CHAPTER 8

MINT

MODELING INCOME IN THE NEAR TERM: THE U.S. SOCIAL SECURITY ADMINISTRATION NEAR TERM RETIREMENT INCOME MODEL

MINT

MODELING INCOME IN THE NEAR TERM

TABLE OF CONTENTS

CHAPTER 8

MINT

MODELING INCOME IN THE NEAR TERM: THE U.S. SOCIAL SECURITY ADMINISTRATION NEAR TERM RETIREMENT INCOME MODEL

I. OVERVIEW

The MINT (Modeling Income in the Near Term) model is one of three retirement income and social security models now being developed in the Office of Research, Evaluation, and Statistics (ORES) of the Social Security Administration.[1] MINT is essentially a static microsimulation model, but it represents an important innovation in microdata modeling and simulation.

The purpose of the MINT project is to develop a large, representative microdata file with information useful for analysis of social security policy proposals. MINT projects the income in retirement of social security retirement beneficiaries who were born from 1931 through 1960. The members of these cohorts reach age 62 in 1993 through 2022 and include the cohorts born during the relatively low fertility years of the 1930s and World War II, as well as the first three-quarters of the baby boom. MINT focuses on income at the age of retirement and in the year 2020. To project retirement income from social security and other sources, MINT projects lifetime earnings of each sample person. To obtain a complete picture of the distribution of income of social security beneficiaries, MINT also projects income from other sources, including pensions, non-pension wealth, and partial retirement earnings.

While MINT produces a simulated microdatabase with projected future incomes, it is not a true dynamic microsimulation model. MINT uses statistically estimated regression equations to project future demographic states of the individuals in a sample of the population and to estimate incomes from various sources and assets at specific retirement ages and years for individuals in later cohorts in the sample, based on the lifetime path of incomes and assets of individuals in earlier cohorts in the sample and other microdatabases. The projections assume that the future patterns of income and assets of younger individuals will replicate the past patterns of income and assets of similar individuals in earlier cohorts.

---

[1] The others are POLISIM, which is based on the CORSIM model described in Chapter 5, being developed in the Division of Policy Evaluation (DPE) and the Office of the Chief Actuary (OCACT), and the Cohorts Model, described in Chapter 12, being developed in the Division of Economic Research (DER).

II. BACKGROUND

The predecessor to MINT was developed in the mid-1990s by analysts in the Social Security Administration Division of Policy Evaluation (DPE), who did several studies of social security policy using a database developed by matching records from the Survey of Income and Program Participation (SIPP) panels of 1984 and 1990 with Social Security Summary Earnings Records (SER).[2] In early 1998, SSA contracted with the Urban Institute and the Brookings Institution and with the Rand Corporation to do extensive data development and estimation to significantly expand and enhance the SSA/DPE version of MINT. The initial contract work was delivered in September 1999. Under a second contract with The Urban Institute and Brookings, revisions and additional enhancements are being done. A significantly revised version of MINT is scheduled for completion in June 2001.[3]

III. DESCRIPTION

MINT is essentially a large, comprehensive data file with detailed information about the socioeconomic characteristics of a large, representative sample of the U.S. civilian, non-institution population in 1990-1995, their earnings taxable for social security from 1951 through 1996, and their projected demographic characteristics, earnings, wealth, asset income, pension coverage and pension income, and social security benefits from 1996 through 2020. MINT focuses on projecting the incomes of the individuals in the data file at age 62 and 67 and in the year 2020, although it can project incomes in any year from 1996 through 2020.[4]

---

[2] See Iams and Sandell (1997), Sandell and Iams (1997), and Sandell and Iams (1998). The Survey of Income and Program Participation and the Social Security Summary Earnings Records are described in Appendix B. The Panel on Retirement Income Modeling of the National Research Council in a comprehensive study of retirement modeling needs and capabilities noted the 1996-1997 SSA modeling work and concluded that "this approach could offer some useful policy modeling capability without requiring a full-fledged dynamic microsimulation model." (Citro and Hanushek, 1997, p. 161).

[3] The model developed under the first Urban Institute/Brookings and Rand contracts, delivered in September 1999, is refered to by SSA as MINT I. A revised version of MINT I, referred to by the Urban Institute as MINT II, was completed in March 2001. A significantly enhanced version of MINT scheduled for completion in June 2001 is referred to as MINT III. Documentation of MINT II and MINT III was not publically available in May 2001. This chapter describes primarily MINT I but describes modifications being implemented in 2001 where information was available in May 2001.

[4] Retirement in MINT I is defined as take-up of social security retirement benefits. The definition of retirement age in the different modules of MINT I is ambiguous. The documentation from the Urban Institute states: "The purpose of the MINT project is to estimate the baseline distribution of income of the population of Social Security retirement beneficiaries from the 1931-
(footnote continued on next page)

The MINT project comprises three major analytical steps:

1. Development of a large, comprehensive microdatabase with actual and imputed historic information on demographic and socioeconomic characteristics, family structure, earnings histories, pension coverage, wealth, and social security benefit receipt;

2. Estimation of equations which predict the future demographic characteristics and incomes from four sources of the individuals in the database who were born in 1931-1960, including mortality, marriage/remarriage, divorce, widowhood, disability, earnings, assets, asset income, pension coverage and accruals, pension benefits, retirement decision, and earnings after retirement;

3. Projection of the demographic and socioeconomic characteristics of the individuals in the database that are needed to project all the major components of their retirement incomes and to estimate the effects of alternative social security policy proposals.

Data File

The basic data file of MINT is composed of four panels of the Census Bureau Survey of Income and Program Participation (SIPP) of 1990, 1991, 1992, and 1993, matched with the Social Security Administration Summary Earnings Records (SER) for 1951 through 1996 and the Social Security Master Beneficiary Records (MBR). The SIPP core survey data files provide extensive information about the demographic and socioeconomic characteristics and incomes of persons in the sample age 15 and older on a monthly basis, including age, sex, race, ethnicity, education, marital status, housing tenure, income from various sources, labor market activity, and disability condition. The 1990 SIPP covers 1990 through mid-1992; the 1991 SIPP covers 1991 through mid-1993; the 1992 SIPP covers 1992 through 1994; and the 1993 SIPP covers 1993 through 1995. The data file also includes information from SIPP Topical Modules on marital history, assets and liabilities, and retirement expectations and pension coverage. The records of the persons in the SIPP samples have been matched by the Census Bureau with the Social Security Administration Summary Earnings Records for those persons, which have data on their annual earnings taxable for social security from 1951 through 1996. These records have also been matched with the Social Security Master Beneficiary Records, which have information on the types, timing and amounts of social security benefits received. The records were also matched with social security Numident files, which provide dates of death of members of the sample through mid-1998.

In the four SIPP panels there are 113,071 records for persons born between 1926 and 1965. Of these, 78,161 had positive panel weights, indicating that they participated in all

---

1960 birth cohorts at the age of retirement (either 62 or 67) and in the year 2020." (Toder *et al.,* p.1). The MINT pension module and assets module estimate pension benefits and wealth assuming that retirement is either at age 62 or age 67. The social security retirement benefit receipt module permits the age of first benefit receipt to be any age from 62 through 67.

interviews. This sample was used for estimation of several of the projection equations. There are 84,497 records for persons born in 1931-1960. These records comprise the sample that is projected. Of these, 59,537 have positive full panel weights.

Projection Components[5]

MINT projects three sets of demographic events – mortality, marital status changes, and disability– and four sources of retirement income – social security retirement or survivor benefits, defined benefit pension benefits, income from assets, and earnings. In order to project retirement income, MINT projects annual lifetime earnings, defined benefit and defined contribution pension accruals, wealth accumulation at retirement age, and age at take-up of social security benefits (referred to as retirement age). MINT includes a tax module that estimates federal income, FICA, and state and local income taxes. The MINT project also included a task that developed a set of stylized earnings patterns that characterized the earnings data and projected those patterns. The stylized earnings patterns have not been used for policy projections.

Table 8-1 shows the major components that are projected by MINT and the sequence in which they are projected. Table 8-2 shows the determinants of the major events simulated by MINT.[6]

[5] MINT is described in detail in Panis and Lillard (1999), Toder, *et al.* (1999), and Butrica, *et al.* (1999).

[6] To facilitate comparison with other models, Table 8-2 groups and numbers the events simulated by MINT somewhat differently than Table 8-1 and lists some events not simulated in the current version of MINT.

TABLE 8-1

MAJOR COMPONENTS OF MINT

1. Demographic projections
    - A. Mortality
    - B. Marriage/remarriage
        - 1. Spousal characteristics
        - 2. Imputed spouses' earnings
    - C. Divorce
    - D. Disability Condition
2. Earnings before retirement
    - A. Individuals in the SIPP/SER
    - B. Receipt of social security disability insurance benefit
    - C. Individuals not in the SIPP
3. Pension coverage, accruals, and benefits
    - A. Defined benefit
    - B. Defined contribution
    - C. Future plan participation
4. Wealth
    - A. Housing
    - B. Non-housing
5. Timing of receipt of social security retirement benefits
6. Earnings after retirement ("partial retirement earnings")
7. Income in retirement from total non-housing assets
8. Projections of total income
    - A. Social security
    - B. Defined benefit pensions
    - C. Financial assets
    - D. Imputed rental income
    - E. Earnings
9. Income taxes
10. Stylized earnings patterns

TABLE 8-2

DETERMINANTS OF MAJOR EVENTS SIMULATED BY MINT

Event or Characteristic[a]	Variables Used to Determine Event[b]
(1) *Demographic Events*	
Death	Sex (separate equations for male, female), age, race, marital status, education, permanent income, year (time trend)
Marriage/Remarriage	Sex (separate equations for male, female), age, duration unmarried, number of previous marriages, race, Hispanic ethnicity, education, previous marital status (widowed), permanent income, year (time trend)
Spousal characteristics imputed	Birth date, race, Hispanic ethnicity, education, disability status, date of disability onset
Mate matching (future spouses and divorced ex-spouses)	Difference in age, marriage start date, marriage end date, marriage termination status, disability status, disability date, death date, race match, Hispanic ethnicity, difference in education, difference in permanent income
Divorce	Sex (separate equations for male, female), age, marriage duration, number of marriages, race, Hispanic ethnicity, education, year (time trend)
(2) *Disability*	
Onset of disability condition	Age, sex, race, Hispanic ethnicity, education
Onset of Social Security Disability Insurance benefit receipt	Sex (separate equations), age group, race, Hispanic ethnicity, education, disability condition, average earnings in previous 10 years
(3) *Labor Market Activity and Earnings before Retirement*	
Lifetime age-earnings profiles (earnings relative to average earnings)	Sex (separate equations), education (separate equations), age (13 age group categorical variables, ages 24-67)

