's Allowance stops if a widow or widower remarries. Spouse's Allowances are not considered as taxable income.

Allowance stops when the recipient becomes eligible for an OAS pension at age 65. The benefit also ceases to be payable if the pensioner spouse becomes ineligible for GIS or the couple becomes separated or divorced.

The maximum full monthly Spouse's Allowance is equal to the full basic OAS pension plus maximum GIS at the married rate. The maximum amount for a Widowed Spouse's Allowance is somewhat higher.

3.11.7 Registered Pension Plans

Registered Pension Plans (RPPs) are retirement benefit programs provided voluntarily by employers or by unions in both the public and private sectors of the economy and are funded both by employees' and employers' contributions. There are two basic types of RPPs: defined contribution (DC) and defined benefit (DB). Under the DC plan, the employer and employee are committed to a specific contribution rate. Under the DB plan, benefits are determined by a formula that is stipulated in the plan text.

The first step LifePaths takes in modeling RPPs is to determine whether a person will be a new RPP member. To be considered at all, a person must first be a non-student, a paid worker, without any current RPP plan, and from 15 to 55 years of age. The chances and timing of becoming a new member are re-assessed each year or whenever a transition to employment occurs. The variables influencing the probability of becoming a member include the sector of the job (public or private), age, sex, year, year of birth, earnings, and, for the private sector only, occupation group, education, previous year of labour force attachment, age at immigration and province of residence.

If the person is determined to become a new member, he or she will be assigned one specific plan in all its detail, including the type of the plan (DB or DC) and plan provisions (e.g., contribution rate, benefit rate, integration and indexation). The variables influencing the type of plan include the sector of the job, year of birth, and, for the private sector only, sex, education, province of residence, occupation group and wage placement in earnings distribution. The portability of the plan is also determined at this time. A portable plan is one that assumes the continuity of the plan membership—i.e., allowing accumulation of years of service from job to job.

A new member's starting RPP status is always 'Current Active.' With this status, along with the plan's portability setting, the member enters cycles of RPP plan status changes that are tightly interlaced with their labour force status and transitions (modeled in the career employment module). The RPP module uses the following five categories of RPP status, taking into account the portability of a plan:

- Current Active: Member currently employed.
- Current Inactive: Member not employed, plan portable.
- Deferred: Member not employed, plan not portable, years of service are less than or equal to two.
- Closed: Member not employed, plan not portable, years of service are greater than two.
- In-pay: Member or survivor receiving benefits.

These status changes occur whenever there are labour force transitions from employed to non-employed or from non-employed to employed, when retired, and in the event of a person's death. All other maintenance activities for the plan, such as updating years of service and other information about the member, making contributions, calculating pension adjustment amount, determining DC asset accrual, checking if it is time to take-up benefits and calculating benefits are performed at the end of every month for each given plan status and type. Scores of detailed assumptions have to be made along the way in accomplishing these activities.

New membership probabilities and probabilities of obtaining a specific type of plan were initially estimated from the Survey of Financial Security. Pension adjustment (PA) and Registered Retirement Savings Plan (RRSP) data (tax data) were then used as a first step adjustment to the new member (PA) probability for the dimensions sex, age, year and earnings. The final adjustment used total member counts by sex and sector from the Pension Plans in Canada Survey.[1]

3.11.8 Registered Retirement Savings Plans

RRSPs began in 1957. In that year, an amendment to the *Income Tax Act* permitted individuals to make deposits into personal savings plans for future retirement income and receive tax savings not only through deducting the tax of contributions, but also through exempting the tax of investment income earned within RRSPs. RRSP savings are accessible anytime. However, if withdrawals are made, the amount withdrawn will be taxed the same as one's income. Accumulated savings must be converted into annuity income or registered retirement income funds (RRIFs) before the end of the year in which the beneficiary reaches a specified age (currently 71 years).

