September 8, 1998

CHAPTER 6

DYNACAN: A CANADIAN MICROSIMULATION MODEL FOR PUBLIC PENSION ANALYSIS

September 8, 1998

CHAPTER 6

DYNACAN: A CANADIAN MICROSIMULATION MODEL FOR PUBLIC PENSION ANALYSIS[1]

I. OVERVIEW

DYNACAN is a dynamic microsimulation model developed and maintained in the Office of the Chief Actuary (OCA) of the Canada Pension Plan.[2] DYNACAN has been developed by the federal government of Canada primarily to analyze alternative policy proposals for the Canada Pension Plan (CPP). The CPP is a mandatory, public, pay-as-you-go pension system.[3] It's role is similar to that of the Old Age and Survivors and Disability Insurance (OASDI) System in the U.S. The Office of the Chief Actuary performs the actuarial valuation of the CPP. DYNACAN is designed to be closely aligned and used with the CPP Actuarial Valuation Model, which is also maintained in OCA.

DYNACAN currently ages the 1971 Canadian Census Public Use Files one percent sample, operating on a sample of up to 213,000 person records organized into families. It ages the database in annual periods from 1971 through 2100. The model depicts most of the important demographic events and characteristics and economic information for persons and families required to calculated CPP contributions and benefit payments. The model can produce simulated cross-section population databases for any simulation year, and longitudinal histories of demographic events, employment and earnings, and retirement for the persons in the database. It has a detailed module for calculating CPP contributions and benefits, with comprehensive representation of CPP rules and program parameters, but little behavioral analysis. DYNACAN does not include models of private or public employer

---

[1] This description draws on OSFI, *DYNACAN High Level Technical Documentation* (1997a), Morrison (1997), OSFI (1997b), Chenard (1996), Neufeld (1997), Swan (1997), Wilde (1997), correspondence and discussions with Richard Morrison of OSFI.

[2] The Office of the Chief Actuary of the Canada Pension Plan is in the Office of the Superintendent of Financial Institutions (OSFI) of the federal government of Canada. In July 1998 it was announced that DYNACAN would be transferred from OSFI to Human Resources Development Canada (HRDC).

[3] The Canada Pension Plan (CPP) is a contributory, earnings-related social insurance program. It provides benefits to a contributor and his/her family for retirement, disability, survivors, and death. The CPP covers all workers in Canada, except in the province of Quebec, where a similar program, the Quebec Pension Plan (QPP) applies. The operation of the two programs is coordinated. The CPP covers virtually all employed and self-employed persons in Canada (except in Quebec) between ages 18 and 70, who earn more than a prescribed minimum level of earnings. The CPP is financed by contributions from employees, employers and self-employed persons, and interest from the CPP fund.

pensions or retirement savings arrangements. It does not model any income source other than earnings and CPP benefit payments, and it does not model savings or wealth accumulation. It does not include occupation or industry affiliation of workers. It does not depict behavior or characteristics of employers or retirement plan sponsors.

II. BACKGROUND

DYNACAN has been developed by the Social Policy Unit of the Office of the Chief Actuary in the Office of the Superintendent of Financial Institutions (OSFI) of the federal government of Canada with input from the Department of Human Resources Development Canada (HRDC). It is the product of a five-year project to develop a policy model to inform CPP policy. DYNACAN is used to estimate the financial impacts on individuals and families of CPP policy proposals.

In 1990, based on a review and initial needs assessment, the Income Security Programs Branch of HRDC concluded that to perform distributional analysis of CPP policy proposals and analysis of interactions of CPP with other programs, a longitudinal dynamic microsimulation model would be required. HRDC and OSFI decided to construct a prototype model and to use an existing model as a template for the prototype. A feasibility confirmation and cost-estimation study was conducted in 1994, and the project was approved. Development was projected to take five years. Models from three countries were investigated for their suitability and adaptability to HRDC requirements. In 1995 the CORSIM 2.0 model developed at Cornell University was acquired to serve as the template for the Canadian microsimulation model.[4] The CORSIM simulation system and, initially, many of the CORSIM modules were adopted to serve as the basis for the new Canadian model. During 1996 the DYNACAN project team began the tasks of assembling the Canadian database for the model, substituting Canadian data and coefficients in the model's demographic, labor force, and earnings modules, and creating new modules. Development continued in 1997 and 1998, including development of a CPP module and initial validation of the model and its application for policy development purposes.

From its inception through 1998, the DYNACAN project team operated under the direction of the Chief Actuary in the OSFI. The Strategic Policy Branch and the Income Security Programs Branch (ISP) of HRDC are heavily involved as the primary prospective users, funder, and provider of data for DYNACAN. DYNACAN's funding comes from the CPP. In October 1998, DYNACAN is to be transferred to HRDC.

The primary objective of DYNACAN is to provide analysis of the effects of CPP policy proposals on persons and families, including distributional analysis and winners-losers. A longer run objective is to analyze the CPP in the context of other sources of retirement income and the interactions of the CPP with other government programs.

[4] CORSIM is described in Chapter 5 of this report.

III. DESCRIPTION[5]

DYNACAN is a large-scale dynamic microsimulation model.[6] It simulates the basic demographic and economic events relevant for calculation of CPP contributions and benefits. The basic units of DYNACAN are persons and families. DYNACAN does not depict firms, industries, or governmental entities. It simulates demographic and economic events, states, and changes occurring for each person and family in its base microdata file on an annual basis. It depicts approximately 14 events, processes, or characteristics of persons and families and operates on a database of up to 213,000 records. The model is usually run over the period 1971-2030, but it can be run for 1971-2100.

Components

DYNACAN has three major components. DYNACAN-A is a set of programs and procedures that generate the starting database for the simulation. DYNACAN-B is a single program that ages the database one year at a time and simulates the demographic, labor force, and earnings events relevant to the calculation of CPP contributions and benefits. DYNACAN-C calculates CPP contributions and benefits, produces reports of the simulation results, and compares simulations of alternative policy proposals. Figure 6-1 shows the major components of DYNACAN and their roles and relationships.

DYNACAN-A

DYNACAN-A prepares the initial input database for the simulations. DYNACAN uses the Public Use Microdata Files of the 1971 Canadian Census to provide the base year database. For this database to be suitable for simulation, several modifications and enhancements were required. The information for persons in the Individuals file was merged with information on persons in the Families file. The resulting file consists of individuals grouped in family units and is representative of the Canadian population in mid-1971.

Items that were not present in the Individuals file were added to the records of individuals from that file, and several variables were recoded (modified without changing their meaning) to make the two files consistent and more useful. Specific ages and sexes of children were imputed to records where they were not present. The age distributions in the sample data were smoothed by sex to correspond to the Census data, and ages were imputed for individuals for whom the recorded age had

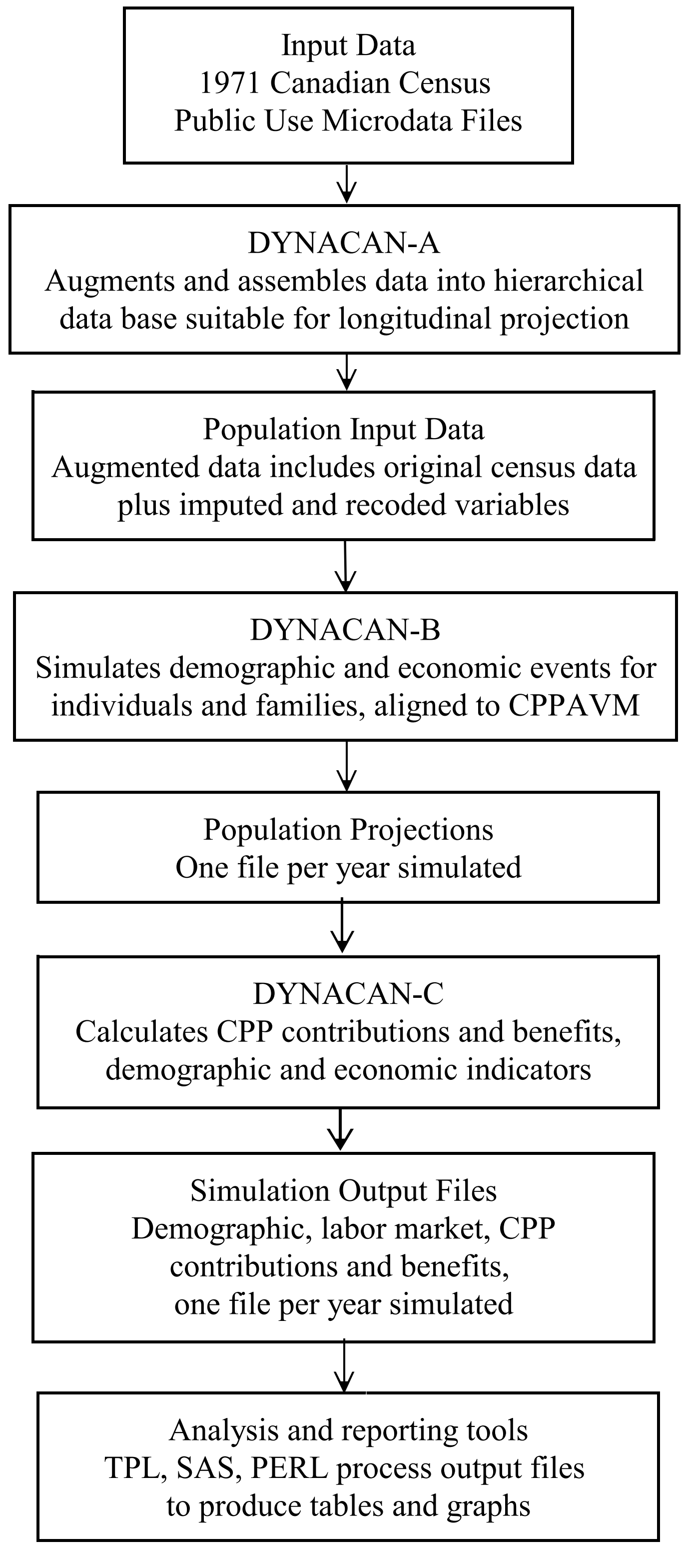
---

[5] The basic features of DYNACAN, including sources of data, are summarized in the Summary Description Table in Annex 6-1 at the end of this chapter.