TABLE 8-2 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY MINT

Event or Characteristic[a]	Variables Used to Determine Event[b]
Labor force participation	not explicitly determined in MINT
Employment	not explicitly determined
Unemployment	not explicitly determined
Industry	not determined
Annual weeks worked	not determined
Hours in the labor force	not determined
Wage	not determined
(4) *Retirement Income*	
(A) *Social Security Retirement Benefits*	
OASDI eligibility	Quarters of social security coverage – own and spouse's
Retirement benefit receipt/timing	Sex and marital status (separate equations for married male, married female, unmarried persons), age, race-Hispanic ethnicity, education, earnings, pension coverage, spouse age, spouse education, spouse earnings, spouse pension coverage, widowed, non-housing wealth, home-ownership, home value
Retirement benefit level	AIME and PIA calculated from worker's and spouses' earnings histories (actual and projected), current earnings, earnings test rules
Spouse benefit	AIME and PIA calculated from worker's and spouses' earnings histories, current earnings, earnings test rules
Survivors Insurance (SI) receipt (widow/er only)	Age, eligibility rules, AIME and PIA of deceased spouse/ex-spouse
Survivors Insurance benefit level	PIA of deceased spouse/ex-spouse, dependents, earnings, earnings test rules

TABLE 8-2 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY MINT

Event or Characteristic[a]	Variables Used to Determine Event[b]
Disability (DI) benefit eligibility Disability benefit receipt Disability benefit level	DI benefit receipt not simulated. Onset of DI benefit receipt is simulated to terminate earnings
Children's benefits	Not included
(B) *Pensions*	
Defined benefit pension eligibility	Coverage reported in base data file (1990-1993 SIPP) or assigned to non-participants based on age, gender, earnings, using SIPP coverage rates
Defined benefit pension benefit	Years of service, final salary, age at retirement, occupation, sector of employment (private, federal government, state and local government, military), job change frequency assigned based on years of service until retirement, expect benefits from prior job, COLAs based on sector of employment
Defined contribution account balances (401(k), non-401(k), IRA, Keogh)	Balance reported in base data file (1990-1991 SIPP), allocation assumptions, employee contribution reported in base data file, employer match rates assigned based on employee contribution rate, assumed (stochastic) rates of return and administrative costs, future participation assigned to non-participants based on age, gender, earnings, using SIPP coverage rates
Defined contribution benefit payments	Annuitized balances at retirement, based on (1) unisex mortality, (2) mortality by gender, birth year, race, education

TABLE 8-2 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY MINT

Event or Characteristic[a]	Variables Used to Determine Event[b]
(C) *Non-pension Wealth before Retirement*	
Housing wealth	
Home ownership	Ownership in base data file (1990-1993 SIPP), age of head, race, marital status, two-earner couple, average earnings, family size
Value of home equity (conditional on ownership)	Age of head, age squared, sex of head, race, marital status, family size, age youngest child, health of head, average earnings, current earnings less average earnings, two-earner couple, fraction wealth held in equities, receive pension income, birth cohort
Non-housing wealth	
Presence of positive other wealth	Age of head, race, marital status, two-earner couple, average earnings, family size
Value of other wealth (conditional on wealth being positive)	Age of head, age squared, sex of head, race, marital status, family size, age youngest child, health of head, average earnings, current earnings less average earnings, two-earner couple, fraction wealth held in equities, homeownership, receive pension income, birth cohort
Asset transfers	
Divorce	Rule-based: split equally
Bequests and inheritances	Rule-based: to spouse (other heirs/dependents not modeled)
(D) *Wealth Change Post-retirement*	
Non-housing wealth	Marital status (separate equations for married, single); age interaction with: homeowner, race, pension receipt, high earnings, widowed (different age slope terms); sex, race, ethnicity, divorced/widowed, education, age at initial social security entitlement, cohort, social

TABLE 8-2 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY MINT

Event or Characteristic[a]	Variables Used to Determine Event[b]
	security benefit, housing wealth, DB pension benefit, earnings age 50-60, death within 27 months
Non-housing asset income	Annuity payment recalculated each year, based on (1) unisex mortality, (2) mortality by gender, birth year, race, education, applied to 80 percent of non-housing wealth
Housing wealth	Assumed constant real
Housing asset income	Assumed 3 percent real rate of return on housing wealth
(E) *Earnings after Receipt of Social Security Retirement Benefit (Partial Retirement Earnings)*	
Probability of being in one of four earnings groups (zero earnings, three positive earnings groups)	Age groups (three separate equations for two age groups), age, race (non-Hispanic white), marital status and sex (separate equations for male/female, married/unmarried for age 65-68), married female, married female/spouse age, education, wealth, pension benefits, pension/married female, retirement wealth, past earnings, social security exempt amount (for 1984 sample)
Earnings (conditional on earnings group)	Drawn randomly from SIPP earnings distributions (Equations estimated using social security exempt amount, age, average past earnings, Mills ratio)
(F) *Supplemental Security Income (SSI)*	Not included in existing model. Under development for MINT III.
(G) *Total Family Income*	Sum of income of unmarried person or married couple from social security, earnings, DB pensions, non-housing and housing wealth

TABLE 8-2 (continued)

DETERMINANTS OF MAJOR EVENTS SIMULATED BY MINT

Event or Characteristic[a]	Variables Used to Determine Event[b]
(5) *Taxes on Income or Assets*	
FICA	Earnings
Federal income tax	Family earnings, family asset income, pension income, social security benefits, 1998 tax rules
State and local taxes	Family federal income tax
Property	Not modeled explicitly; included in state and local taxes
(6) *Family Consumption and Savings*	
Consumption expenditures	Not included
Savings	Not included (wealth assigned directly)

[a] Events/characteristics are grouped into categories by the author. Categories and order of presentation in this table do not indicate logical relationships in the model or solution order.
[b] Data used for each variable are described in Annex 8-1.

Demographic Projections

The first step in developing the MINT projected data file is to project the sample of the birth cohorts of 1931-1960 that is alive at retirement ages and in 2020 and their marital statuses and marital histories. Because social security age benefits depend not only on a worker's earnings history, but can depend on the past earnings of her/his current or former spouse(s), it is necessary to know the marital status and past marriages and marriage durations of each individual. The lifetime earnings projections and retirement income projections algorithms are applied to a demographic file which contains the projected date of death, dates of marriages and divorces, characteristics of spouses not observed in the database, and periods of disability of each person in the base year projection sample (the 1931-1960 birth cohorts in the 1990-1993 SIPP panels). This demographic projection file was developed by the Rand Corporation in the first task of the MINT project.

To create the demographic projection file, five transition events are estimated for each individual:

- Marriage and remarriage
- Divorce
- Death of spouse
- Death of the individual
- Onset of disability

The timing of the sequence of the first four events determines marriage histories. Marital status and mortality are projected jointly.

Mortality

Death of the individual and deaths of spouses are modeled using the mortality model. Mortality is modeled using a continuous time proportional hazard model. The logarithm of the hazard of dying at time t is specified to be a piecewise linear function of age, with different slopes for different age groups, a calendar time trend, marital status and socioeconomic characteristics (race, education, permanent income) (see table 8-2).[7] [8] The equation is estimated using data from the Panel Study of Income Dynamics (PSID) for 1968-1994 for individuals age 30 and older. The

---

[7] The hazard of dying is very similar to what is referred to in the actuarial literature as the *force of mortality*, usually written $\mu(x)$, or, more completely, the select force of mortality, $\mu_x(t)$. It is the (unobserved) instantaneous age-specific mortality rate – the conditional probability density function of dying at exact age x, given survival to that age (or the conditional probability of dying at age x+t, given selection at age x). (See Bowers, *et al.*, p. 55, 79, 83; Ross, p. 240.) The probability that an event has not happened yet as of time t is given by the survival function (which equals one minus the cumulative probability function of occurrence). The hazard rate (force of mortality) equals the proportional rate of decline of the survival function. Hazard rates and hazard models are discussed in Allison.

[8] The mortality model is a piece-wise linear Gompertz model. See Bowers, *et al.*, pp. 77-78; Spiegelman, pp. 132, 163-165, 168, 169; Shryock and Siegel, Vol 2, pp. 382, 690, 691.

estimates based on the PSID are adjusted to correspond to the U.S. population of 1901-1994 using the United States Vital Statistics data. Mortality hazard models were estimated using the Vital Statistics data and the PSID data using sex, age, year, and race, and the coefficients for those variables of the full model were adjusted to correspond to the estimates based on the Vital Statistics.[9] This mortality model is used to estimate the probability of death for individuals in the baseline SIPP sample and their spouses each year and to simulate the year of death.[10]

Marriage and Remarriage

Marriage and remarriage are modeled using a continuous time hazard model – an equation that specifies that the logarithm of the hazard of getting married (remarried) at time t for a particular individual is a piecewise linear function of age, duration unmarried, time, number of previous marriages, race, ethnicity, education, and permanent income. The model was estimated on data from the 1990 and 1991 SIPP panels, including Marital History modules, and the 1992 and 1993 panels were used to assess the goodness of fit.

Divorce

Divorce is modeled using a continuous time hazard model with a specification similar to that for marriage. The logarithm of the hazard of divorcing for a particular individual is a piecewise linear function of age (two age slopes), marriage duration (five duration slopes), time trends (two time slopes), number of previous marriages, education, race, ethnicity. The model was estimated on data from the 1990 and 1991 SIPP panels, including Marital History modules, and the 1992 and 1993 panels were used to assess the goodness of fit.

Onset of Disability Condition

This equation models self-reported functional work disability. The hazard of onset of disability for a particular individual is specified to be a function of age, sex, education, race, ethnicity. The disability model was estimated using the 1990 and 1991 SIPP Work Disability Topical Modules for persons age 30 and older. Disability is modeled as an absorbing condition, i.e., once an individual is projected to be disabled, he/she remains disabled for the remainder of the simulation period.

---

[9] As a consequence, the MINT mortality time trends are based on the time trends prevailing over the longer period 1901-1994, rather than the more recent period reflected in the PSID data, and differences by sex and race reflect differences over the longer period rather than the more recent period .

[10] The mortality and disability condition algorithms were modified significantly in MINT III. MINT III imputes death and disability through age 67 by matching each person to one or more older persons in the SIPP/SER database with similar characteristics, in sets of five-year intervals. This matching technique, referred to as hot-decking, is used to project earnings after 1998 (see footnote 19). If the older (donor) individual whose record was matched to the younger (recipient) individual dies or becomes disabled in the interval, the recipient individual is imputed to die or become disabled. The mortality and disability rates through age 67 are adjusted to match projections of the SSA Office of the Chief Actuary. Mortality after age 67 is projected using the procedures described for MINT I.

Simulation and Projection

The interrelated demographic transitions of marriage, divorce, widowhood, and death, and the resulting marital histories are projected using the estimated hazard functions. At the beginning of the projection process each individual in the base data file has a set of characteristics, including age, marital status, duration of that status, number of previous marriages and divorces, etc. Depending on his/her demographic state, he or she is subject to two or more transition hazards. For example, an unmarried person may become (1) married or (2) deceased. A married person may become (1) divorced, (2) widowed, or (3) deceased. The next state is determined by whichever transition takes place first. The model generates durations until all possible events. The various demographic states affect each others' transitions hazards and, therefore, the generated durations.