The RRSP module in LifePaths consists of four components:

- contributors and contributions
- asset growth
- cash withdrawals
- conversions of RRSP assets to RRIFs or annuities.

1. Canadian Socioeconomic Information Management System (CANSIM) table 280-0010.

LifePaths places all RRSP activities at the end of each year. First of all, a person's new RRSP room is calculated according to the legislated rules, taking into account whether a person has a PA. This new room is to be used for RRSP purposes for the coming year (year+1). To be eligible to make an RRSP contribution, a person's total RRSP room at this time must be positive—i.e., a positive sum of new room from the previous year (year-1) plus all accumulated room that was previously unused, if applicable. If this requirement is met and the person is from 15 to 69 years of age, then they will go through a process that determines whether they will contribute to an RRSP this year, and if yes, the contribution amount. This amount is then deducted from the total room this year.

People with positive RRSP assets can withdraw cash. There is no information available, however, to determine whether a person is withdrawing their assets down to zero; this behavioural detail is absent in LifePaths for people younger than age 65. For people from 65 to 69/71 (of conversion age), LifePaths has a withdrawal plan. Variables influencing the probabilities and size of contributions and withdrawals are sex, age, marital status, earnings, whether that person is an RPP member with a PA, lagged contributions and withdrawals, province of residence, an individual specific random component with zero mean, and status change variables such as whether the person was just married last year; just separated, divorced or widowed last year; and whether they moved from another province last year.

Conversions from RRSP assets to either an RRIF or an annuity take place from age 65 to the mandatory conversion age of 69/71. The decision to convert is according to a set of proportions cross-classified by couple or single status, sex, tax year and age (65 to 69/71). If a conversion is determined, then the annual pension amount is calculated according to the type of conversion (RRIF, annuity or both). For an annuity, a fixed factor is applied to the assets to determine the annual pension. For an RRIF, the minimum percentage required at each age is applied upon the remaining assets to obtain the annual pension.

The asset growth component posed challenges for LifePaths. Using a set of historical rates of returns for some financial instruments such as treasury bills, GICs, bonds and Canadian and U.S. stock indices, almost all other aspects of investments are implemented according to some reasonable assumptions. LifePaths classifies the financial instruments into high-risk (e.g. stocks), medium-risk (e.g., corporate bonds) and low-risk (e.g., GIC) groups. LifePaths then assumes varied risk tolerance scores among people in different sex, age and earnings groups, and translates these scores into a varied mix of high-, medium- and low-risk investment portfolios. The historical rates of returns are the net of Management Expense Ratios and an additional penalty rate that compensates for the inefficiencies in the individual retail investment market. For future rates of returns, options are made for either using historical average rates of returns by risk class or a random pick of intervals of years within historical data series.

PA and RRSP longitudinal tax files (1990 onward) were used to estimate the probabilities and amounts of RRSP contributions and withdrawals. Greenbook Tax Files were used both for inferring the proportions to have a withdrawal plan within ages 65 to 69/71, and the proportions to convert RRSP assets to RRIFs or annuities. Historical patterns of contributors and contributions before 1990 are adjusted using information from Statistics Canada's publication "RRSPs: Tax-assisted retirement savings."[2] Data on historical rates of returns for various financial instruments are from Canadian Socioeconomic Information Management System (CANSIM) tables and the "Report on Canadian Economic Statistics 1924-2002" of the Canadian Institute of Actuaries.

3.12 Saving and wealth

The inclusion of (non-pension) saving and wealth is a new area of module development, currently focusing on income adequacy at retirement. At this time, models are available for home ownership and non-registered assets and debts. LifePaths does not model the ownership of businesses. An important component still missing are intergenerational and family transfers and bequests.