[6] Dynamic and static microsimulation are described in Chapter 2. As described in Chapter 2, a microsimulation model simulates social and economic behavior by depicting events, conditions, and changes in the information recorded on each individual (person or family) record in a large database. A microsimulation model depicts the aggregate conditions of the population by aggregating or tabulating over all the modified individual records. Dynamic means the model simulates events for each individual one year at a time, and the outcome for each event each year depends on current and past year's outcomes for that and other events.

FIGURE 6-1

DYNACAN COMPONENTS AND FLOW DIAGRAM



been truncated at 75. This was done for the national sample and by region, for Quebec and the rest of Canada. Employment earnings were allocated to family members other than the head and spouse, where appropriate. Individuals' earnings histories were imputed back to 1966, the first year of operation of the CPP, by deflating the reported 1970 earnings by the index of average wages. Disability status for 1971 was imputed. Records of selected females were marked to be excluded from entering the labor force, to assure that the simulations were consistent with survey results concerning the proportions of women who report themselves as never having worked. Education variables in the census were converted to the specific variables used in DYNACAN-B. Individuals in the 1971 database were assigned a variable for "permanent luck" to introduce a systematic divergence of earnings from the values otherwise predicted by model equations, in order to achieve realistic variation in the longitudinal simulations of earnings.

After completing these adjustments, DYNACAN-A produces a population file suitable for input into DYNACAN-B. The full version contains records for 212,935 individuals grouped into families. DYNACAN-A also produces a 20 percent subsample and a 40 percent subsample, for both of which the age distributions are independently smoothed by sex and region. These population files are input into DYNACAN-B for the simulations.

DYNACAN-B

The function of DYNACAN-B is to simulate longitudinal histories of demographic events, employment and earnings for each individual in the database to be used in DYNACAN-C to calculate CPP contributions and benefits. DYNACAN-B is a single integrated simulation program that operates on a microdata file. It executes a set of operations for each year in the projection period that ages by one year an input population. The structure and operations of DYNACAN-B is shown in Figure 6-2.

DYNACAN-B is fully dynamic. Each event, characteristic, or process for each person or family is determined for each year, one year at a time, with reference to variables representing demographic and socioeconomic characteristics of the person or family in that year and previous years, including lagged (previous years) values of the variable being determined. Each of the behaviors, events, or processes depicted in DYNACAN-B is included in an individual module. A module may determine outcomes using rule-based processes or behavioral equations or both. DYNACAN includes 17 behavioral and rule-based or bookkeeping modules. Table 6-1 shows the events and processes depicted in DYNACAN and the variables used to determine each event. The order in which the modules are solved is shown in Table 6-2.

DYNACAN-B simulates demographic and labor market events for each individual in the file each year. The first event/process is determined for all individuals, then the next event is determined for all individuals, until all events and outcomes have been determined for all individuals and families in that year. Then DYNACAN-B repeats this process for the next year with the simulated event values for the previous year and those already determined for the current year as inputs. The process is repeated for each year until the final year of the simulation period. The output of DYNACAN-B

FIGURE 6-2

DYNACAN-B STRUCTURE AND OPERATIONS

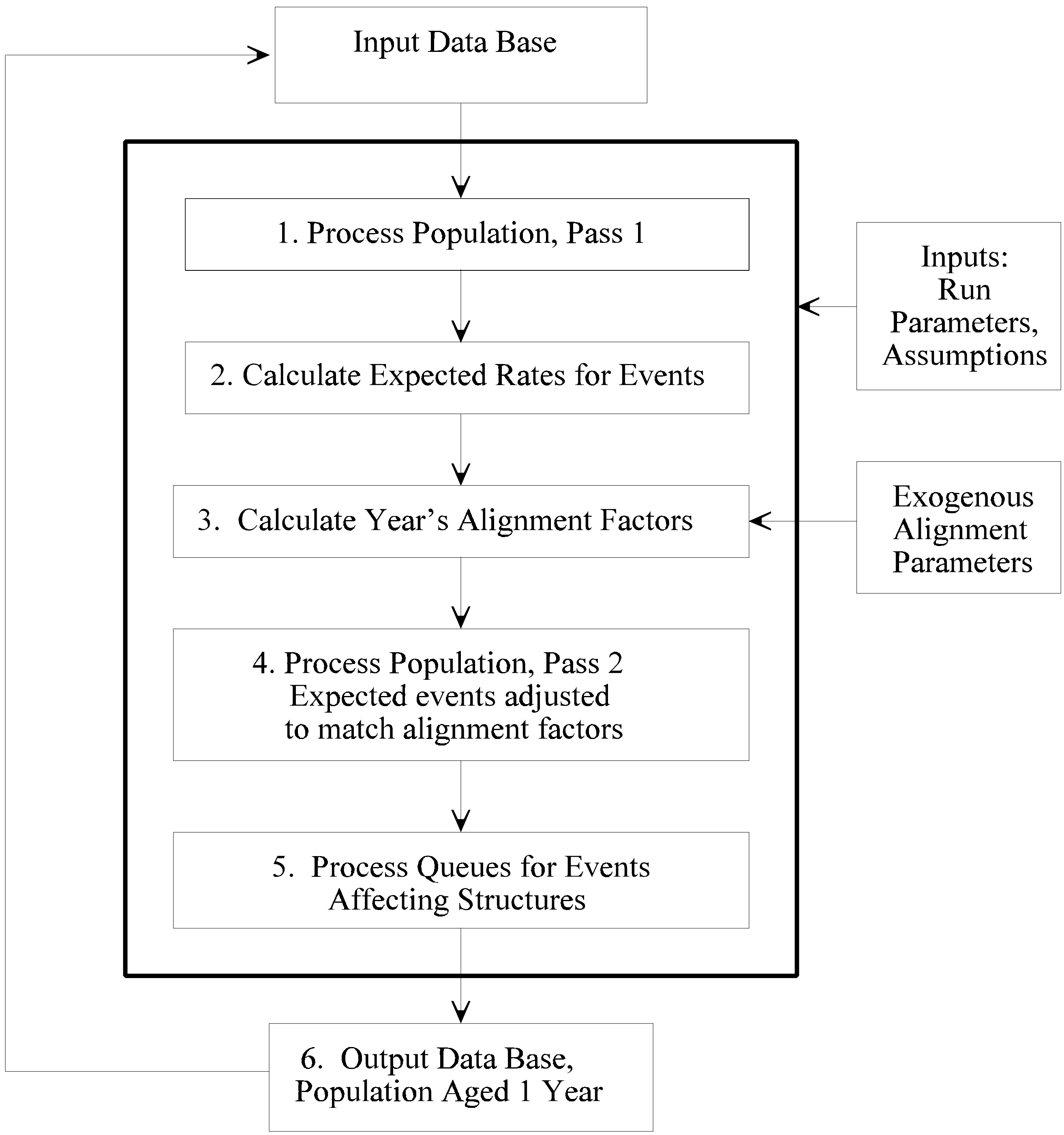


TABLE 6-1

DETERMINANTS OF MAJOR EVENTS SIMULATED BY DYNACAN

Event or Characteristic	Variables Used to Determine Event[a]
DYNACAN-B	
(1) *Family demographic events*[b]	
Death	Age, sex, marital status, labor force participation, employment, earnings, family income, education, disability status, region (Quebec, RoC), place of birth (Canada, other), year (DYNACAN will not permit both spouses in a family to die in same year.) Equations originally are from CORSIM, adapted to Canadian data.
Custody of children at death of parents	Not documented
Birth	Age, marital status, duration of marriage, birth(t-1), birth(t-2), number of children, education, school attendance, labor force participation, employment, living arrangements, family income, earnings, home ownership, work status (full/part-time), region (Quebec, RoC) Equations from CORSIM.
Sex of newborn	Aligned to CPP Actuarial Valuation Model (CPPAVM)
Multiple Births	No
Assignment of "Permanent Luck" to newborn (constant deviation of individual wage from value predicted by wage equation)	Random
First marriage	Age, sex, education, earnings, presence of children
Remarriage	Age, sex, previous marital status (divorced/ widowed), earnings, presence of children, education

TABLE 6-1 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY DYNACAN

Event or Characteristic	Variables Used to Determine Event[a]
Mate matching	Age, sex, education, previous marital status
Leaving family of origin[c]	Age, sex, school attendance (lagged), number of parents, parents' education, employment, earnings (lagged), presence of children, family income, family structure, number of younger siblings (Uses CORSIM data)
Divorce	Duration of marriage; presence, number, and ages of children; wages; does wife have earnings; earnings level; education attainment (aligned to U.S. Vital Statistics data)
Custody of children at divorce of parents	Biological ties, if equal ties random assignment of 90 percent to mother
Emigration from Canada	Rates by age, sex, marital status from 1991 Statistics Canada data; aligned by age, sex, source region (Quebec, RoC) Aligned to CPPAVM.
Immigration (cloning of existing family)	Rates by age, sex, marital status from 1991 Statistics Canada data; aligned by age, sex, destination region (Quebec, RoC) Aligned to CPPAVM.
(2) *Education, location, disability*	
Education (grade completed)	Sex, age, parents' education, living arrangements, parents own home, marital status, has child (Equations from CORSIM.) No alignment.
Interprovincial migration (Quebec, RoC)	Age of family head, number of children, family income, marital status, education, weeks worked, characteristics of family; aligned by age, sex, marital status, region, year. Rates from Statistics Canada 1991 data. Trend to zero in 2010 and after. Aligned to CPPAVM.