From the estimated hazard function of an event, the survival function can be calculated. The survival function gives the probability that an event has not happened yet at time t (or duration t).[11] From the survival function, the distribution of potential durations until an event and the expected duration of a state for an individual can be calculated (e.g., expected remaining years of life, length of marriage). To predict the time of a transition, the model draws randomly from the distribution of durations. This is accomplished by generating a random number between 0 and 1, and solving for the duration corresponding to that number (probability). For each potential transition, the model draws a duration. The shortest duration determines which transition occurs. After that transition has occurred, the individual becomes subject to a new set of competing hazards and has different demographic characteristics. The model then draws a new set of durations corresponding to the new set of hazards and characteristics. The shortest duration again determines which event occurs next. The process continues until the individual becomes deceased[12].

The date of onset of disability is projected in the same way. In the current version of MINT I disability is not affected by marital status and does not affect the hazard of mortality. Consequently, disability is projected independently of the other demographic characteristics and does not affect them.[13]

---

[11] See footnote 17.

[12] For example, the projected duration until marriage of a never married man may be five years, while the projected duration until death may be 40 years. The algorithm decides that the man will marry first. He then becomes subject to the competing hazards of divorce and widowhood, as well as death. Because mortality is a function of marital status, the man now faces different survival probabilities. A new set of durations is drawn – until divorce, spouse's date of death, and his date of death – and the shortest one determines the next transition.

[13] Rand estimated and reported a mortality hazard equation that includes disability, which indicated that disability is a significant predictor of mortality. The mortality equation which takes account of disability could be incorporated into MINT. (See Panis and Lillard, p. 46 and Appendix A.7.) MINT III uses disability status to predict mortality.

This simulation algorithm differs from an alternative approach, which is used in dynamic microsimulation models, such as DYNASIM, PRISM, and CORSIM, which are described in earlier chapters. These models project individuals' life paths in discrete periods, such as months or years. For example, the probability that an unmarried man marries during the next year is estimated as a function of his characteristics. If this probability exceeds a randomly drawn number from a uniform distribution between 0 and 1, he is projected to become married in that year. Otherwise, he remains single. Similarly, the probability of death is estimated, and whether death occurs is projected. The model selects the dominant transition(s), generally by the order in which the algorithms are solved. The probability of each event may be influenced by various socioeconomic characteristics and states. Consequently, the occurrence of events simulated earlier in a period may influence the probability of other events simulated later in the same period. Later events in the solution sequence can influence earlier events in the sequence only in the following period. The model projects all potential states and events that it encompasses for each individual for each period, before beginning the simulation of the next period. This process is repeated for each subsequent period until the death of the individual and is repeated for the sample of survivors until the end of the simulation period. This alternative approach may permit simulation of more interactions among covariates that vary frequently in time. Note that this approach requires assumptions about which transitions are dominant.

Characteristics of Former and Future Spouses

Calculation of social security retirement benefits and projection of widowhood and divorce require information about characteristics of spouses, including spouses from marriages that ended before the 1990-1993 SIPP data were collected and from marriages projected to occur in future years. If those spouses were not present in the base SIPP data file, their characteristics must be imputed. Characteristics of unobserved former spouses and future spouses were imputed based on empirical distributions of characteristics of couples in the SIPP data. These characteristics include sex, date of birth, race, Hispanic ethnicity, education, disability status, and date of onset of disability. Characteristics of unobserved spouses are recorded as variables on the SIPP individual's record. Unobserved spouses (i.e. spouses who were not in the base data file) do not have their own records.

Simulated Former and Future Spouses

After the specified characteristics of unobserved spouses were imputed, the records of actual spouses in the SIPP data file were matched to individuals with unobserved spouses, based on the imputed characteristics of the unobserved spouses and the closeness of the actual persons to the imputed characteristics.[14] Closeness was determined in terms of minimizing a distance function – a weighted sum of squares of the normalized difference between the imputed characteristic and the actual characteristic of the person in the SIPP database. The characteristics measures in the distance function were spouse's birth date, marriage start date, marriage end date, marriage termination status (divorce, widow, death), disability status, disability date, race, Hispanic ethnicity, education, and permanent income. Weights were assigned based on analysts'

[14] This statistical matching technique of imputing proxy individuals or characteristics where information is missing for some individuals in a database from other individuals in the same database is referred to as "hot-decking."

judgement of which characteristics were more important. The primary reason for imputing spouses was to provide earnings records of the unobserved spouses, to be used to calculate social security retirement benefits and other elements of retirement income.

Retirement Income

Four sources of retirement income are projected:

- Social security retirement benefits and survivor benefits of elderly widows and widowers;
- Employer pensions;
- Asset income;
- Earnings.

Retirement is defined to mean receipt of social security Old Age benefits. MINT estimates the income from the four sources for individuals and married couples who have begun to receive social security Old Age benefits at age 62, for those who receive social security benefits at age 67, and for the population of beneficiaries in the year 2020. MINT assumes that all who are eligible for social security benefits begin to receive them by age 67. The relatively small proportion of the elderly population that does not receive social security benefits at any age are not included in the MINT universe.[15]

MINT does not project the absolute level of real or nominal incomes. All economic variables in MINT are expressed as percentages of the average wage in the economy. Thus, MINT projects the distribution of incomes and wealth relative to the average wage rather than the path of incomes and levels of wealth. MINT uses the projections of the Social Security Administration Office of the Chief Actuary (OCACT) to derive the future values of the average wage and price level in the economy. By applying the MINT projections of relative incomes and wealth to the OCACT projections of the path of average wages and prices, the user could derive projections of real and nominal income levels, but the focus of MINT is on income distributions.

To provide the basis for the calculation of social security benefits, as well as the accrual of pension benefits and non-pension wealth, MINT projects the lifetime earnings profile of each individual and the year of retirement (defined as take-up of social security retirement benefits).

Earnings Before Retirement

To calculate social security retirement benefits MINT applies the social security benefit formula to the lifetime earnings of the individuals in the projected data file. The first task of the retirement income model is to project lifetime pre-retirement earnings profiles. Lifetime earnings profiles are projected both for individuals in the SIPP/SER sample and for individuals who are not in the SIPP sample.

---

[15] A Supplementary Security Income module is currently under development, which will include some of the low income population that currently does not receive social security.

The matched SIPP/SER data file provides earnings histories for each individual for the period 1951 through 1996. Based on these earnings histories, MINT projects social security covered earnings from 1997 to 2020 and calculates Average Indexed Monthly Earnings (AIME) at the ages of predicted first benefit receipt.[16]

Lifetime age-earnings profiles are projected using an age-earnings equation that specifies that each individual's annual earnings, relative to average earnings, are a step-function of age. A fixed effect earnings equation was estimated separately for men and women in five educational attainment classes. A fixed-effect specification assumes that each person in a sub-population differs from other workers in the group by a fixed average amount. The individual-specific difference persists over a worker's entire career and is captured by an individual-specific constant deviation term in the equation.[17] The equations were estimated using the matched 1990-1993 SIPP/SER data for 1987-1996. The SER records workers' earnings only up to the social security taxable maximum amount, which after 1989 was adjusted each year to be 2.46 times average earnings. For the 1987-1989 data, men's earnings at the taxable wage ceiling were adjusted upward based on CPS data to reflect an estimate of an average wage in excess of the taxable maximum but below 2.46 times average earnings.[18]

---

[16] In MINT III annual earnings from the SER are observed through 1998. Projection of earnings are made from 1999 through date of first social security benefit receipt.

[17] A number of functional specifications are used to analyze hypothetical relationships that may characterize the behavior of individuals in panel data. In general, the specifications can be written as: $y_{it} = \alpha_i + \gamma_t + \sum\beta_k x_{kit} + e_{it}$,

where y_{it} = the value of the dependent variable for individual i in time period t,
α_i = an intercept term that varies over individuals,
γ_t = an intercept term that varies over time periods,
β_k = a set of slope coefficients to be estimated,
x_{kit} = a set of independent (nonstochastic) variables,
e_{it} = a stochastic error term that varies over individuals and over time.

The α_i or the γ_t can be constrained to equal zero, or the intercept can be constrained to be equal for all individuals and time periods. The α_i and/or the γ_t can be specified to be constant (fixed) coefficients or they can be random error components. If the α_i and/or the γ_t are specified to be constant, the model is called a *fixed effect* model, or a dummy variable model, or a covariance model. If the α_i and/or the γ_t are specified to be random individual-specific or period-specific errors, the model is called a *random effect* model, or an error components model. Both fixed effect and random effect specifications are used to estimate the equations of the MINT model. Specification and estimation of these models is discussed in Johnston (1984), pp. 396-407; Kmenta (1986), pp. 625-635; Judge *et al.* (1985), pp. 515-560; and Chamberlain (1984).

[18] All of the SER earnings data of 1951-1989 were adjusted to take account of censuring at the taxable maximum, but only the 1987-89 data were used for estimation.

Earnings are projected for each person in the projection sample (SIPP respondents born 1926-1965) by adding an estimate of the person-specific fixed effect to the average age-education specific earnings predicted by the equation. A time-varying random error term is added to the prediction for each individual each year to generate projections which have a variance similar to actual earnings, based on the errors observed for the individual during the estimation period 1987-1996.[19]

Disability Insurance Benefit Receipt

The onset of a functional disability health condition was projected as part of the demographic projections. A separate model, estimated in conjunction with the generation of lifetime age-earnings profiles, projects the take-up of Disability Insurance (DI) benefits. The probability of first receipt of DI benefits is estimated using a probit specification to be a function of age, race, Hispanic ethnicity, education, average earnings in the previous ten-year period (in six or eight earnings groups), and functional work disability condition. The disability benefit receipt model was estimated using the matched SIPP/SER data files matched with the Master Beneficiary Records for 1987-1994. Onset of DI receipt was projected using actual data from the MBR for 1995-1997 and the predictions of the model for each year 1998-2025. Each year after 1997 for each person, the probability of onset of DI receipt predicted by the equation was compared to a random number between 0 and 1 to determine if the person was assigned to begin DI in that year.

After the onset of DI benefit receipt, projected earnings are set to zero for the remainder of the person's life. After beginning to receive DI benefits, individuals do not return to employment. Earnings are set to zero in the year that an individual is projected to die.

Earnings for Individuals Not in the SIPP

As explained above, in each case where an individual in the SIPP/SER projection database was projected to have a spouse or spouses before or after the data collection period, for whom there is no SIPP or SER record, a record of an eligible person in the SIPP/SER database was assigned to provide information about the unobserved spouse, based on that spouse's imputed characteristics. The historic and projected work history and lifetime earnings on this record provide the earnings of the unobserved spouse. The number of imputed spouses is considerable. The demographic projection file has 113,071 persons. Altogether, there are 149,445 observed or imputed marriages. For about 42 percent of marriages (62,102) the spouse is observed. All other spouses are imputed (87,343).