3.12.1 Housing

The LifePaths housing module models home ownership, house value (changes) and house equity. The model operates in yearly steps and is cross-sectionally consistent with census distributions of house values by income and household type. Longitudinal consistency is enforced by two main simplifying assumptions: First, individuals usually stay in ownership. Second, besides the general change of house values observed in census data, house value changes over the life course are driven by changes in household type (e.g., marriage) and change in income (5% income quantile). Distributions of house values are parameterized by constructing 5% house value quantiles within each income quantile and household type. People are assumed to stay within the same house value quantile over their life course. Thus we allow upward and downward mobility because of changes in income and household type, but assume that the (unobserved) factors (e.g., taste and regional factors) placing a person into a house value quantile within a given income quantile are age invariant.

Updates are on a yearly basis and are processed in five steps:

1. Mortgage payments are made, if applicable. Payments depend on historic and projected interest rates and the duration of the mortgage, which is assumed to be 25 years for the initial mortgage taken up at home purchase, and 10 years for additional mortgages taken up for upgrades.
2. It is decided if a house is being sold. This only occurs in the case of union formation or dissolution. In the case of dissolution, the equity is divided between partners. Cashed out equity is used the next time each person purchases a house.

---

2. Frenken, Hubert. 1990. "RRSPs: Tax-assisted retirement savings." *Perspectives on Labour and Income.* Vol. 2, no. 4. Statistics Canada Catalogue no. 75-001-X1990004121. http://www.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=75-001-X1990004121&lang=eng (accessed 2013-06-03).

3. It is decided if a house is purchased. For those who never owned a house, we use the increments observed in census data. Those who sold a house alongside a union formation event immediately buy a new home according to the updated household characteristics. In the case of union dissolution, the probability to buy a new house is the age-specific probability of a single person to own a house. The algorithm accounts for different probabilities of former renters and owners to enter ownership after a divorce.

4. The house value is assigned or updated. House values are drawn from distribution tables by 5% income quantiles and 5% house value quantiles within each income quantile, by calendar year and family status.

5. Finance decisions are made. For a first purchase, a down payment is sampled by age and household type. For following purchases, the cashed out equity is used as a down payment. Upgrades are financed by a second mortgage. Whenever an upgrade occurs, the duration of the second mortgage is re-initialized to 10 years. Equity cashed out because of downsizing is used to pay off first the second, then the first mortgage.

Ownership rates and their increments as well as the distribution of house values by income quantile are based on census data. The distribution of down payments (as a proportion of the house value) for first-time buyers by family characteristic and age at purchase is based on the 2005 Survey of Financial Security. We assume that all changes in house price values outside the range of +/-5% of the Consumer Price Index are because of investment or downsizing and that a fixed proportion of investments is financed by a second mortgage.

3.12.2 Residual (non-registered) assets and debts

The module for non-registered assets is built on top of the existing LifePaths model. The modeled assets are comprised of financial assets (other than pension savings) and properties (excluding primary homes). Both stocks and annual flows are modeled. The module closes a gap in the modeling of savings in LifePaths which was previously restricted to various vehicles of registered pension savings (RRSPs and RPPs) as well as equity in housing. Saving behaviours are modeled conditionally on choices concerning registered pension savings (RRSPs), as well as home ownership and equity.

The three main data sources available and used for the modeling of non-registered assets are (1) the Survey of Financial Security (SFS) and its predecessor Asset and Debt Surveys for stocks of wealth, (2) the Survey of Household Spending (SHS) and Family Expenditure Survey (FAMEX) for information on the annual flows of saving and dissaving, and (3) Longitudinal Administrative Databank (LAD) for data on longitudinal investment income.

Concerning data, serious challenges had to be faced:

- The absence of longitudinal data (except for dividends, interest, rental income and capital gains in LAD).

- Separate datasets for stocks and flows.
- The incomplete collection of data in the case of mortgages, where there's no distinction made between principal and interest payments.
- Some potential saving components are not well defined, for instance personal insurance which can contain pension saving components.
- Possible issues with the quality of some of the data due to small sample sizes and the possibility of larger than average non-sampling error for some data.