TABLE 6-1 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY DYNACAN

Event or Characteristic	Variables Used to Determine Event[a]
Disability	Disability status in previous period, age, sex, year, duration of disability. Rates from CPPAVM. No additional alignment.
(3) *Labor force events and earnings*	
Employment status (worked less than 48 weeks, worked more than 47 weeks)	Age, sex, marital status, presence of children, disability status, retirement status, worked in previous year (full-time, part-time, none)
Exclusion of selected females from labor force	Age (21), sex (female), marital status, presence of children, previous labor force participation
Annual weeks worked (3 classes: 0, 1-47, 48 or more)	Age, sex, marital status, lagged weeks worked class, education, presence of children (women), number of children (women), age of children (women), unemployment rate. Aligned by age, region.
Wage (per week)	Age, sex, marital status (women), presence of children (women), education, unemployment rate, previous year earnings, employment status, lagged employment status, "permanent luck" Aligned to annual earnings by age, sex, marital status, presence of children (women).
Income from employment	Product of wage times weeks worked
Retirement age	Age (60+), disability status, sex current retirement status, year. Rates from CPPAVM.
Labor force participation	Not modeled explicitly. Labor force participation included in employment status.
Hours in the labor force	No
Unemployment	No
Industry	No

TABLE 6-1 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY DYNACAN

Event or Characteristic	Variables Used to Determine Event[a]
(4) *Income from transfer payments*	None modeled other than CPP benefits.
Supplemental Security Income (SSI)	No (Comparable Canadian program (Old Age Security pension, Guaranteed Income Supplement) not modeled.)
Aid to Families with Dependent Children	No (Comparable Canadian program not modeled.)
Earned Income Tax Credit (EITC)	NA
Unemployment Insurance	No (Comparable Canadian program not modeled.)
(5) *Social security eligibility and benefits* (DYNACAN-C)	
CPP coverage	Randomly assigned based on aggregate coverage ratio. Aligned to CPPAVM.
Pensionable earnings	CPP program rules applied to DYNACAN-B longitudinal files
CPP contributions	CPP coverage, earnings, tax rules, year
Retirement Pensions	
Retirement benefit eligibility	Age, covered earnings, retirement status
Retirement benefit receipt/timing	Age, sex, coverage, earnings Aligned to CPPAVM.
Retirement benefit level	Coverage and earnings history of both worker and dependent, retirement timing, year
Disability Pensions	
Disability benefit eligibility	Age, covered earnings, disability status
Disability benefit receipt	Age, sex, coverage, disability status, receipt of disability benefit in previous period, work hours in previous period, change in work hours, earnings, asset income, income, education, home-ownership, live alone, marital status, marital disruption (Take-up rates not modeled)

TABLE 6-1 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY DYNACAN

Event or Characteristic	Variables Used to Determine Event[a]
Disability benefit level	Earnings, coverage and earnings history of both worker and dependent, year
Child of Disabled Contributor (CDC) Benefits	
Benefit eligibility	Age, parents benefit status
Benefit receipt	Program rules (Take-up rates not modeled)
Benefit level	Program rules (flat-rate amount)
Dependents Pensions	
Spouse benefit	Benefit history of wage-earner/ head of family, covered earnings, age, program rules
Children's benefits	Parent's eligibility, age, school status
Survivors Pensions	
Survivors Insurance receipt	Age, coverage, covered earnings of deceased parent or spouse, earnings, timing of parent/spouse's death
Survivors Insurance benefit level	Covered earnings of deceased parent or spouse and dependent, earnings, timing of receipt, year
Orphan's Benefits	
Benefit eligibility	Age, school attendance, mortality of parents
Benefit receipt	
Benefit level	
Death Benefit	
Benefit eligibility	Coverage, mortality
Benefit receipt	
Benefit level	Flat benefit

TABLE 6-1 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY DYNACAN

Event or Characteristic	Variables Used to Determine Event[a]
(6) *Family assets and debts*	Not modeled
(7) *Taxes on income or assets*	Not modeled (except for CPP contributions)
CPP Contributions	Covered earnings, year
Federal income tax	No
Provincial income tax	No
Property taxes	No
(8) *Family consumption and savings*	Not modeled

Footnotes:

[a] Data used for each variable are described in Annex 6-1.

[b] Events/characteristics are grouped into categories by the author. Categories and order of presentation in this table do not indicated logical relationships in the model or solution order. Solution order is shown in Table 6-2

[c] People leaving home for reasons other than marriage, birth of a child, divorce, or death.

Sources: Morrison (1997), *DYNACAN High-Level Technical Documentation* (January 1997), Swan (1997), Wilde (1997)

TABLE 6-2

ORDER OF MODULE SOLUTION IN DYNACAN

Emigration (from Canada)
Pseudo-immigrants (both into Canada and leaving Canada -- reconciliation between Statistics
    Canada's treatments of population stocks and flows)
Internal Migration (between Quebec and the Rest of Canada)
Immigration (to Canada)
Mortality
Fertility
Aging (adding one to the ages of those who have survived)
Marriage (including remarriage)
    Enter marriage pool
    Mate matching
Divorce (including termination of common law unions)
Leaving Home (relevant for family formation)
Education
Disability (including rehabilitation)
Retirement
Selection of certain persons to be permanently excluded from the labor force
Earnings -- major submodules:
    Labor force participation/employment
    Weeks worked
    Wages

Source: Richard Morrison correspondence.

is a set of simulated cross-section databases, one for each year of the simulation period, with demographic and labor market information for each individual in the sample. DYNACAN-B simulates all the years in the projection period before DYNACAN-C is run.

Most discrete events in DYNACAN-B (e.g. death, birth, marriage) are simulated stochastically using a *Monte Carlo* technique.[7] That is, the equation for each event determines a probability that the event will occur for each individual. To determine if the event is assigned to occur, the probability is compared to a random number.[8] Wages, a continuous variable, is also simulated stochastically. The wage equation produces an expected value, and a random term distributes the simulated outcome for each individual in each year around the expected value.[9] This means that different runs of the model, using identical equations and assumptions, will generally produce different sets of outcomes, both for individuals and for distributions and aggregates. Each outcome could be seen as one particular observation from a distribution of possible outcomes. Because of computation costs, DYNACAN, like other microsimulation models, is not usually run numerous times using the same set of input assumptions, to produce such a distribution of outcomes.

Figure 6-2 shows the major steps carried out in the simulation of each year.

(1) **First pass through the data, calculate preliminary event probabilities and outcomes.** DYNACAN-B processes the input database, passing each individual and family through each module. It calculates preliminary probabilities of each event occurring for each individual, but it does not implement any events in this step.

(2) **Calculate expected rates for events.** In the first pass, DYNACAN-B accumulates counts of the expected numbers of events and expected values that the unaligned equations would generate.

(3) **Alignment factor calculation.** DYNACAN-B compares the expected event rates and expected values generated in the first pass to target rates specified by exogenous alignment values. It calculates alignment factors for each equation that are required to ensure that the expected numbers of events and expected values equal the target rates.

(4) **Second pass through the data, assign event outcomes.** DYNACAN processes the input database through the modules a second time. In the second pass the probabilities generated by the module equations are adjusted by the alignment factors calculated in step (3). For each individual and each event a random variable is drawn and compared to the adjusted event

---

[7] Monte Carlo simulation is discussed in Chapter 2, p. 2-5, and in Rubinstein (1981).

[8] In DYNACAN, as in most microsimulation models, "pseudo random numbers" are actually used. Pseudo random numbers are generated by a computer algorithm and are not truly random, but for an individual equation in a given period, they appear to be random. In fact, the stream of pseudo random numbers is determined by the initial "seed" of the random number generator.

[9] Simulation of wages in DYNACAN is discussed in Swan (1997).

probability to determine if the event occurs. Events that do not affect family structures are implemented immediately. Events that affect family structures are put into queues to be implemented in step (5).

(5) **Update family structures.** DYNACAN updates the database to reflect the events that affect linkages between individuals' records, e.g. births, deaths, marriages, divorces, leaving home, or that affect the numbers of data records.

(6) **Output database.** DYNACAN writes out a new database that has been aged one year, reflecting the demographic consequences of the events and states that are simulated to occur during the simulation year. This output serves as the input database for the next year's simulation. When the specified number of simulation years are complete, these output databases, one for each year, are processed by DYNACAN-C.