---

[19] The earnings projection module was modified significantly for MINT III. In MINT III annual earnings are projected by matching each person in the SIPP/SER database who does not have a complete lifetime earnings history to one or more older persons in the database with similar characteristics, and assigning the older person's observed earnings (in five-year age intervals) to the younger person for future years when the younger person reaches the current age of the older person. This imputation technique is often referred to as hot-decking.

Pensions

The MINT pension module projects income from employer defined benefit (DB) pensions and income and balances of defined contribution (DC) pensions, IRAs and Keogh plans at age 62 assuming retirement at age 62 and at age 67 assuming retirement at age 67. Pension income includes income from private, federal employee, military, and state and local government pensions. Pension benefits are projected based on information reported for each individual in the 1990 and 1991 SIPP Topical Modules on Retirement Expectations and Pension Coverage and on Assets and Liabilities. MINT assumes that each worker remains in the same type of pension plan with the same characteristics until retirement. To account for workers who do not have pension coverage in the SIPP but may have future coverage, and persons who may enter the labor force in future years, participation rates are projected based on participation rates in the 1990-1993 SIPP, and future coverage is randomly assigned based on age, gender and income. MINT provides information on type of pension, years of pension participation, value of current defined contribution type accounts, and contribution rates by employee and employer.

Defined Benefit Pensions Information on pension coverage on a SIPP respondent's current jobs is used to project pension benefits. MINT assumes each worker reporting participation in a DB plan in the SIPP continues DB participation until retirement, even if he/she changes jobs. The DB pension benefits at age 62 and age 67 are projected using replacement rates by years of service, final salary, age at retirement, and occupation. Replacement rates for private sector and state/local government pensions are from BLS tabulations. Replacement rates for federal civil service and military pensions are estimated based on plan provisions. All workers covered by a DB plan are assumed to begin receiving a pension at ages 62 or 67, even if their projected earnings fall to zero at younger ages. The model assigns to each worker a number of job changes and the years of service on each job, then reduces the pension benefit according to total years of service, number of job changes, and occupation. MINT estimates survivor benefits assuming that all married workers have a 50 percent joint and survivor benefit. MINT estimates pension benefits expected from prior jobs reported in the SIPP assuming that these are all taken as DB pensions at age 62 (67) and that none were taken as lump-sum distributions.[20] The model applies cost-of-living-adjustments to a portion of future pension benefits and to pension income already being received in the base year data file, to a randomly assigned portion of private and state/local government beneficiaries and to all federal civil service and military beneficiaries.

Defined Contribution Pensions, IRAs, and Keoghs MINT projects account balances for 401(k) plans, non-401(k) defined contribution plans, Keogh plans, and Individual Retirement Accounts. The first three are estimated and reported together as DC plans; IRAs are reported separately. The 1990-1991 SIPP provides account balance information for 401(k)s, IRAs and Keoghs for most respondents with these plans. For non-401(k) defined contribution plans MINT estimates the base period account balance by accumulating assumed contributions over the entire

---

[20] MINT assumes lump-sum distributions from prior pension plans which were rolled over into IRAs are reported as IRAs rather than as pensions expected from prior jobs. MINT developers recognize that this assumption may result in some double counting, which they estimate may be 10 percent, based on 1991 SIPP data.

period of plan participation. Employee contributions for each type of plan are reported in the base period SIPP. The model assumes the percentage of earnings contributed to the plan in the base year data remains constant until retirement and applies this percentage to the earnings projections of each year. MINT randomly assigns employer match rates separately for 401(k)s and for non-401(k) plans based on reported employer contribution rates in the 1995 Survey of Consumer Finances (SCF). MINT specifies that employer match rates decrease with increases in the employee contribution rate. Account balances reported in the SIPP and new contributions are allocated equally between stocks and bonds. MINT assigns a mean rate of return and standard deviation to each type of asset parametrically, then draws the particular rates of return to each asset for each participant in each year stochastically. An administrative cost is subtracted from each rate of return (one percent in the current version). At age 62 (67) DC balances, IRAs and Keoghs are combined with other non-housing wealth for the projection of asset drawdown and asset income.[21] The pension model also reports illustrative retirement income from pensions by calculating annuities from account balances at the assumed retirement ages. Two sets of annuities are reported, based on (1) unisex mortality assumptions based on 1989-1991 Decennial Life Tables from the National Center of Health Statistics, and (2) mortality assumptions used in the mortality model based on PSID data, which vary by gender, birth year, race, and education. The annuities assume that married persons have 50 percent joint and survivor annuities.

Non-Pension Wealth

MINT projects the level of housing wealth and non-housing wealth at age 62 and age 67. MINT does not attempt to model wealth accumulation directly, with reference to consumption, saving, and rates of return on assets. Rather, it uses reduced form equations to estimate wealth of individuals as a function of age, earnings histories, and demographic characteristics.[22] For both housing wealth and non-housing wealth, two-stage models are estimated, using data from the PSID for 1984, 1989, and 1994. The first stage estimates the probability of having positive housing wealth and the probability of having non-housing wealth. The second stage predicts the level of each type of wealth conditional on having positive wealth.

The probability of owning a home was estimated with a random effect probit model using PSID data, as a function of age, marital status, race, whether there were two earners, family size, and average income.[23] Conditional on owning a home, the value of home equity (house value less

---

[21] The MINT documentation indicates that the model does not recognize the different tax treatment of withdrawals from tax-preferred pension and retirement accounts vs. ordinary asset accumulations.

[22] A *reduced form* model or equation describes a relationship between a dependent variable and a set of independent or predetermined variables and other endogenous variables but does not attempt to characterize the structure of the underlying system, or the underlying causal relationships of the system. Reduced form and *structural form* models are discussed in Chapter 1.

[23] A *random effect* specification assumes that there is an individual-specific random error term. See footnote 17 for a description of fixed effect and random effect models.

A *probit* model is a statistical model that permits estimation of an underlying probability of mortgage debt) was estimated with a random effect model using PSID data, as a function of age of head, marital status, gender, race, family size, age of youngest child, health status, average earnings, whether there are two earners, and the fraction of wealth held in equities.[24]

The probability of having positive non-housing wealth[25] was estimated with a random effect probit model using PSID data, as a function of age, marital status, race, whether there were two earners, family size, and average income. Conditional on having positive non-housing wealth, the value of non-housing wealth was estimated with a random effect model using PSID data, as a function of the same explanatory variables included in the housing wealth equation plus housing tenure.

The housing and non-housing wealth equations were used to project the wealth of the individuals in the SIPP sample at age 62 and age 67. If an individual was a homeowner in the 1990-1993 SIPP he/she was assumed to remain a homeowner at age 62 and 67. For individuals who were not homeowners in the base year sample, homeownership status at age 62 was predicted using the homeownership model by comparing the probability estimated by the homeownership equation with a uniformly distributed random number. For individuals who were assigned to be homeowners at age 62, the value of housing wealth was imputed. First, the housing wealth equation was used to predict a value of housing wealth for that individual in the base year SIPP file. An individual-specific "residual" was calculated as the difference between the actual housing wealth reported in the SIPP and the predicted wealth from the estimated equation. This residual was used to center the predicted wealth to match reported wealth on the SIPP. The residual was then added to the value of housing wealth predicted by the equation at age 62 and age 67 to project housing wealth.

Non-housing wealth was projected in the same way. Individuals were predicted to have positive non-housing wealth at age 62 (67) using the equation for the probability of having positive non-housing wealth. The value of non-housing wealth in the base year SIPP data was predicted using the estimated wealth equation, compared to the actual wealth reported, and an individual-specific residual was calculated. That residual was then added to the value of wealth at age 62 (67) estimated by the equation to project non-housing wealth.

---

a state or event for an individual (such as the probability that the individual owns a house), as a function of characteristics of the individual, using observations on the binary (yes/no) choices made by the individuals in the sample (such as whether each individual owns or does not own a house).

[24] The housing wealth and non-housing wealth equations are actually mixed fixed effect/random effect models. Both equations include an individual-specific error term (random effect) and dummy variables for year (fixed effect). See footnote 17.

[25] Non-housing, non-pension wealth is defined as the sum of vehicle equity, equity in other real estate, own business equity, stocks, mutual funds, bonds, checking accounts, savings accounts, money market accounts, and certificates of deposit, less unsecured debt.

Receipt of Social Security Retirement Benefits

The age of initial receipt of social security retirement benefits is projected for individuals age 62 through 67 who are eligible for social security and who have never received social security disability benefits at ages younger than 62. The probability (or hazard) of receiving social security at the current age, given that the individual did not receive benefits at any earlier age, is estimated as a logit function.[26] Probability of onset of social security receipt at the current age is specified to be a function of age (dummy variables for each year of age 62-67), education, race/Hispanic ethnicity, pension coverage, earnings during various ages, spouse age, spouse education, spouse earnings, spouse pension coverage, gender and widowed status for unmarried, non-housing and housing wealth. Separate equations were estimated for married men, married women, and unmarried persons. The equations were estimated using the 1990-1993 SIPP matched with Summary Earnings Records (SER) and Master Beneficiary Records (MBR). Age at first benefit receipt is projected by comparing the probability predicted by the model to a random number for each person eligible for a benefit at each age from 62 to 67 until the person is designated to receive a benefit. All eligible individuals are designated to receive a benefit by age 67.[27]

Earnings After Retirement

MINT projects earnings from employment of individuals who are receiving social security retirement benefits at age 62 and at age 67 (referred to as "partial retirement earnings"). The projections are carried out in two steps. First, a statistical model is used to assign each individual to an earnings group: no earnings or one of three positive earnings groups. Second, each individual who was assigned to one of the three positive earnings groups is assigned a level of earnings based on a random draw from the 1990-1992 SIPP data. To determine if a social security benefit recipient has employment earnings and to which earnings group he/she is assigned, a set of ordered probit models are estimated.[28] An ordered probit model was estimated for social security retirement beneficiaries age 62 and 63 using the 1990-1992 SIPP data matched with the SER and MBR data. Ordered probit models were estimated for social security beneficiaries age 65-68 using two data sets – the 1990-1992 SIPP-SER-MBR data and the 1984 SIPP data matched with the SER and MBR. All of these data sets have information on covered earnings for 1951-1996 as well as year that the individual first received social security benefits. The 1984 SIPP data are used to estimate an ordered probit model because they include

---

[26] The *logit* is a statistical model, similar to the *probit,* that permits estimation of an underlying probability of a state or event for an individual using observations on the actual events in the sample. The logit differs from the probit in terms of the probability distribution function assumed. The probit assumes that the underlying distribution function is normal, while the logit assumes it is logistic.