The most important consequence for the development of the Non-Registered Assets module is that a statistical approach in the way it is done in many other LifePaths components cannot be followed. Specifically, observed saving rates and their distribution can only provide calibration targets, while theoretical assumptions are needed to produce individual level saving rates. SHS data reveal a very wide range of saving rates with almost 50% of people age 20-64 not saving at all or having negative saving. Such snapshot pictures contain no information how saving rates change over time for individuals. The variance in savings rates stems from several factors which are unobservable in cross-sectional data. For example, it might be reasonable to assume that observed negative saving rates are an effect of consumption smoothing (e.g. are produced by individuals who experience a temporary income loss due to unemployment) or result from purchases of consumer durables like vehicles which typically do not occur every year.

Saving decisions are modeled as yearly events for persons up to age 64. The following describes the basic algorithm;

1. Calculate a "statistical" saving rate. This rate can be interpreted as the expected saving rate in absence of income change and irregular purchases of vehicles. Individual level characteristics taken into account are age, calendar year, sex, family characteristics, RRSP savings, home ownership and income quantile. The calculation is based on regression models estimated from SHS data.
2. Apply a "taste factor". We assume, that people are different in "taste", thus some always save more while other always save less in otherwise identical circumstances.
3. Consumption smoothing: We assume that people make additional savings in case of increased income and make withdrawals in the case of income loss. Withdrawals are subject to budget constraints.
4. Account for vehicle purchases. Vehicle purchases follow a statistical model based on SFS data and can occur at any time of the year.
5. Calculate resulting savings or withdrawals. Finalize the saving decision (if saving is positive, make a final decision taking into account a probability of additional expenditures).
6. Make portfolio choices.

Assets savings and debts are implemented by a system of individual level accounts. For instance, each individual has an account for financial assets (and debts) and for properties. Accounts are linked to individualized rates of return and costs of borrowing. Positive returns are subject to taxation. All models are parameterized separately for single males, single females, and couples. In the case of couples, income is split and both partners save the same amount holding separate accounts. Besides the regular yearly saving events, deposits and withdrawals are made by various events, e.g. when purchasing a vehicle; when purchasing, upgrading, or selling a home; when taking out RRSP savings for home purchases; and or for the financing of post-secondary studies (student loans).

While highly stylized and operating on the national level only, the module very well reproduces the wealth distributions observed in data especially of people in the age-range 60-64. Detailed documentation and validation results are available.

3.13 Health

LifePaths models health status in two ways: the first based on a five-level scale of disability; the second on the Health Utilities Index (HUI), which is a continuous measure of health.

3.13.1 Disability

LifePaths models transitions between five levels of disability—the most severe being permanent institutionalization, an event that depends on the level of a person's disability as well as their age and sex and the current year. Disability is defined in terms of the characteristics most likely to be associated with the need for assistance in performing everyday activities, concentrating on four activities: everyday housework, grocery shopping, meal preparation and personal care. In turn, these activities were posited to depend primarily on the following kinds of disabilities or impairments: mobility, dexterity, cognitive capacity and pain. In effect, the disability indicator has been tailored to reflect aspects of the assistance that may be required to maintain independent living.

Reasonable disability data are available mainly starting from the 1990s. In particular, the National Population Health Survey (NPHS) provides longitudinal data from 1994 to 2000, as well as cross-sectional prevalence. The NPHS covers both the household population and those in institutions, and the survey included several measures of disability. As a result, the NPHS has provided the basis for both prevalence distributions of disability by level of severity, and microlevel estimates of transition dynamics.

Transition probabilities for movements between each possible combination of these disability states were estimated using longitudinal NPHS data spanning the 1994, 1996, 1998 and 2000 cycles of the survey. To allow for transition probabilities varying not only with the current disability state but also with the individual's prior disability trajectory, the estimation draws on both lagged and current disability status.