DYNACAN-C

DYNACAN-C calculates CPP contributions and benefits and provides analysis tables and reports requested by the user. It carries out five processes.

(1) It first processes the information in the output file produced by DYNACAN-B to identify persons required to make CPP contributions and eligible to receive CPP benefits, and indicates the years for which each individual is relevant for either CPP contributions or benefits.

(2) It processes the information in the set of annual cross-section population data files to create for each individual a longitudinal data file, converting the organization of the data from individuals within each simulation year to individual-level information ordered by year for each individual, required to implement the CPP calculations.

(3) DYNACAN-C then reads the person-level longitudinal files created in step (2) and calculates all CPP contributions and benefits for those persons. These calculations are designed to mirror closely the actual program rules used to administer CPP contributions and benefits. The information is ordered chronologically for each individual.

(4) The longitudinal file produced in step (3) is processed to retain records relevant for CPP and produces a data file containing each person's contribution, benefits by type, age, sex and identification number and year.

(5) Finally DYNACAN-C produces tables and reports requested by the user, extracting information from the data files produced in DYNACAN-C and DYNACAN-B, and does comparative policy analysis if appropriate.

Controls and Alignment

Because DYNACAN has been developed to be run in tandem with the CPP Actuarial Model and to produce outcomes consistent with the aggregate outcomes of that model, alignment is a central feature of its operations. DYNACAN has a set of alignment procedures for most modules. As shown in Figure 6-2, alignment of the individual probability or expected value is performed as part of the simulation process itself for most variables. Because DYNACAN begins with a 1971 data file and simulates the period 1971-1997, extensive historical data are available to compare and align the simulated historical data. The role of DYNACAN's projections for future years is to provide demographic and distributional detail for policy analysis of future CPP alternatives, for which the official aggregative projections are provided by the CPP Actuarial Model. Consequently, DYNACAN projections of the future are aligned to projections of the CPP Actuarial Model.

Most module outcomes are aligned in two stages. In the first pass through the data file the initial outcomes or probabilities are estimated, which are used to calculated the alignment factors. Individual micro outcomes of unconstrained equations (usually event probabilities) are aggregated or tabulated for demographic groups for which independent historical data or projections are available. Group alignment factors are calculated by comparing the group expected value or the group mean of the initial simulated outcomes with the historical or projected group data. Individual outcomes or probabilities are then adjusted by the group alignment factor so that the group mean or expected value after alignment corresponds to the historical data or projected target. The adjustments are usually multiplicative. This alignment procedure requires a second pass through the data in the file. In the second pass the initial micro outcomes (probabilities) are adjusted to align to the historic data or projected targets. Because probabilities must be constrained to be between zero and one, the adjustment may take the form of an additive adjustment to the logit of the probabilities, rather than a multiplicative adjustment to the probabilities themselves, which could make some probabilities greater than one.[10] This non-linear alignment procedure may require more than one iteration.

For some variables a second alignment is performed, in which the outcomes aligned to group data are aggregated and compared to historic or projected aggregate data for all groups combined

---

[10] Most probabilities in DYNACAN are estimated using a logistic specification: $p = (1/(1+e^{-Bx})$, where x is a vector of attributes and other explanatory variables and B is a vector of coefficients. The coefficients are estimated by transforming to the logit: $Bx = \ln(p/(1-p))$. The micro equations are then aligned by comparing the logit calculated using the historic group data with the logit calculated using the mean of the initial simulated probabilities. Adding the difference between these two "group-level" logits to the logit equation for each individual will not, in general, produce an exact alignment on the first pass, but ordinarily converges after several iterations. DYNACAN documentation refers to the group mean of the initial simulated probabilities as the *MRATE* and the historical group average rate as *HISTRATE*. The adjustment constant to be added to the Bx is GP = ln(*HISTRATE*/(1-*HISTRATE*)) - ln(*MRATE*/(1-*MRATE*)). See Wilde (1997). Neufeld (1998) has proposed a technique to accelerate this process.

(e.g. national aggregates). The individual outcomes may then be aligned so that in the aggregate they match the historic aggregate data, while preserving relative variation among groups.[11]

These alignment procedures are designed so that the historic simulations track historic data, by aligning to historic group and aggregate data, and future projections of group and aggregate variables match closely CPP Actuarial Model projections of aggregate variables, while the "distribution" of individual outcomes over individuals within each group reflects individual-level variables.

Variance Reduction

As noted, DYNACAN is a stochastic model. The set of outcomes of any simulation is uncertain. Probabilities and expected values can be aligned to match historical or target rates from external data sources, but actual outcomes cannot be guaranteed to match the targets because of the random nature of Monte Carlo simulation. Random variation in simulation outcomes also is problematic for policy analysis. One wants to compare simulations of two policy alternatives to evaluate the effects of the different policy measures. However, differences between two simulations may be due to random variation in outcomes.

If a large number of simulations were to be performed using identical specifications and assumptions (e.g., identical alignment targets), the average values of the outcomes would be expected to approach the unknown expected values. Generally, there are not sufficient resources and time to permit performing the required number of simulations and tabulating and analyzing the results. Consequently, DYNACAN (and other microsimulation models) have developed techniques to reduce the variance of the outcomes of any single simulation.

DYNACAN initially used three variance reduction methods that were used in CORSIM, which served as a template for its development: (1) selective sampling, (2) sidewalk method, (3) common random number stream.[12] To avoid serious drawbacks inherent in the first two of these techniques, the DYNACAN team developed a variance reduction technique referred to as the hybrid method, which is a combination of the sidewalk method and a regular (unconstrained) Monte Carlo technique with pure random (or pseudo random) numbers.[13] The hybrid method uses a table of genuine random numbers (or genuine pseudo random numbers) to determine the occurrence of each event, and keeps a counter of the number of events that have occurred and the number expected to that point, based on the estimated expected values. As each individual is processed, a random number is drawn, and the algorithm compares the number of events that would have occurred using that

---

[11] This "superalignment" in fact is performed as part of the second stage alignment to group controls. See Neufeld (1997).

[12] These methods are described in Chapter 5. Variance reduction in Monte Carlo models is discussed in Pidd (1988), Hillier and Lieberman (1995), and Rubinstein (1981).

[13] The drawbacks of the selective sampling and sidewalk methods, and the advantages of the hybrid method are discussed in Neufeld (1997).

number and the number of events that would be expected. If the two differ by more than a specified threshold, it adjusts the value of the random number slightly to increase the likelihood of the outcome that would bring the number that have occurred closer to the number expected. The two counters retain their values from one year to the next, so that if the number of events simulated in one year differs from the number expected, the algorithm will act to reduce the difference in the following year. In experiments and trial simulations, this hybrid technique has been shown to reduce the variance of outcomes.

Databases

DYNACAN operates on the June 1971 Canadian Census Public Use Microdata Files. The full public use microdata file is a one percent sample of the 1971 Canadian Census and includes about 213,000 person records. Because it is a random sample of the census, each of the person records in the full sample has the same weight, about 100, meaning that each person record represents 100 persons in the Canadian population. Often, DYNACAN is run on smaller subsamples of the full public use file, such as 42,000 (20 percent subsample) or 84,000 (40 percent subsample).

Data from other Canadian and U.S. microdata files were used to estimate the equations of the model, and data from other sources, including survey data and CPP program administrative data, were used to develop rule-based algorithms and to validate the model. The data used to develop and operate the various modules of the model are listed in the Summary Description Table in Annex 6-1. Data bases used include the 1971, 1981, 1991 Canadian census data, the Canadian Survey of Consumer Finances, Canadian and U.S. Vital Statistics data, and CPP program data.

DYNACAN's use of the 1971 Census one percent sample as its base data file is an issue. It is not clear why a data file that is more than two decades old was chosen to serve as the base file, other than to provide an opportunity to validate the model by permitting DYNACAN to simulate the evolution of the population over 25 years for which historic data are available and compare its simulations to the actual data.[14] DYNACAN developers argue that, because the model is closely aligned to the CPP Actuarial Model, which is based on recent data, its projections of group and national rates and aggregates are determined by that model, and would be unaffected by the choice of the microdata base for the simulation. However, two issues may be raised by use of the old database. (1) Even if many group and aggregate outcomes can be aligned to recent data exactly, there is no way to assure that the joint distributions based on the 1971 data remain accurate after 25 years. (2) DYNACAN developers, as well as other microsimulation modelers, have discovered that simulations of individual earnings fail to reproduce the patterns of cross section (annual) and longitudinal (lifetime) variation that characterize CPP administrative data. This is a severe drawback for a model with the purpose of analyzing the distributional effects of social security policies, where benefit levels are based on lifetime earnings patterns. Efforts to address this problem have generated

[14] Note that DYNACAN has been developed primarily to simulate future CPP policy, rather than primary for research.

useful research and imaginative techniques. However, DYNACAN reports indicate that a fully satisfactory solution has not been attained.[15]

The CPP has extensive earnings data for participants. These data have been analyzed in the attempts to develop a methodology to produce satisfactory annual and lifetime earnings simulations and to evaluate simulations which have been produced. If it is feasible and permissible, it would seem that a superior database for simulating lifetime earnings and retirement income policies could be developed by linking CPP administrative data on earnings and other characteristics, such as employment and disability, to records of individuals in a recent census sample or cross section survey.[16]

Documentation

The DYNACAN team has produced considerable documentation of the model and papers discussing its features, methodologies, and problems. This documentation is extensive, clear, and informative. Both detailed and summary descriptions are available. Because the model is still under development, and the developers also face demands both to validate the model and to begin to use it for policy analysis, the existing documentation is incomplete. In particular, there is no written, publicly available description of the CPP model (DYNACAN-C). In addition, the existing documentation provides almost no empirical estimation results, so it is impossible for the reader to evaluate the validity of the model equations. (Some of the equations were originally taken from the CORSIM model.)