[27] MINT III will include a retirement model for persons aged 50 and over, based on analysis of the Health and Retirement Survey (HRS). The model will predict change from a career job to no earnings or reduced earnings.

[28] An *ordered probit* model is a statistical model that permits estimation of an underlying probability of an individual being in one of two or more mutually exclusive groups which can be ordered (or ranked), as a function of characteristics of the individual and other variables.

employment and earnings data for individuals age 65 and older during the period 1977-1983, when the social security exempt amount was changing (from $3,000 in 1977 to $6,600 in 1983)[29] Ordered probit models were estimated to predict the probabilities that an individual would be in one of four earnings groups: (1) no earnings, (2) earnings greater than zero but less than 85 percent of the social security exempt amount, (3) earnings between 85 percent and 115 percent of the exempt amount, (4) earnings greater than 115 percent of the exempt amount. Explanatory variables in these equations include average earnings during various ages, pension income, wealth, IRA, Keogh, 401(k) balances, age, race, education, gender and marital status, and the social security exempt amount. Social security recipients are assigned to one of the four earnings groups at age 62 (67) using the appropriate ordered probit model and each individual's projected characteristics by drawing a random number between zero and one and comparing that number to the predicted probabilities that the individual would be in each of the four groups.[30] After the individual is assigned to one of the earnings groups, individuals assigned to one of the three positive earnings groups are assigned a level of earnings based on a random draw from the distribution of actual earnings of persons in that earnings group in the 1990-1992 SIPP.[31]

Asset Income After Retirement

At retirement age (62 or 67) defined contribution pension balances, including 401(k) and Keogh plan assets, and IRA balances are combined with non-housing non-pension wealth. MINT projects the income from these assets by carrying out two steps. First, MINT projects the change in the stock of these assets each year for each family or individual. Second, MINT estimates the income the assets would generate each year.

MINT projects the rate of change of total non-housing assets with age using a reduced form (descriptive) equation that relates the level of non-housing assets of individuals age 62 and older and married couples headed by persons age 62 and older to a set of demographic and economic variables. The level of non-housing assets is specified to be a function of social security benefits; defined benefit pension income; average earnings of head and spouse at ages 50-60; housing wealth; gender, marital status, race and ethnicity of unmarried persons; education; age at

---

[29] The social security exempt amount is the amount that a social security retirement beneficiary can earn without any reduction in the benefit payment.

[30] The documentation does not indicate which model was used for age 67 beneficiaries. The coefficient on the social security exempt amount in the ordered probit equation estimated for age 65-68 using 1984 SIPP data was not statistically significant and was negative (not the expected sign).

[31] During the model development, equations were estimated to predict the level of earnings, with reference to demographic and economic characteristics and the social security exempt amount, using the 1984 and 1990-1992 SIPP data. The coefficients in these equations in general were not statistically significant and the equations had little explanatory power, so they were not used to impute earnings levels. The MINT documentation asserts that the model simulates the effect of changes in the social security exempt amount on projected earnings, but it does not indicate how this is achieved.

onset of social security benefits; cohort; death of person within 27 months; and a set of interaction terms between age and homeownership, race, pension receipt, high earnings, widowed status. The model was estimated using data from the 1984, 1990, 1991, 1992, and 1993 SIPP panels, including pension and wealth topical modules, which were matched with the SER and MBR. The sample was limited to families that included persons born in or before 1922. The model was estimated for married couples and single individuals separately. The estimated coefficients on the interaction terms between age and other characteristics show how wealth changes with age for couples or persons with each of the interaction characteristics, i.e., how the presence of each of the characteristics changes the age-wealth slope. For example, the coefficient on the age-homeownership term estimates how wealth changes for each additional year of age of homeowners, compared to non-homeowners.[32] Using the relevant combinations of age-interaction terms for each couple or single person, the level of non-housing assets at the age of initial social security receipt is updated each year until 2020 or death.[33]

Income from the stock of non-housing assets each year is estimated as the annuity income those assets would provide that year, given the life expectation of the person.[34] Life expectation is calculated using the mortality rates of the mortality model, based on the individual's age, gender, race, and education. Married couples use a joint life expectancy and assume a 50 percent survivor annuity. Each year the stock of assets is decremented and the potential annuity income is recalculated. At the death of a spouse, the surviving spouse inherits all of the non-housing and housing assets. Non-housing assets then are decremented at the widow(er) rate.

The level of housing assets is assumed to be constant in real terms. Housing asset income is imputed based on a three percent per year real rate of return on home equity.

Projection of Total Retirement Income to 2020

Total income in each year after retirement is the sum of projected income from social security, earned income, defined benefit pension income, and income from non-housing and housing assets. Social security retirement benefits are determined by calculating the Average Indexed Monthly Earnings (AIME) from the projected lifetime earnings and calculating the Primary Insurance Amount (PIA).

---

[32] The dependent variable is the natural logarithm of wealth, so the exponentiated coefficient on an age-interaction term can be interpreted as a measure of the percent change in wealth per year of age for individuals or couples with the particular interaction characteristic.

[33] There are 12 different combinations of age-interaction terms for married couples and 16 combinations of age-interaction terms for single persons.

[34] The current version of MINT assumes that 80 percent of the assets are used to calculate the value of an annuity. This is a parameter that can be changed.

Income Taxes

MINT includes an auxiliary individual tax model that permits calculation of after tax income. The model calculates federal individual income taxes, FICA taxes[35], and state and local taxes (including average sales taxes and property taxes per person) based on the projected income of each person and couple in each of the four income components. The model is a simplified representation of federal tax laws of 1998. A parameter determines the proportion of total asset income (withdrawals) that is taxed (from both retirement savings balances and non-pension assets).[36] Individual state tax regimes are not identified. MINT represents total state and local tax burdens per person as a constant fraction of federal income tax liability, based on the median state and local tax burden.[37]

Stylized Earnings Patterns for Cohorts

As an ancillary study, the MINT I project developed a set of stylized lifetime earnings patterns for the birth cohorts 1931-1960. The stylized lifetime patterns reflect levels and changes in earnings over three periods in each worker's career: ages 31-40, 41-50, 51-60. Workers were classified into 27 groups based on the level and shape of their career earnings. Workers were classified in terms of three characteristics: (1) their lifetime average earnings over the 30 year period (low, middle, high), (2) the trend change in average earnings from youngest age group to oldest (falling, level, rising), (3) a profile over the three periods (sag, level, humped). Stylized earnings patterns were first developed based on historical data for the 1931-1940 birth cohort.[38] Stylized patterns were then developed for six five-year birth cohorts using the historical and projected earnings. Earnings patterns based on projected data, in general, display less variation than those based largely on actual data. The stylized lifetime earnings profiles were developed for two primary purposes: (1) to ascertain if the distributional implications of current social security benefits and proposed reforms differ when analyzed using lifetime earnings patterns that may more accurately reflect the variety of experience of actual workers than the smooth illustrative patterns often used, and (2) to provide simulated earnings data and patterns to the public for retirement income research and model development.[39] The stylized earnings patterns are not being used in the MINT III project.

---

[35] FICA is the sum of Old-Age, Survivors, and Disability Insurance (OASDI) and Hospital Insurance (HI) taxes.

[36] In the current version the parameter is set to equal 1. This parameter can be varied. Both tax sheltered assets (such as IRAs, Keoghs, and DC pension balances) and after tax savings are combined. The proportion of total asset income that is taxed is treated as ordinary income.

[37] In the current version state and local taxes are calculated as 83.5 percent of federal income taxes. This parameter can be varied

[38] Persons born in 1940 were age 56 in 1996, the last year of historical earnings data.

[39] The Summary Earnings Records data and the matched SIPP/SER database are not available to the public because of concerns about confidentiality. The data were used by SSA and contractor personnel only under highly restrictive conditions at government facilities.

Social Security Benefits and Policy Analysis

The demographic and retirement income projections provide a comprehensive data file that is processed by a set of algorithms developed at the Social Security Administration to estimate social security retirement benefits under current law and under alternative policy proposals. The input data file provides detailed information on socioeconomic characteristics, earnings histories, marital histories, assets, retirement age, post-retirement earnings, and pension coverage and pension benefits for the 1931-1960 birth cohorts, focusing on incomes at age 62 and age 67 and in the year 2020, when those cohorts are age 60-89. The social security benefits module calculates social security retirement and survivor benefits under current law and a variety of alternatives and compares the effects on income distributions. The social security benefits module also can analyze the effect of alternative individual accounts proposals. The social security benefits module is not documented.

Controls and Alignment

MINT projections, in the current version, have virtually no controls or alignment. Underlying macroeconomic variables are not projected by MINT but are assumed based on projections done by the SSA Office of the Chief Actuary (OCACT). These include the rate of change of the Consumer Price Index, the path of average earnings covered by social security, real rates of return on government securities and on equities. MINT does not project any complete populations of workers, contributors or social security beneficiaries, so it provides no aggregates which could be compared to other aggregate forecasts or projections for control or alignment. The MINT I mortality projections differ from those of OCACT. MINT I does not attempt to align disability or retirement benefit take-up rates to those projected by OCACT.[40]

Stochastic Features

MINT incorporates stochastic features in the determination of many individual-level events and states. No aggregate-level or macroeconomic variables are determined stochastically, but rather a single projected path is adopted based on exogenous projections. Many individual events and states are determined stochastically.

- Date of death is determined stochastically based on an individual-level hazard of mortality.
- Dates and durations of marriages and divorces are determined stochastically based on individual-level hazards of marriage and divorce.
- Onset of disability is determined stochastically based on the projected hazard of disability.
- A random error term is added to the estimated earnings of each worker in each year.
- Workers without pension coverage in the base data file are assigned coverage in future years randomly, based on projected future coverage rates.

---

[40] The modules for projection of mortality and disability have been revised in MINT II and MINT III. See footnote 10. In MINT III cohort-specific mortality and disability prevalence rates up to age 67 are aligned to OCACT projections. For ages older than 67 they are not aligned.

- Workers in DB pension plans are randomly assigned COLAs.
- Rates of return to individual DC pension assets, IRAs and Keoghs are determined stochastically based on assumed probability density functions, means and standard deviations for various types of securities.
- DC employer contribution match rates are assigned to workers randomly based on assumed distributions for 401(k) and non-401(k) plans.
- Individuals are randomly assigned to be homeowners (if not homeowners in the base data file) and to have positive non-housing wealth at ages 62 and 67, based on predicted probabilities. The values of home equity and non-housing wealth, for those determined to have positive wealth, are not determined stochastically.
- Timing of retirement (initial receipt of social security retirement benefit) between ages 62 and 67 is determined stochastically based on predicted probabilities of retirement.
- Whether an individual has earnings after retirement and in which earnings group is determined stochastically. Level of earnings is determined by a random draw from actual earnings distributions.