3.13.2 Health Utilities Index

The other health indicator modeled in LifePaths is the HUI (see www.healthutilities.com/HUI.htm). This index provides a description of an individual's overall functional health using eight attributes: vision, hearing, speech, mobility (ability to get around), dexterity (use of fingers and hands), cognition (memory and thinking), emotion (feelings) and pain. Based on responses to a standard set of questions, the HUI provides a summary functional health score that lies between -0.360 and 1.000. For instance, an individual who is nearsighted, yet fully healthy in the other seven attributes, receives a score of 0.973. The most preferred health level on the HUI scale (perfect health) is rated 1.000 and death is rated 0.000, while negative scores reflect health states that are considered worse than death. The HUI score is derived using a weight that has been assigned to each level of the eight attributes. The weights embody the views of society concerning health status. These views are termed 'societal preferences,' and are based on preferences about various health states as elicited from a representative sample of individuals.

As such, the HUI measure may be useful in a broader range of applications than would the disability measure. Nevertheless, neither the disability measure nor the HUI measure would be completely satisfactory as an indicator of disability for the purposes of the Canada and Quebec Pension Plans. For those purposes, it would be necessary to have an indicator that embodies a medical doctor's judgements of health in combination with a respondent's self-reported health.

As with the disability measure, the LifePaths model of HUI dynamics draws on NPHS data: using longitudinal data spanning the 1994, 1996, 1998, 2000, 2002 and 2004 cycles of the survey. To initialize each simulated health trajectory, it is assumed that each newborn is in perfect health (HUI equals 1.000). Thereafter, functional health status is updated on each birthday in two steps:

1. First, it is determined whether or not any change is to be made to the HUI score for the current age. This determination is based on probabilities that depend on the individual's current age, current and lagged HUI score, immigration status, student or employment status, presence of a spouse and broad education level. In addition to direct influences on the likelihood of change, there are random influences representing the degree of variability among otherwise identical individuals and reflecting differences in variance by immigrant and education statuses. This random term is imputed at birth and remains fixed over an individual's lifetime: in effect, it represents an individual's intrinsic susceptibility to change.

2. Then, if it is determined that a change is to be made to the HUI score, the magnitude of the change is imputed using an equation influenced by the same set of characteristics (e.g., age) as in the first step. In addition, there is another random influence that reflects unexplained differences between individuals that represent a person's intrinsic frailty, but with differing variance by immigrant and education statuses. Finally, there is a random influence (with constant variance) that represents the unexplained year-to-year variability experienced within an individual's life.

As a consequence of imputed HUI, a simulated individual's mortality risk is changed: those with a lower HUI are more at risk than those with a higher HUI. For that purpose, relative risks that depend on the simulated HUI score and age have been estimated from NPHS data. These relative risks are unlikely to be accurate for historical simulations (i.e., for periods prior to the NPHS), but are likely to be useful for projections into the future. Similarly, there is a risk of 'double counting' the effects of education and marital status, if those risk factors are directly applied to mortality with HUI turned on.

4 Summary

LifePaths aims to create the best possible empirical picture of the past and future Canadian population's socio-economic status from a life course perspective. The simulation of detailed life histories placed into the evolving tax-benefit and social insurance policy context makes LifePaths a generalized, policy-oriented simulation model capable of framing and answering "what if" questions in many policy domains. While LifePaths has ongoing utility, its greatest importance lies perhaps in its ready availability to analyse proposals that find their way onto the policy agenda with little prior warning. The potential impact of such proposals is often large and is frequently in the multi-billion dollar range. On the other hand, the political time frame of the proposals is often short. This means that the analytical capability to develop or evaluate the proposals needs either to be created quickly or be already in place. The complexity of longitudinal modeling means that there is typically not enough time to develop from scratch the tools necessary to adequately and credibly analyse policy for specific proposals. The LifePaths project fills this gap by providing a foundation on which to build specific, specialized models designed to evaluate new policy proposals.