Computer Hardware, Software and Portability

The primary DYNACAN computing environment consists, in 1998, of a dual-processor 200 Mhz Pentium Pro desktop computer running parallelized code, 384 megabytes of RAM, and a four gigabyte, fast hard disk, Jaz drive, and writeable CD-ROM drive to meet the very heavy computational and input/output (I/O) demands associated with dynamic microsimulation and large data bases. Most of the model's components are coded in standard C language, and run under the Linux and Windows NT operating systems. A wide variety of other software is used in support of the model and for the post-processing of model results, including SAS, SPSS, SQL Server, Excel, Quattro Pro, Perl, Cold Fusion, TPL, PFE and Brief (editors), LaTex, and other Linux support utilities. A sophisticated user-friendly interface has not been developed. DYNACAN is highly modular and designed so that existing modules can be modified and new modules can be added relatively easily.

---

[15] This issue is discussed in Swan (1997).

[16] This was the approached used in the development of DYNASIM and PRISM. See Chapters 3 and 4. DYNASIM used an exact match of 1972 Social Security earnings records (1937-1972 data) with the March 1973 Current Population Survey. PRISM used a match of 1937-1977 Social Security earnings data with the March 1978 CPS. Unfortunately, current U.S. regulations prohibit public use of Social Security earnings data, so development of future databases in this way for the U.S. is precluded for non-U.S. government institutions. There are no such restrictions on internal Canadian government use of CPP individual earnings records..

The costs per computer work station range from $15,000 (Canadian dollars) to $30,000, with specific costs depending on the speed of processor, amounts of RAM and hard disk storage, presence of writeable CD-ROM drive, nature of backup device, and array of support software for the machine.

The DYNACAN Model has not been transported to any other location, nor are there any plans at present to do so. The computing environment was chosen to facilitate transportability. Installation of DYNACAN on multiple additional machines within OSFI has proven easy.

IV. APPLICATION TO RETIREMENT POLICY ISSUES

DYNACAN was developed to analyze CPP policy issues, particularly distributional effects of changes in contribution rates and benefit formulae. DYNACAN models individual person and family demographic and earnings behavior. It can be used to simulate individual work histories, CPP contributions, benefits receipt and amounts. It can analyze some policy issues for which information on personal histories is required. Examples of issues suitable for analysis with DYNACAN include CPP policy issues such as earnings sharing, changes in benefit formula and indexing, effects of changes in marital histories on the distribution of CPP benefits.[17]

The current version of DYNACAN does not model employer pensions, so it cannot address policy issues pertaining to private or government employee pensions. It currently does not model savings or wealth, so it cannot model individual retirement saving. It does not model income other than earnings, so it cannot address means testing, or broad income distributional issues.

DYNACAN does not simulate employer or plan sponsor behavior, so it does not have the capability to analyze the effects of policy measures on employers or on their offering of pensions or other benefit plans. It does not depict the behavior of producers or industries, so it cannot be used to analyze industry outcomes. Other than very simple labor input variables, individual weeks worked and wages, DYNACAN does not depict any aggregate economic behavior, so it cannot analyze effects on the aggregate economy, such as saving, investment, GDP growth, or interest rates.

V. ACCESSIBILITY AND EASE OF USE

DYNACAN is owned and maintained by the federal government of Canada. It is designed specifically to analyze CPP policy and, perhaps eventually, other income maintenance programs of Canada. It ultimately could be used to do analysis for other Canadian government agencies, but there are no plans to make the model itself available to users outside the government of Canada. Contracting with HRDC to conduct specific studies may be possible after the model development phase is completed.

---

[17] Because DYNACAN is designed to model the Canadian population and Canadian policy issues, no policy matrix for DYNACAN is provided, nor is there a discussion of its suitability for analysis of the policy benchmarks described in Appendix C.

VI. STAFFING AND COSTS

DYNACAN's current core staff consists of eight persons on a full-time or nearly full-time basis: project manager, administrative assistant, senior model analyst, two model analysts, two systems analysts, and liaison to HRDC and other users.[18] Other individuals, including the Chief Actuary and outside consultants, contribute smaller amounts of time.

The initial plan for DYNACAN anticipated a development phase of 5-6 years, with a projected total development budget of about $6.3 million 1992 Canadian dollars. This budget forecast was based on past modeling experience in HRDC and OSFI and consultation with other model developers in government, academia, and the private sector in Australia, Canada, the Netherlands, Sweden, the U.K., and the U.S. Through the end of fiscal year 1997/1998 (March 31, 1998), covering a period of about three and three-quarters years, estimated expenditures were about $2.8 million current Canadian dollars, shown in Table 6-3. This includes the cost of acquisition of the CORSIM model, development of the Canadian database, development of the modules described in Table 6-1, and development of capabilities to simulate CPP contributions and benefits at the micro level that are consistent with those of the CPP Actuarial Valuation Model for both historical and projection periods. Remaining work to complete the original plan includes improvement of the longitudinal individual earnings histories, modeling take-up rates for CPP benefits, and development of new modules to simulate employer pensions, tax-sheltered retirement savings (Registered Retirement Savings Plans), other financial assets, selected real assets (housing), income-tested benefits for the elderly, and income taxes. DYNACAN staff estimate that these remaining tasks may cost an additional million dollars over the remainder of the planned development period (1998/1999 and 1999/2000 fiscal years). This total, about $3.8 million, would leave the project substantially under budget for its complete development.

Part of the total development costs of DYNACAN may reflect research and implementation costs that would not have to be paid by another developer of a similar microsimulation model, who could make use of techniques and methodologies developed by DYNACAN, CORSIM, and other microsimulation models. For example, the DYNACAN team chose to purchase the right to use the CORSIM code, minimizing unnecessary development costs by building on the CORSIM experience. DYNACAN staff estimate that such costs may represent one-half million to one million dollars of the DYNACAN development expenditures to date.

Future applications and operations costs will arise from requirements to develop new modules and modify existing modules to address particular issues and proposed policy options, maintenance of the model, adaptation to use new data as they become available, adaptation to changing computer platforms, ongoing validation against new administrative and survey data, improvements to existing modules, and extensions to new areas. DYNACAN staff estimate that annual expenditures for applications, maintenance, and extensions will be approximately equal, on average, to the annual costs that have been incurred for the initial development phase, i.e. in the range of $600-750 thousand

[18] This staff configuration may change after DYNACAN is transferred from OSFI to HRDC, scheduled for October 1998.

Canadian dollars a year (which would represent about 0.2-0.3 percent of annual CPP administrative costs).

TABLE 6-3

DYNACAN DEVELOPMENT COSTS

Expense Category	Budget 1994-1999 (1992 Canadian $)	Actual Costs Total 1994/95-1997/98 (sum of current Can $)[1]
Salaries and Benefits	2,431,000	1,392,162
Contracts, including EDP	1,900,000	619,191
Travel	96,000	42,940
EDP (hardware, software, data & communication)	683,000	260,462
Other	457,000	20,921
Overhead[2]	729,000	417,649
Totals[2]	**6,296,000**	**2,755,319**

[1] During the period 1994-1997 the exchange rate varied between 1.35 and 1.42 Canadian dollars per U.S. dollar. Simply converting DYNACAN costs from Canadian dollars to U.S. dollars may not provide an accurate cost estimate.

[2] Estimated overhead charges (for accommodation, utilities, etc.) calculated at the rate of 30% of salaries, do not appear directly in the DYNACAN accounts.

VII. CRITIQUE

DYNACAN is a large scale microsimulation model which simulates several demographic and labor market events relevant to the CPP, and calculates CPP contributions and benefits. DYNACAN includes a comprehensive CPP module which depicts in detail the provisions of the system and has extensive report and tabulation capabilities.

Very few of the demographic or labor market behaviors depicted in DYNACAN explicitly represent the effects of economic variables, such as prices, wages, incomes, or assets, or expectations about future values of these variables. DYNACAN has no behavioral retirement function, in the sense of depicting responses to economic incentives.

For modeling broader retirement policy issues beyond the CPP, DYNACAN currently has several limitations. The only two aspects of labor market behavior that are depicted explicitly are weeks worked and wages, and these are modeled primarily with reference to demographic attributes and previous labor market status (see Table 6-1). There is no depiction of the labor force participation decision, *per se*, or of unemployment or hours worked. There is no identification of industry affiliation or size of firm of workers, two attributes that are important for pension analysis. Employment class (self-employed), occupation, and labor union affiliation are not depicted.

DYNACAN's input database does not include individual historical earnings records. Rather, the model imputes individual earnings for the period 1966-1969, based on reported 1970 earnings, and simulates earnings after 1971. This is an area where individual earnings history data may be available, and where accurate data have proven to be both important and difficult to simulate.

In the current version, there is no feedback from changes in CPP policy or provisions or other aspects of retirement income or retirement saving to affect labor market behavior or retirement. Demographic and employment and earnings histories are simulated for the entire simulation period in one model (DYNACAN-B), and CPP provisions and operations are simulated in a separate model (DYNACAN-C), which operates on the output database produced by DYNACAN-B. There is no feedback from changes in the CPP depicted in DYNACAN-C to any demographic or economic behavior depicted in DYNACAN-B.