Documentation

The initial version of MINT (MINT I) is well documented. Comprehensive reports from the two original development contractors – the Urban Institute and the Rand Corporation – describe each component of the model and the research that went into its development.[41] Component models, data used for estimation, and statistical estimates and tests of significance are described. The Rand report also contains documentation and code of many of the SAS routines used to produce the MINT files. The SAS programs contain internal documentation. SSA has prepared a summary description of MINT, which provides a good, more concise description of its major components.[42]

Computer Hardware, Software, Portability

The MINT data file was created and is processed on personal computers. Efficient processing of the file at the Social Security Administration is done on PCs with 850 MHz to 1.2 GHz processors, 550 MB RAM, and hard drives of two GB. All the algorithms are written in SAS (although some developmental work was done in other languages). Developmental work was done for the most part in SAS running under Windows NT.

While, in concept, MINT is highly portable, in reality there are substantial barriers to its use outside SSA or SSA contractors. It does not require unusual or specific hardware or software (other than SAS). However, MINT is not an integrated, self-contained model. It is a set of algorithms that develop a data file, one step at a time, and that process that file for social security policy analysis. It may be difficult for a potential user who was not part of the development group to use or modify the system without considerable assistance from the

---

[41] Toder *et al.* (1999) and Panis and Lillard (1999).

[42] Butrica, *et al.* (1999).

developers. The greatest barrier to portability is that the base data file is not in the public domain. The matched SIPP/SER data are not available to the public because the individual earnings records are confidential (even though personal identifiers are not included in the data). Because the base year data file is an integral part of MINT, its use in its current form outside SSA would not be permitted.

IV. APPLICATION TO IMPORTANT RETIREMENT POLICY ISSUES

MINT is designed to analyze the income distributional effects of social security retirement benefit policies. Although it depicts all the major sources of income in retirement, those representations are all of a descriptive or "reduced form" nature – that is, the underlying structural or behavioral processes or institutions are not depicted. For example, employer pension coverage and benefits are assigned and estimated with reference to demographic and economic characteristics of workers and assumptions about years of service and job change behavior, without reference to characteristics of employers or of pension plans and how they may be affected by laws, regulations, and other institutions, or even by the social security system itself. So MINT has limited capability to address policies concerning employer pensions. MINT estimates lifetime earnings based on two attributes of workers, age and education, without reference to labor market activities or institutions or worker objectives or decisions. Retirement age and retirement earnings also are assigned using descriptive ("reduced form") equations based on worker characteristics, without reference to the operation of labor markets or labor market behavior. So MINT has limited capability to address issues concerning employment or earnings or older workers. MINT estimates the accumulation of wealth at retirement age and changes in wealth after retirement based on observations of relationships between individual demographic and economic characteristics and reported stocks of assets, without reference to savings behavior, tax laws, retirement savings incentives, rates of return, or other potential behavioral determinants. So it has limited capability to analyze retirement savings policies or the effects of social security policies on retirement savings. It can describe potential future employment and earnings patterns, and the wealth and income of the elderly that would provide the environment for future labor market and retirement issues.

The policy matrix tables in Annex 8-2 indicate the types of issues which MINT is suitable to address. There is a table for each of six major areas where changes in retirement income policy could have important implications. The rows, which are the same for each table, list several important areas where policy changes could occur. The columns, which differ in each table, show aspects which are relevant or of concern in each of the major areas. Entries in each table show areas where MINT is suitable for analysis of the impacts of the types of policy changes indicated in the rows on the aspects of the pension system indicated in the columns.

Examples of issues suitable for analysis with MINT include social security benefit policy issues such as changes in benefit formulae and indexing, earnings sharing, and adoption of individual accounts. MINT 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. MINT 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.

Benchmarks

Appendix C describes seven illustrative policy issues that can be used as benchmarks to assess and compare the suitability of various models for analysis of retirement income policy in various areas. This section reviews the suitability of MINT for analysis of each of these illustrative benchmark policy issues.

1. Effects of increase in Social Security Normal Retirement Age on:

OASI revenues, benefit payments, trust fund balances -- The MINT data file includes birth cohorts of 1931-1960 who ultimately receive social security retirement benefits. It does not include the complete population of social security contributors or beneficiaries. It is not designed to estimate social security revenues, aggregate benefits payments, or trust funds.

DI benefit payments and trust fund balances -- MINT does not simulate the complete population of DI beneficiaries or DI benefit payments. It is not designed to estimate aggregate DI benefit payments or trust funds. The current version of MINT estimates the prevalence of a disability health condition, and the onset of DI benefit receipt, primarily to adjust the earnings of DI beneficiaries.

Employer pension accruals and benefit costs -- MINT does not model employer pension plans or employer behavior. Pension accruals and benefit costs are not depicted.

Social security retirement replacement rates and total retirement income replacement rates -- MINT has the formal mechanics to depict the statutory effects of change in social security normal retirement age (NRA) on individual benefit payments, for given retirement behavior. Change in eligibility for OAI full retirement benefit and change in early retirement reduction can be programmed and will change the retirement benefit amount. However, MINT does not have a behavioral retirement model. None of the variables affected by the change in normal retirement age are represented in the MINT retirement benefit receipt timing equation. The MINT documentation states: "Unfortunately, we cannot directly estimate the effects of the Normal Retirement Age (or changes in the NRA) on retirement benefit receipt timing patterns because it has not varied in the years for which we have reliable data. We must therefore instead make assumptions . . . about whether changes in retirement behavior will accompany this change in policy and, if so, how large these changes would be." (Toder *et al.*, p. 134). *Ad hoc* assumptions about potential changes in retirement behavior could be made. MINT's lifetime earnings records and social security benefit calculator provide the capability to calculate social security replacement rates, for various assumptions about retirement behavior. The asset model does not depict effects of changes in retirement behavior or expectations on assets, nor does the pension model depict the effects of changes in retirement behavior on pensions. Consequently, MINT could not analyze effects of changes in the NRA on total retirement income. Assuming no changes in other sources of retirement income (or changes specified exogenously), MINT could calculate changes in total retirement income replacement rates.

2. Means testing of Social Security benefits.
MINT could be used to estimate the effects of means testing of social security benefits on individual benefit payments of persons represented in the data file, but not on aggregate benefit payments.[43] MINT depicts earnings, pensions, and asset income. It currently depicts the effects of the social security earnings test on individual benefits. MINT I does not have behavioral models of DI or OAI benefit receipt.[44] *Ad hoc* assumptions about benefit receipt and about the effects of means testing on other sources of income would be required. Effects on income replacement rates could be simulated. MINT does not depict employer costs, so the effects on employer costs could not be simulated.

3. Mandatory minimum employer pension.
MINT could be used to estimate effects on workers of a mandatory employer pension, for specified assumptions about the behavior of workers and employers. MINT includes information about current and projected employer pension coverage and lifetime earnings. MINT could not simulate the effects on employers or the potential changes in other pensions, wages, or employment. MINT does not depict DB employer pension accruals.

4. Expansion of individual retirement account eligibility.
MINT could estimate effects of expansion in IRA eligibility on individuals with *ad hoc* assumptions about the effects of changes in eligibility on IRA participation and contributions. MINT does not have a behavioral Individual Retirement Account model. Individual Retirement Account assets in the pension model are not modeled with reference to IRA rules. Because MINT does not depict saving behavior or macroeconomic effects, it could not estimate any policy effects on total retirement savings, personal savings, national savings, capital formation, or GDP. It does not depict federal expenditures, deficits, or debt, so it cannot analyze effects of IRA expansion on federal revenues or finances.

5. Effects of value added tax on pension contributions and accruals.
MINT cannot address issues concerning the effects of taxes on pension offerings, contributions, or accruals. The model does not depict the effects of taxes on individual or firm behavior. It does not depict the behavior of corporations or other employers, so it does not depict pension plan sponsors.

6. Effects of construction industry benefit accrual rates on funding.
MINT is not suitable to analyze issues concerning the condition or behavior of employer pension funds. MINT does not model employers, industries, sponsors, or pension funds.

---

[43] Note that MINT depicts social security benefits relative to average wages, so any means test must be represented accordingly.

[44] The MINT III retirement module will include social security and pension wealth as determinants of the age of retirement in a reduced form equation, which may permit depiction of behavioral response to changes in potential social security benefits.

7. Effects of alternative macroeconomic scenarios on social security and employer pensions. MINT is not designed for analysis of effects of alternative macroeconomic scenarios on either individual economic variables or other aggregates, such as social security fund balances. Aggregate economic variables do not enter any individual equations. The earnings, wealth, pension benefits, and other income amounts in MINT are all expressed relative to average earnings in the economy. MINT does not depict absolute individual or aggregate economic variables. It does not explicitly depict employment or the effects of changes in aggregate employment. MINT does not depict employer pension fund balances or accruals.

V. ACCESSIBILITY, EASE OF USE, AND COST

MINT is a product of the U.S. government and in principal is in the public domain, and it is well documented. However, direct accessibility and use by the public of the complete modeling system may be limited. The main reason is that the base historic data file is not available to the public. The core data file is the 1990-1993 SIPP panels matched with the Social Security Summary Earnings Records (SER) for 1951-1996 and Master Beneficiary Records (MBR). Neither the SER, which provides actual earnings records for the individuals in the SIPP, nor the MBR, which provides dates of entitlement to social security benefits and benefit amounts, are available to the public. The MINT projected data files, for 1998-2030, may be available to the public.

MINT is not a self-contained model. It is a set of algorithms that develop a data file, one step at a time, and that process that file for social security policy analysis. Although all the component algorithms are in the public domain and the developmental algorithms are documented, it may be difficult for a potential user who was not part of the development group to use or modify the system without considerable assistance from the developers.

MINT is fairly narrowly specialized and focused on its objective, which is to provide a tool for analysis of the distributional effects of social security policy on the cohorts born 1931-1960. MINT focuses on the income and characteristics of social security retirement beneficiaries of those cohorts at age 62 and 67 and in year 2020. Although MINT provides other information, including career annual earnings and simulation of several income sources annually for ages after 62, its focus is ages 62, 67 and year 2020.

The predecessor to MINT was developed by SSA staff of two to three persons during the period 1996-1998. Contracts for MINT development during the 1998-2001 period have totaled about $3.3 million. In addition to contract assistance, SSA has had staff of approximately seven to ten full-time-equivalent personnel working on MINT during the 1998-2001 period.

MINT provides a new and innovative approach to social security and retirement income policy analysis. MINT was proposed and initiated with the expectation that it could provide the information needed for analysis of a number of retirement income policy issues more quickly and at lower cost than development of a complete dynamic microsimulation model. The National Academy of Sciences National Research Council endorsed the approach in an important 1997 study of retirement income policy data and modeling needs, concluding, "This [SSA modeling] approach could offer some useful policy modeling capability without requiring a full-fledged dynamic microsimulation model." (Citro and Hanushek, 1997).