In this respect DYNACAN is similar to DYNASIM and PRISM. All three models create a detailed simulated longitudinal database with the demographic and economic information needed to simulate social security policy (and employer pensions in DYNASIM and PRISM), then use that simulated database as input into a separate pension model, treating the longitudinal demographic and labor market data as exogenous to pension policy. None of the three can depict simultaneous feedbacks of social security or retirement income policy on labor market behavior. Such feedbacks can be simulated "manually" by the model user modifying input variables or parameters and rerunning the demographic and labor market model (e.g. DYNACAN-B). Changes in social security and pension arrangements may affect labor market, saving, and demographic behavior. Those effects are not well understood, so it is not clear whether they could be represented accurately in a simulation model.[19] CORSIM simulates social security contributions and benefit receipt and demographic and labor market behavior in a single integrated model, so, in concept, it has the capability of building in feedbacks of social security changes on economic and demographic behavior. Such feedbacks are not depicted in the current version of CORSIM.

---

[19] Changes in social security and pension policy may affect current labor market and saving behavior by changing workers' expectations about future income and wealth. The affects of outcomes in future periods, or expectations about such future outcomes, on current economic and demographic behavior may be very difficult to represent endogenously in a microsimulation model.

A crucial aspect of DYNACAN is its reliance on the 1971 census sample for its basic data file. DYNACAN simulates a lengthy historical period (1971-1997) to bring its population to the present. This provides opportunities to investigate the accuracy of various modules, theories, and behavioral representations, and to improve the model's alignment to historical data. However, the ability of the model, even with extensive alignment, to provide simulated historical data and earnings histories that are accurate and capture the many changes in family structure, socioeconomic patterns, and joint distributions that have occurred over the past 27 years is a concern. The use of simulated earnings histories for the 1971-1997 period may be a particular concern for retirement income policy analysis. Despite alignment of simulations to historic group averages, DYNACAN developers report that simulation of individual earnings histories has not yet satisfactorily reproduced observed variation in cross section and lifetime earnings histories data over the 1971-1996 period. This attests to the inherent difficulties of microsimulation modeling. It also reinforces the concern about using simulated historic earnings data as the base for projections of the future and analysis of CPP policy.

Of the microsimulation models reviewed in this report, DYNACAN and CORSIM have the most recently developed code and simulation system. DYNACAN does not have a private pension module, although one is planned. DYNACAN does not provide important information relevant to broad retirement policy, such as worker's industry, occupation, hours worked, firm size, and labor union affiliation. DYNACAN does not model saving or wealth, although these are also planned for future development.

VIII. RELATIONSHIP BETWEEN DYNACAN AND CORSIM

For the Society of Actuaries, one significance of DYNACAN lies in its relationship with CORSIM. DYNACAN represents a successful and fruitful adaption of CORSIM to serve as the basis for the development of an independent microsimulation model. Not only did the DYNACAN development team use the existing CORSIM simulation system and software, taking advantage of CORSIM advances in simulation algorithms, alignment, and variance reduction, and demonstrating CORSIM's portability and adaptability to use with a completely different database and ultimately different operating modules. The two teams have developed a fruitful ongoing relationship, from which both have benefitted. The two teams meet regularly and share information on problems, solutions, and development of new techniques.

ANNEX 6-1

DYNACAN

SUMMARY DESCRIPTION TABLE

DYNACAN
Summary Description

DYNACAN is a Canadian, Monte Carlo, longitudinal, dynamic microsimulation model developed for generating longitudinal and cross-sectional, policy-oriented analyses of the Canada Pension Plan (CPP).

- DYNACAN simulates Canadian demographics and earnings, focussing on the Canada Pension Plan. It is being developed by the Office of the Superintendent of Financial Institutions of the Canadian federal government.
- Its fundamental unit of analysis is the individual. Individuals are treated in the context of the families in which they reside. Collectively and distributionally, the individuals are representative of the total Canadian population in each year of the projection period.
- DYNACAN is longitudinal in the sense of projecting the individuals forward through time, synthesizing life histories as regards the events relevant for CPP contributions and benefits, with the characteristics of the representative individuals changing over the course of the simulation.
- It is dynamic in the sense of simulating individual behaviors and events year by year, based on current and past values for states, behaviors, and events, generating an evolving population, whose size and membership changes over time.
- It is a Monte Carlo model in the sense of using "random" numbers in simulating the events associated with that evolution.
- DYNACAN presently simulates demography, earnings, and CPP contributions and benefits. Plans call for it to be extended so that policy analyses can address the CPP in the context of the broader retirement income system.

Subject: Size, distribution, structure, and socioeconomic conditions of the population of Canada; socioeconomic conditions and behavior of individual persons and families with respect to events and conditions of significance to the Canada Pension Plan specifically, and eventually the retirement income system more generally, with a major emphasis on CPP contributions and benefits.

Purpose and Objective of Model: Analysis of hypothetical changes to the Canada Pension Plan, including estimation of longitudinal, distributional impacts, assessments of winners and losers and their characteristics, eventually in the context of the broader retirement income system, including interactions with taxes and income-tested benefits.

Period of historical analysis: 1971-1996.

Forecast/simulation horizon: Historical simulation: 1971-1996 (1996 being the last year for which authoritative CPP contribution data are available); forecast simulation: 1997-2100, though more commonly 1997-2030. The historical period is simulated both to provide the necessary earnings and event histories on which future projections depend, and to permit the validation of the model's algorithms and parameters.

Frequency: Annual

Base year: 1971

Simulation technique: Monte Carlo dynamic microsimulation with alignment to exogenous assumptions and variance reduction.

Solution algorithms and structure: Sequential/recursive solution of single equations, with Monte Carlo assignment of events outcomes and states. Each year all individuals and families pass through the several modules. The population is processed by one module before processing starts for the next module. Most modules utilize a two-pass algorithm. The first pass gathers the information necessary for alignment to external controls. The second pass generates and implements the events.

Unit(s) of analysis: Individuals within the context of families, with explicit attention to family formation and dissolution. All migration is simulated on a family basis.

Cell Structure: DYNACAN is not a "cell-based" model, in the sense of depicting the behavior of aggregates or group averages. Its analytic units (cells) are (initially) the individual person and family records from 1971 Canadian Census, weighted (equally) to represent the Canadian population. Raw data are not aggregated or summarized before model operations. Module-specific sets of cells or bins are used, year by year, to align event probabilities to their counterparts in the cell-based CPP Actuarial Valuation Model.

<u>**Databases**</u>:

Base Year Database: 1971 Canadian Census Public Use Microdata file 1 in 100 sample, and 20% and 40% subsamples.

Population/demographics: 1971, 1981, 1991 Canadian Census microdata file from the full census tabulations, supplemented by large sample versions of the Canadian Survey of Consumer Finances, Labour Market Activities Survey, CPP administrative data, data from CPP Actuarial Valuation Model, Canadian vital statistics, and other data. In some instances where Canadian microdata are not available, U.S. data have been used, aligned to Canadian totals.

Individual/family/household characteristics: Canadian Census data, Survey of Consumer Finances, Vital Statistics, U.S. data.

Employer characteristics: Employers not explicitly simulated.

Industry characteristics: Industries not simulated**.**

Retirement plan coverage, participation: Canada Pension Plan coverage simulated as universal for those persons with sufficient earnings to make contributions, and not living in the province of Québec. Private sector plans are not presently modeled, but planned for future extensions.

Retirement plan vesting: Immediate vesting and full portability for the CPP, including explicit provision for movement between the Québec Pension relevant for residents of the province of Québec and the CPP relevant elsewhere in Canada. Private sector plans not currently modeled.

Retirement plan characteristics: CPP contribution rules, benefit rules, and administrative provisions are simulated based on program rules and data, designed to mirror the legislation, guidelines and institutional practices.

Individual Retirement Account (IRA) participation, contributions: Not presently modeled. Development of a model of the Canadian general equivalent to IRAs, Registered Retirement Savings Plans (RRSPs), is planned for future, with some work underway.

Supplemental Security Income (SSI) eligibility, participation: Not modeled. Implementation of the closest Canadian income-tested benefit equivalents to U.S. SSI (Old Age Security, the Guaranteed Income Supplement, and the Spouses Allowance) is planned after the other income sources required for their computation are modeled.

Family assets: Not included. Planned eventually to be included in a wealth module as part of an effort to model the CPP in the context of the broader retirement income system. Background work underway, but no attempts at implementation.

Home ownership: Not included. Planned eventually. A small amount of background work has been done for inclusion in the planned wealth module, but there has been no attempt to implement it.

Macroeconomic data: Provided exogenously in the form of parametric assumptions about the general economy, e.g. inflation and interest rates.

Labor market data: Exogenous assumptions about labor force participation, wage growth, and the conditional distributions of earnings, generally taken from the CPP Actuarial Valuation Model (ACTUCAN) and Labour Market Activities Survey.

Retirement behavior: Specified as exogenous assumptions, based on tabulations of CPP program administrative data, with worker retirement decision a function of age, gender, and disability status.

Taxes: Not included (other than CPP contributions). Planned for future versions. CPP contributions are simulated. Unemployment Insurance premiums could be approximated. Income sources other than earnings and CPP benefits are not included.

Health conditions, disability: No simulation of health conditions. Disability status, including rehabilitation and the increased mortality associated with the disabled population, is simulated using probabilities drawn from CPP administrative data.

Health insurance coverage: Not simulated, in part because of Canada's universal health care coverage, as administered and largely funded by the provinces.

Institutional population characteristics: Not included as a separate or identifiable population.

<u>Data Quality</u>: Generally good. 1971 Census sample includes about 213,000 individuals with data on several attributes relevant for simulation of CPP contributions and benefits, including relationships to others in the same family unit. The data have undergone validity checks by Statistics Canada, as well as DYNACAN's adjustments to ensure that the data base's distributions reflect full Census distributions throughout the historical period.

Completeness: Base year data base contains original or synthesized data sufficient for simulation of CPP contributions and benefits. Data may not include several elements relevant for simulation of employer pensions, e.g. industry, occupation.

Accuracy: DYNACAN staff believe that with the combination of Statistics Canada's care in collection and in validity and consistency checks, plus their own validation relating to distributions and averages, the data are sufficiently accurate for the intended purposes.

Representative: The use of the 1971 Canadian Census one percent sample as the starting data base, and alignment during the simulations (to actual distributions during the historical period and to the CPP Actuarial Valuation Model projections during the projection period), ensures that the data base is, throughout the simulation period, sufficiently representative. Substantial use of administrative program data and with survey data over the historical period, including its use to set alignment targets, ensures that the data remain representative.

Currency: The base year data for 1971 are 27 years old. The data are aged year by year to simulate events aligned to reflect known historical rates. The evolving data base generally reproduces known stock totals and distributions over the historical period. Considerable validation is undertaken. DYNACAN staff believe that the evolving data base offers, for all years in the historical period, a reasonable picture of persons and families relevant for the CPP. However, it is not possible to assure that important covariances that prevail in 1998 are represented in the aged 1971 data, even with extensive alignment. Additional validation is planned to ensure that important covariances, e.g. relations between spouses' ages or earnings, that prevail during the historical period are reproduced in the simulated historical data base.

Applicability to other contexts: With proposed future extensions, DYNACAN will be applicable to several contexts other than the CPP. Health and housing are being considered.

Gaps: Incomes other than earnings and CPP benefits, private pension coverage and entitlements, and asset stocks that will be necessary to simulate income taxes and income-tested benefit programs. DYNACAN is being developed in stages. The filling of these gaps is planned for future stages.

Applicability of other private/consulting firm data: Given the intended purposes, there is relatively little need for data from the private sector. Data on private pension plan participation and entitlements, and on registered retirement savings plans would be the most useful, but there is no comprehensive private source for such data, and the degree of representativeness of firm-specific data, even if they could be obtained, would be suspect.

Characteristics, activities, behaviors that are modeled:

Demographic characteristics: All characteristics and activities relevant to CPP benefits and contributions are modeled. They include: births, gender, age, deaths, marriage and remarriage, divorce, marital status and status duration, schooling and educational attainment, immigration, emigration, interprovincial migration, Quebec versus Rest of Canada residence, leaving home, entering the labor force, retirement, disability and recovery/rehabilitation, disability status and status duration.

Economic activity:

Short-run/cyclical: Exogenous.

Long-run growth, productivity: Exogenous

Inflation: Exogenous

Industrial sector detail: None.

Open or closed economy: There is no simulation of the economy beyond employment of individuals, aligned to ACTUCAN's assumptions relating to the overall economy. Later stages of model development should add the aggregates and distributions for income taxes and savings.

Labor market behavior: Labor-force entry, lifetime exclusion from the labor force (for a small proportion of individuals), annual labor force participation, weeks worked during the year, full-time/part-time employment, wage level for the portion of the year worked, and retirement age. Plans to add more detail involving transitions from one employer to another and explicit consideration of paid versus self-employment.

Capital markets: None. After implementation of proposed wealth module, capital markets outcomes will be the result of exogenous assumptions about savings rates and savings channels, rather than any simulated market clearing mechanism.

Retirement program characteristics: Individual's eligibility for various CPP benefits, updating eligibility and benefit determinations each simulated year, indexation of benefits in pay. No private plan characteristics.

Retirement behavior: The model draws on year-specific administrative data relating to retirement benefit applications to simulate retirement for CPP purposes, conditional on year, gender, and disability status.

Savings and asset accumulation: None at present. These are included in the planned wealth module intended to mirror the wealth module in CORSIM.

Government behavior:

Federal budget: None at present. Plans to simulate income taxes.

Social Security and Health Insurance Trust Funds: All CPP contributions and benefits (retirement benefits, survivors benefits, disability benefits, death benefits, orphans benefits, benefits to children of disabled contributors, including the impact of limits on combined benefits) are modeled. Benefit-specific take-up will be added in the near future. The CPP fund itself is not modeled in DYNACAN but is in the CPP Actuarial Valuation Model, ACTUCAN. Canada does not have health insurance trust funds to be modeled..

Regulations: Only the regulations pertaining to CPP benefits are modeled.

Taxes: None at present; a tax module is planned.

Public retirement income programs:

Social Security Retirement, Disability and Survivors Insurance: All CPP contributions and benefits (retirement benefits, survivors benefits, disability benefits, death benefits, orphans benefits, benefits to children of disabled contributors, including the impacts of limits on combined benefits) are modeled.

Means tested old age or disability income transfers: No means-tested programs are modeled. None of the CPP benefits, including the disability benefits, is means tested, except indirectly via inclusion of the full amount of the benefits in taxable income for income tax purposes. Applicants for retirement benefits, if they are less than 65 years of age, must withdraw from the labor force at the time they apply for benefits, but can later re-enter it with no reduction in CPP retirement benefits. In practice, disability beneficiaries with any significant continuing level of earnings (more than about $6,000 U.S. per year) would likely be deemed to have recovered, and their benefits cease.

Government employee pension programs: DYNACAN does not at present include any pension plans other than the mandatory Canada Pension Plan. Inclusion of employer plans is part of DYNACAN's planned staged development. A fraction of such employer pensions will reflect the characteristics of government employee pension programs.

Federal civil service: In future extension of model may be treated as a special case of "private" pension plans, one of which will have the characteristics of this type of plan.

Military: In future may be treated as a special case of "private" pension plans, one of which will have the characteristics of this type of plan.

Provincial and local government, types: May be treated as a special case of "private" pension plans, one of which will have the characteristics of this type of plan.

Private pensions: None at present. Inclusion of such plans is part of planned extensions. Preliminary work is underway.

Defined benefit: Planned model of private pensions will include present defined benefits plans and their characteristics.

Defined contribution: Planned model of private pensions will include present defined contribution plans and their characteristics.

Supplemental: Planned model of private pensions may include supplemental plans and their characteristics, to the degree they exist in Canada.

Individual retirement saving arrangements: DYNACAN does not model individual level tax-sheltered (registered) savings arrangements. These are to be included in the planned wealth module.

Public sector health care finance programs: DYNACAN does not model either the premiums or the benefits from public sector health care finance programs. In Canada, with its largely universal health care system, modeling benefits as financial flows to individuals is not an issue. Modeling contributions, e.g. Ontario's Fair Share Health Care Levy, would be done in conjunction with modeling provincial income taxes. There are at present no plans to model demands for health services; this could be a possible extension of DYNACAN.

Medicare: Not Applicable (NA)

Medicaid: NA

Military: NA

Veterans: NA

Indian Health Service and others: NA

Private sector health care finance programs: Not modeled. (Except for a fairly narrow range of services deemed non-essential, provincially run health care systems provide most traditional health care, and private competition is not permitted. Some employers provide supplemental health care, e.g. dental plans, eyeglasses or semi-private accommodation if the employee is hospitalized. Modeling such relatively small programs is not a priority in DYNACAN.)

Private health insurance, especially retiree health insurance: Not modeled. (Beyond the overall government provision of health care, there are, in many provinces, separate benefits paid out of public monies targeted to seniors, e.g. payment for a variety of prescription drugs. Modeling such programs is not currently a priority for DYNACAN.)

Employer/plan sponsor behavior: DYNACAN does not include employers and thus does not model their behavior, beyond accounting for both the employee and employer portions of CPP contributions.

Worker behavior: Labor force participation, weeks worked, earnings, retirement decision. No behavioral equations reflecting responses to economic variables. The equations for these variables include characteristics such as education, marital status, presence of children, and thus do respond to the occurrence of the associated events in the individuals' lives.

Health care provider behavior: Not included.

Insurer behavior: Not included.

Institutionalization: Not included.

Assumptions, Parameters, Methodology

Key Assumptions: To align aggregate expected values and rates, DYNACAN uses numerous assumptions, e.g. average wage growth, inflation, age group specific birth rates, mortality rates, disability and recovery from disability rates, entry into the labor force, labor force participation rates, distributions of earnings by year, age and gender, gender-specific distributions of ages for retirement, immigration, emigration, and interprovincial migration, educational attainments, marriage, remarriage and divorce rates, etc. Virtually all of these are specified parametrically and can be varied. Due to the importance of ensuring consistency with the CPP Actuarial Valuation Model (ACTUCAN), DYNACAN typically uses the same values assumed in that model, assumptions that are typically heavily informed by recent Canadian experience. Rate of return on investment is not an explicit variable or parameter presently. There is an implicit assumption via CPP contribution rates that have been calculated in ACTUCAN assuming a particular rate of return on monies in the CPP fund.