Model development projects sponsored by SSA may provide a valuable opportunity for observation and evaluation of alternative modeling approaches. In addition to MINT, SSA has acquired the CORSIM model, described in Chapter 5, and is currently sponsoring a major revision and augmentation of the model, for the purpose of simulating the effects of alternative social security policies. CORSIM is a full dynamic microsimulation model. With sponsorship of SSA and considerable internal staff resources, a comprehensive social security benefits module has been developed in CORSIM. The acquisition and contract costs of CORSIM have been comparable to those of MINT, about $3.1 million over the 1998-2001 period. SSA staff of about two to three full-time-equivalents work on the CORSIM project. The two models are not substitutes. They have different base data files, different approaches, and different simulation methodologies. They may well provide complementary approaches and information. It may be very useful for the policy modeling community to observe the parallel development and application of the two modeling systems, including comparisons of costs and capabilities.

VI. CRITIQUE

MINT is an innovative and promising approach to the development and projection of information for retirement income analysis. The approach is essentially to project directly the future demographic states and incomes of a significant range of birth cohorts for which considerable recent historic information is available on the characteristics and interrelationships that will affect retirement income. This approach emphasizes the maximum use of actual, recent, historic information. MINT uses and, in turn, projects the largest, most comprehensive, most up-to-date micro database available, with a broad range of data elements and a large sample. The base data file is composed of the 1990-1993 SIPP panels, which have comprehensive information on socioeconomic characteristics, incomes, employment, health, assets, pension coverage and pension plan characteristics, work and disability histories, and marital histories of approximately 84,500 persons in the 1931-1960 birth cohorts.[45] These data are linked with social security earnings records for the period 1951-1996 and social security Master Beneficiary Records with data on dates of disability, retirement, and survivor benefit entitlements and benefit amounts, and social security Numident files with dates of death through 1998. This data file is projected into the near-term future – about the next 20 years – to the year 2020 when the members of the 1931-1960 birth cohorts will be age 60-89.

This approach has a number of limitations. The population being projected – cohorts of 1931-1960 – does not provide the entire population affected by the social security system at any time during the projection period. Consequently, the model cannot analyze the effects of any policy on all beneficiaries. It cannot analyze issues concerning contributors or revenues. MINT cannot estimate or analyze any effects on the aggregate condition, revenues, or benefit

[45] Of these, approximately 59,500 have positive weights, indicating that they were present for the entire SIPP data collection period.

expenditures of the social security system. The projections are limited to the near term. Simulations after the year 2030, when the sample will be age 70-99, have limited use. MINT may address this limitation by adding later birth cohorts as data become available in later SIPP panels. The 1996 SIPP panel will provide data for 1996-2000, during which period the 1961-1970 birth cohorts will be ages 30-39.

MINT cannot analyze issues of employer pension policy. It has no information on employers or plan sponsors. The pension module of the current version (MINT I) has been criticized for a number of limitations. It has a very simplistic defined benefit projection methodology, relying on average replacement rates from BLS surveys. It has simplified and perhaps unrealistic job change assumptions, which do not take into account worker characteristics. It does not incorporate cash balance plans, which are becoming an important proportion of plan offerings. It has very simplified and perhaps unrealistic assumptions concerning employer contributions to defined contribution plans, including 401(k) plans, and employer match rates. It has very simplified and perhaps unrealistic assumptions concerning asset allocation of contributions and balances in DC plans, 401(k)s, IRAs, and Keogh plans. It does not take into account information on employee cash-outs of pension benefits prior to or at retirement, which may become increasingly significant with the increased prevalence of cash balance plans. Workers are assigned only one type of pension plan, ignoring the significant proportion of workers with more than one employer plan. Many of these issues may be addressed in revisions to the pension model.

MINT projects sequentially the demographic states, then earnings, pension accruals, asset accumulation, retirement, and retirement incomes of the population, then uses the projected data file to analyze the effects of social security policy. This precludes incorporating any feedbacks of social security policy on labor market behavior, earnings, pension offerings or accruals, retirement savings and asset accumulation, retirement ages, or retirement earnings. Most of the income and asset projections are based on descriptive, or "reduced form" equations, which incorporate little or no economic behavior. The wealth projection models do not provide for feedbacks or links between social security changes and savings behavior. There is no behavioral retirement model. This limits the ability of the model to depict the effects of changes in social security policy on the retirement decision and on post-retirement earnings. While recognizing that estimation of structural or behavioral models is difficult and the prospects are uncertain, users of the model should be aware of the potential limitations of failing to take into account behavioral responses to program and policy changes.

MINT has very limited depiction of labor market behavior. There is no explicit representation of labor force entry or participation, employment, job change, job or employer characteristics, or industry identification. This limits the range of social security policy effects that can be analyzed, and limits the potential richness of any employer pension model.

MINT is a work in progress. It represents a promising and innovative approach to retirement policy modeling. In particular, it represents the most comprehensive effort currently underway to utilize to the maximum extent the most comprehensive and most up-to-date microdata available for the analysis of retirement income policy. The use of this large and comprehensive data set, which incorporates the actual covariants among a large number of demographic and socioeconomic variables and recent trends and behaviors of the individuals in the population for near term projections, is of potentially great value.

ANNEX 8-1

MINT

SUMMARY DESCRIPTION TABLE

MINT

Summary Description

Subject: Economic well-being of members of the 1931-1960 birth cohorts from age of first Social Security receipt through 2020

Purpose and Objective of Model:
- Policy analysis: distributional effects of Social Security policy changes
- Descriptive analysis: baseline projections of aged economic well-being

What is Represented in the Model (Universe): Demographic status and income of persons born 1931-1960 at retirement (age 62 and 67 and year 2020)

Period of historical analysis: earnings – 1951-1998 (Social Security Summary Earnings Records); mortality – 1968-1994 (PSID), 1901-1994 (vital statistics); wealth – 1984-1994 (PSID); socioeconomic status, pension coverage, disability – 1990-1995 (SIPP)
Forecast/simulation horizon: Targeted to 2020, though most projections are actually made as far as 2031
Frequency: MINT is not a periodic model. Simulation targets are individuals at age 62 and age 67 and year 2020. Simulation output may be obtained on annual basis for some economic variables. Demographic events occur in continuous time (i.e., more than one event can occur in same month).
Base year: 1990-1995 (1990-1993 SIPP panels)
Simulation technique: Static microsimulation, with some dynamic elements, cross-sectional imputation
Solution algorithms and structure: Sequential/recursive. Each process is solved for complete lifetime, using lifetime solutions to processes solved earlier, but independent of processes solved later. Lifetime demographic outcomes are solved initially and are independent of income or labor market activity processes.
Unit(s) of analysis: Individuals and married couples

<u>**Databases**</u>
Base year database: Pooled panels of the 1990 through 1993 Survey of Income and Program Participation (SIPP), matched to Social Security Summary Earnings Records (SER); 113,716 person records (78,161 with positive weights)
Population/demographics: Mortality - Panel Study of Income Dynamics (PSID) 1968-1994, Vital Statistics of the United States 1901-1994, base year individual mortality updated by Social Security Numident files; Marriage/remarriage and divorce - SIPP 1990-1993 Marital History Topical Modules; Disability - SIPP 1990-1991
Individual/family/household characteristics: 1990-1993 SIPP
Employer characteristics: None (some employer data available on SIPP)
Industry characteristics: None (industry of employment from SIPP)
Retirement plan structure and characteristics: 1990-1993 SIPP Retirement Expectations and Pension Plan Coverage topical modules, 1993 BLS tables, PBGC Pension Insurance Management System (PIMS) model

Retirement plan coverage, participation: 1990-1993 SIPP Retirement Expectations and Pension Plan Coverage and Annual Income and Retirement Accounts topical modules, BLS tables, PIMS
Retirement plan vesting: Default assumption of 5 years for DB plans, none for DC plans
Individual Retirement Account (IRA) participation, contributions: SIPP topical modules on assets and liabilities
Supplementary Security Income (SSI) eligibility, participation: 1990-1993 SIPP matched to SSA Supplemental Security Records (SSR)
Household assets: SIPP Asset and Liability topical modules, 1990-1993; PSID, 1984, 1989, 1994 (from SIPP observation until age 50) calibrated to values on SIPP topical module; Health and Retirement Survey (HRS) (ages 51-67); Social Security New Beneficiary Survey (NBS) (from age of first Social Security receipt to age of first receipt plus ten); Assets and Health Dynamics Among the Oldest Old (AHEAD) database (from age of first Social Security receipt plus ten until death)
Homeownership: SIPP Asset and Liability topical modules, PSID, HRS, NBS, AHEAD calibrated to value on SIPP topical module (same age ranges as above)
Macroeconomic data: None analyzed or simulated, macroeconomic controls provided exogenously
Financial data: No aggregate financial data. Interest rate exogenous. Individual and family financial assets from SIPP topical modules, 1984, 1990-1993; and PSID, 1984, 1989, 1994.
Labor market data: SIPP, job change assumptions (used in pension module) based on data from Gustman and Steinmeier (1995), Holmer
Taxes: Kroes (1997, 1998); Federal income tax rules (Form 1040)
Health conditions: SIPP for date of onset of health condition that limits amount or type of work a person can do; HRS for health ages 51-67; SIPP for health ages 68+
Disability: DI participation from Social Security Master Beneficiary Record (MBR); predicted using SIPP matched to MBR and SER
Health insurance coverage: None
Institutional population coverage: None

<u>Data Quality</u>
Completeness: Comprehensive person and family data available for limited model universe of cohorts born 1931-1960; no immigration
Accuracy: All data considered accurate, with recognized limitations of panel survey databases. SIPP assets data are consistent with other survey data for lower to upper-middle level asset holders. In general, no assets survey databases are highly accurate for high level asset holders. High asset holders are not of importance for this model.
Representative: Base data file is designed to be representative of the U.S. civilian non-institutional population in 1990-1995, born 1931-1960
Currency: Base data file is for 1990-1995; administrative records cover years to 1996, latest revisions in 1999
Applicability to other contexts: Could readily be extended for a variety of applications
Gaps: Models only birth cohorts of 1931-1960. Depicts income at ages 62 and 67 and in year 2020 of social security benefit recipients only. Institutionalized and military populations of 1990-1993 excluded. Immigrants arriving after 1993 excluded.

Applicability of other private/consulting firm data: Data on employer pension plan provisions, lump-sum distributions, DC and 401(k) investment allocations and rates of return would be very useful.

<u>Characteristics, activities, behaviors that are modeled</u>

Demographic characteristics: death, marriage, divorce, disability

Aggregate economic activity

> **Short-run/cyclical:** Economy is exogenous, no cyclical activity.
>
> **Long-run growth, productivity:** exogenous
>
> **Inflation:** exogenous
>
> **Industrial sector detail:** not included
>
> **Open or closed economy:** N/A

Labor force characteristics and labor market behavior: For individuals, age when begin social security benefit receipt, employment and earnings after benefit receipt; no labor demand equations nor supply-demand interactions.