Types of Parameters, Decrements, Transition Rates/Probabilities: Demographic parameters – fertility, mortality, immigration, emigration, inter-provincial migration rates; marriage and divorce rates; labor force participation rates; retirement and disability incidence rates. At the beginning of model development approximations and simplifications have been used that are expected to be improved over time.

> **Experience considered, origins of decrements**: CPP administrative program data, Census and survey data, and vital statistics. Mortality tables and projections are closely related to the Canada Life Tables, following the principles espoused by the Canadian Institute of Actuaries, consistent with the procedures espoused by the Society of Actuaries. (Fuller detail is available in the CPP statutory actuarial reports published regularly by the Chief Actuary.)
>
> **Consistency with other experience and other assumptions of model**: Model is consistent with program data and consistent with the CPP Actuarial Valuation Model. Many of the parameters have been to some degree validated by their consistency with their counterparts in the U.S. CORSIM model. No parameters are known to violate assumptions of the model.

Internal consistency: Model is internally consistent. DYNACAN developers attempt to adopt assumptions that are judgementally consistent (e.g. for the relative rates of inflation, wage growth and interest rates). The considerable attention given throughout the model to respecting the rates and distributions reflected in administrative and historical economic data, vital statistics, and census and survey data, ensures that parameters will be generally consistent internally.

Methodology used to estimate parameters and relationships:

Econometric/statistical: Many of the equations that generate DYNACAN's events are estimated statistically, using a variety of statistical/econometric techniques, including ordinary least squares regression and maximum likelihood, especially for probabilities. Many probability equations are estimated using a logistic specification The outputs from these equations are usually aligned to respect assumptions, as described above. For the most part, the equations supply the dispersion, while the assumptions control the overall levels or rates.

Actuarial: DYNACAN is designed so that all of its important alignment parameters parallel those of the CPP Actuarial Valuation Model (ACTUCAN) and have, as far as possible, the same interpretation. Runs are typically made using the same values of these parameters as are used in ACTUCAN runs.

Judgmental: Some of the alignment values are at least partially judgemental; particular examples include the overall birth rates anticipated for the future, and the future spread between wage and price growth.

Economic/actuarial literature, studies done by others, etc.: Virtually all of the assumptions are viewed in the light of existing literature and studies by others, for example, the Canadian Institute of Actuaries' work on the impacts of AIDS on the distribution of mortality rates. DYNACAN cooperates closely with the CORSIM modeling team, and it uses several of their equations (with coefficient adjustments to ensure that the results align to the Canadian experience) or their forms (with re-estimation using equivalent Canadian data sources), and in some cases their specific parameter values when no reliable Canadian counterparts are available.

Simulation Methodology: Monte Carlo dynamic microsimulation

Stochastic Properties: DYNACAN is a stochastic model. It determines events and state changes using Monte Carlo simulation. An event probability is compared with a random number to determine if the event occurs. The results of any given simulation depend in part on the seed for the pseudo-random number generators it uses. DYNACAN employs a number of variance reduction techniques to reduce the random variation from one run to the another. (It shares these techniques with the U.S. CORSIM model.)

Feedback Phenomena: No simultaneous feedbacks. No feedbacks of pension system provisions on labor market, savings, or demographic behavior. Since DYNACAN is not a market-clearing type of model, and does not have modules for sectors such as capital and labor markets, it does not provide for feedback among them. When the savings and investment and taxation modules are developed, the design will have to address the types of feedback to be included (e.g. the extent to which tax changes will affect labor force participation and savings behavior, and how this will in turn affect the timing of future tax yields).

Microsimulation adjustment ("aging") methodology: Dynamic aging of individuals and families of the population. Individuals, families and their characteristics evolve period by period. Specific identifiable individuals age and die during course of the simulation, and others come into existence via birth or immigration.

Policy levers: CPP contributions and benefits rules, drop-out rules. DYNACAN's purpose is the projection of the current CPP system and impact analyses of CPP policy options. If a tax system module is added to the model, tax policy levers could be included.

Economic/demographic feedbacks:

Employer costs and behavior: None. DYNACAN does not model employers. Labor force participation rates and earnings distributions are given exogenously.

Labor market behavior: Limited. Labor market behavior is influenced by demographic variables, especially marital status. Labor market status and earnings of individuals affect some demographic behavior in the following period. As is the case elsewhere in DYNACAN, however, this affects primarily the dispersion of outcomes over individuals, since the overall levels of wages, wage growth, participation, etc., are governed by assumptions and controlled by DYNACAN's alignment facilities. Pension changes do not affect labor market behavior except to the extent that the effects are exogenously provided for in the alignment/assumption parameters..

Taxes, government deficits, etc: None. DYNACAN does not, at present, have a tax module. The initial version of the planned tax module is likely to be strictly computational, and would involve no impacts on labor force participation, earnings or savings behavior.

Capital accumulation: None. DYNACAN does not, at present, have a wealth, or wealth accumulation module. The initial version of the planned wealth module is likely to be strictly deterministic based on past behavior, and would involve no impacts on labor force participation, earnings or savings behavior.

Interest rates: Interest rates are not included in DYNACAN. In the CPP Actuarial Valuation Model interest rates are specified exogenously as assumptions. If a wealth and wealth accumulation module is added to DYNACAN, returns to all types of investments (stocks, bonds, interest bearing accounts) will have to be specified explicitly within DYNACAN. Initially, the rates, and their evolution over time, would be specified exogenously as assumptions.

Employment, productivity, economic activity, GDP: None. Economic activity and GDP are not modeled in DYNACAN. (Aggregate employment is determined by population distributions and assumptions about labor-force participation and employment rates by age and gender.) Employment-related variables (e.g., the time paths of labor force participation for the various cohorts, rates of growth in wages) are specified exogenously. Currently, DYNACAN uses the same aggregate assumptions as the CPP Actuarial Valuation Model, but the model could in principle use different assumptions.

Sensitivity Analysis: The parametric specification of most of DYNACAN's assumptions permits the convenient generation of sensitivity analyses. When DYNACAN development is further advanced, its sensitivity analyses will be compared against those regularly performed in the CPP Actuarial Valuation Model.

Model Validation Procedures: DYNACAN undertakes validation on several levels: (1) the model's demographic stocks and flows; (2) labor force participation, employment and earnings, including both cross-sectional and longitudinal distributions; and (3) CPP contributors, contributions, beneficiaries and benefits. These are validated against the Canadian Census, survey data, and CPP administrative program data over the historical period, as well as against the CPP Actuarial Valuation Model's outputs over both the historic and projection periods. (The CPP Actuarial Valuation Model is itself extensively validated against CPP program data and a variety of Census and survey data.) In addition, there is a separate set of "collateral validation" standards having to do primarily with the extent to which the family types in DYNACAN, and the characteristics of these family groups, reflect distributions seen in Census and survey data. The collateral validation is seen as particularly important because family characteristics play a major role in assessing projected policy impacts, and because they are not directly modeled or controlled in the model, depending as they do on the interaction of many modules.

<u>Computer implementation</u>:

Hardware requirements: The primary DYNACAN computing environment consists, in mid-1998, of two dual-processor 200 Mhz Pentium Pro desktop computers running parallelized code under the Linux operating system (but with only one of the machines used for any given simulation), each with 512 megabytes of RAM, and large (4GB), fast hard disks, Jaz drives, and a shared writeable CD-ROM drive to meet the very heavy computational and input/output (I/O) demands associated with dynamic microsimulation and large data bases. These machines are networked with other Linux machines, e.g. one operating as a network server, and several Windows NT machines.

Software: Most of the model's components are coded in ANSI standard C language and run under the Linux operating system. Windows NT 4.0 is used as the operating system for the other (non-Linux) components of the support network. A wide variety of other software is used in support of the model, and for the post-processing of model results, including SAS, SPSS, SQL Server, Excel, Quattro Pro, Perl, Cold Fusion, TPL, PFE and Brief (editors), Interbase, LaTex, Applix, and other Linux support utilities.

Computer costs: At summer 1998 prices, the computer configuration costs about $21,000 Cdn (about $14,000 US) per work station, with specific costs depending on the speed of processor and number of processors, amounts of RAM and hard disk storage, presence of peripherals such as writeable CD-ROM drives, nature of backup device, and the array of support software chosen for the machine. Ongoing improvements in computer technology and price reductions have substantially reduced and can be expected to continue to reduce the cost of suitable hardware/software.

Transportability: Model has not yet been transported to any other location. One of the design criteria was to permit satellite user installations, and the computing environment was explicitly chosen to facilitate transportability for this purpose. In July 1998 the project was informed that it must be able to transfer the model to Strategic Policy in Human Resources Development Canada, from its developmental home in the Office of the Chief Actuary in the Office of the Superintendent of Financial Institutions. Installation of DYNACAN on multiple additional machines within OSFI has proven easy.

Applications (Projects, Studies) DYNACAN has been used on a trial basis to analyze the impacts of proposed changes to the Canada Pension Plan. Applications are expected to begin in earnest in late 1998 upon the finalization of cross-sectional earnings alignment and the implementation of benefit take-up algorithms, and its further validation against the CPP Actuarial Validation Model, ACTUCAN. Procedures for validating the impacts of individual options against ACTUCAN are still under development.

Contact Person: Richard J. Morrison
Phone: (613) 990-3955,
Fax:: (613) 990-9900,
E-mail: dynacan@istar.ca