Capital markets: not included

Retirement plan/program characteristics: replacement rates, balance accumulations, COLAs

Retirement behavior: Retirement decision determined by change of social security and pension wealth from additional year of work ("premium value") (Gustman and Steinmeier, 2000). Social Security benefit take-up based on retirement decision.

Individual/family savings and asset accumulation – Savings not modeled explicitly. Assets assigned to families based on socioeconomic characteristics and income in reduced form equations.

> **Pensions:** yes
>
> **IRAs:** yes
>
> **Financial assets:** Aggregate non-housing assets (vehicle, other real estate, farm and business equity, stock, mutual fund, and bond values, checking, saving, money market, and certificate of deposit account balances, less unsecured debt) are assigned.
>
> **Home equity:** values assigned

Plan sponsor asset accumulation: no

Aggregate national saving and capital formation: no

Government behavior

> **Federal budget:** not represented
>
> **OASDI and HI trust funds:** Model does depict contributions and does not include full population of taxpayers and beneficiaries, so it cannot depict trust funds.
>
> **Regulations:** minimal explicit
>
> **Taxes**: Federal - based on tax rules (Form 1040 and tax rate schedules); FICA; state and local taxes modeled as percent of federal

Public retirement income programs – Rules affecting individuals; no behavior or status of programs

> **OASDI:** tax rules, coverage and benefit rules
>
> **SSI:** federal guarantee, state supplements, take-up behavior
>
> **Other means-tested old age or disability income transfers:** none

Government employee pension programs – Behavior and status of programs not modeled.

> **Federal civil service:** benefit accruals and benefit receipt of individual participants
>
> **Military:** benefit accruals and benefit receipt of participants

State and local government, types: benefit accruals and benefit receipt of participants
Private pensions – Behavior and conditions of plans or sponsors not modeled
Defined benefit: benefit payments to participants, including COLAs
Defined contribution: account balances, annuitization
401(k): yes
Non-401(k): yes
Supplemental: no supplemental plans
Individual retirement saving arrangements (IRA, Keogh, etc.): IRAs modeled as part of financial assets. Keoghs are included as a separate category with private pensions rather than with IRAs.
Public sector health care finance programs
Medicare: no
Medicaid: no
Military/CHAMPUS: no
Veterans: no
Indian Health Service and others: no
Private sector health care finance programs: no
Private health insurance, especially retiree health insurance: no
Employer/plan sponsor behavior: no
Worker behavior: Employment, earnings, retirement (pension acceptance coincides with retirement), social security acceptance, employment and earnings after social security acceptance
Health care provider behavior: no
Insurer behavior: no
Sources of family and individual income (pre-retirement): earnings only; not separated by self-employed status
Components of retirement income
Earnings: yes
Social security: yes
Employer pensions: yes (DB)
Means-tested transfer income: yes (SSI only)
Income from assets: yes (including DC pension balances, and imputed housing rent)
Other: none
Consumer/family expenditures/uses of income:
Taxes: Federal, FICA, state and local
Health care: no
Other: none
Institutionalization: no, models transition into nursing homes as absorbing state

Assumptions, Parameters, Methodology

Key Assumptions: Future demographic, labor market, earnings, pension accrual, asset accumulation, and retirement behavior will be like the past, as reflected in 1968-1994 PSID, 1990-1993 SIPP, 1993 BLS data. U.S. mortality trends of 1901-1994 continue unchanged. No changes in social security rules, tax rules in base case after 1998.

Types of Parameters, Decrements, Transition Rates/Probabilities

Experience considered, origins of decrements: Primarily incidence rates, based on analysis of PSID 1968-1994, social security Numident files 1990-1996, U.S. Vital Statistics 1901-1994 (mortality rates), SIPP 1990-1993 and marital history topical modules (marriage, remarriage, divorce, disability)

Consistency with other experience and other assumptions of model: Generally consistent – explicit discussion of consistency issues by Rand called for in initial contract

Internal consistency: Model is generally internally consistent. Most events specified as independent logit or probit functions (competing risks in demographic modules). Pensions affect Social Security take-up. Social Security take-up affects subsequent earnings.

Methodology used to estimate parameters and relationships

Econometric/statistical: single equations estimated by ordinary least squares (OLS); logit, probit, random effects, fixed effects, proportional hazard specifications

Actuarial: Specification of program rules, part of pension methodology

Judgmental: yes

Economic/actuarial literature, studies done by others, etc.: yes (e.g. Gustman and Steinmeier, 1990, 1995)

Simulation Methodology: static microsimulation with dynamic elements, cross-section imputation

Stochastic Properties: probabilistic assignment of demographic events based on random number draws; stochastic assignment of individual deviations from assumed average rates of return

Feedback Phenomena Minimal feedback phenomena in MINT I. More feedback in MINT III is being implemented.

Microsimulation adjustment ("aging") methodology (where relevant): Static aging, dynamic elements; demographic events and labor market status assigned sequentially in life histories (e.g., mortality; marriage, divorce/widowhood, remarriage; disability; lifetime earnings; retirement)

Policy levers: Social Security tax and benefit computation rules, SSI benefit rules

Economic/demographic feedbacks

Employer costs and behavior: none

Labor market behavior: Limited; individuals receiving Social Security target earnings to Social Security exempt amount; pension coverage and pension wealth influence retirement; retirement influences Social Security receipt**.**

Taxes, government deficits, etc: none

Capital accumulation: minimal

Interest rates: exogenous

Employment, productivity, economic activity, GDP: exogenous

Other: none

Sensitivity Analysis: Simulations can be conducted with alternative parameters. Since most economic and demographic outcomes are not aligned or calibrated, sensitivity tests are fairly complex. Since the MINT output data file is produced by carrying out numerous, separate processing tasks (which are not integrated into a single program), performing sensitivity analysis is a relatively lengthy and time consuming task.

Model Validation Procedures: Limited – Early version of SSA/DPE model based on 1974-1983 data was used to project 1984-1993 earnings and compared to actual data (Iams and

Sandell, 1997). MINT simulations of demographic variables (mortality, marital status) have been compared to SSA OCACT projections. Urban Institute proposes to undertake validation by comparing MINT simulations with 1999 and 2000 historical data.
Alignment: None. MINT does not project complete aggregates which can be aligned to other aggregate forecasts. MINT uses SSA OCACT assumptions/projections regarding Consumer Price Index, average social security covered wages, as inputs. MINT is not aligned to OCACT demographic projections.
Is model proprietary, available to public
No. Aging algorithms are in the public domain. Base year database and model output are restricted because of confidentiality restrictions (individual social security earnings records and benefit receipt records are not available to public).
Documentation: MINT I documentation is complete and extensive. Modifications to MINT I (MINT II) have not been documented.

Computer implementation

Hardware requirements	Personal computers (Pentium machines)
Software	SAS language, NT operating system
Computer costs	Running costs marginal
Transportability	High

Applications (Projects, Studies)
Application of existing model for policy analysis has been limited. Descriptive studies of retirement income trends have been presented. Potential applications to social security policy include effects of benefit rule changes on income of various socioeconomic groups, effects of diversion of contributions to individual accounts.

Contact Person
Howard Iams, Office of Research, Evaluation and Statistics, Social Security Administration

ANNEX 8-2

MINT

POLICY MATRICES

MINT -- POLICY MATRIX

1. Effects of Policy Measures on Employer Pensions

Outcome Variable **Policy Input**	Offerings	Types of plans and provisions	Costs of plans	Funding	Contributions and benefits
Tax Policy					
Pension					
General					
Social Security					
Retirement age					
Benefit structure					
Indexation					
Payroll tax					
Trust fund investment					
Individual accounts					
Funding and Guarantees					
PBGC premium					
Funding rules					
Pension Regulation and Policy					
ERISA/IRS					
Employer plans					
Pension and saving incentives/mandates					

Blank cell indicates that the effects of the policy issue or input on that outcome cannot be simulated in this model.

MINT -- POLICY MATRIX

2. Effects of Policy Measures on Employees

Outcome Variable **Policy Input**	Job availability	Portability	DC accumulations, investments, earnings	Benefit accruals	Wage and non-wage compensation levels and mix	Incidence and timing of retirement
Tax Policy						
Pension						
General						
Social Security						
Retirement age						
Benefit structure						
Indexation						
Payroll tax						
Trust fund investment						
Individual accounts						
Funding and Guarantees						
PBGC premium						
Funding rules						
Pension Regulation and Policy						
ERISA/IRS						
Employer plans						
Pension and saving incentives/mandates						

Blank cell indicates that the effects of the policy issue or input on that outcome cannot be simulated in this model.

MINT -- POLICY MATRIX

3. Effects of Policy Measures on Retirees

Outcome Variable / **Policy Input**	Payouts	Funded levels of plans	Retirement income	Replacement rates	Poverty levels	Health care costs and insurance	Retirement age and labor mkt outcomes	Inflation protection	Auxiliary benefits
Tax Policy									
Pension									
General									
Social Security									
Retirement age			x	x	x		x	x	x
Benefit structure			x	x	x		x	x	x
Indexation			x	x	x		x	x	x
Payroll tax			x	x	x		x	x	x
Trust fund investment			x	x	x		x	x	x
Individual accounts			x	x	x		x	x	x
Funding and Guarantees									
PBGC premium									
Funding rules									
Pension Regulation and Policy									
ERISA/IRS									
Employer plans									
Pension and saving incentives/mandates									

Blank cell indicates that the effects of the policy issue or input on that outcome cannot be simulated in this model.

MINT -- POLICY MATRIX

4. Effects of Policy Measures on Industry Outcomes

Outcome Variable **Policy Input**	Financial strength of plans, sponsors, insurers	Labor costs	Profits	Competitiveness
Tax Policy				
Pension				
General				
Social Security				
Retirement age				
Benefit structure				
Indexation				
Payroll tax				
Trust fund investment				
Individual accounts				
Funding and Guarantees				
PBGC premium				
Funding rules				
Pension Regulation and Policy				
ERISA/IRS				
Employer plans				
Pension and saving incentives/mandates				

Blank cell indicates that the effects of the policy issue or input on that outcome cannot be simulated in this model.
Industries are not modeled in MINT. MINT shows effects only on individuals.

MINT -- POLICY MATRIX

5. Effects of Policy Measures on Aggregate Economy

Outcome Variable / **Policy Input**	GDP growth	Saving and capital accumulation	Equity investment	Investment efficiency	Interest rates	Productivity	Inflation	Labor mobility and labor market flexibility
Tax Policy								
Pension								
General								
Social Security								
Retirement age								
Benefit structure								
Indexation								
Payroll tax								
Trust fund investment								
Individual accounts								
Funding and Guarantees								
PBGC premium								
Funding rules								
Pension Regulation and Policy								
ERISA/IRS								
Employer plans								
Pension and saving incentives/mandates								

Blank cell indicates that the effects of the policy issue or input on that outcome cannot be simulated in this model.
Aggregate economy is not modeled in MINT.

MINT -- POLICY MATRIX

6. Effects of Policy Measures on Government Finances

Outcome Variable **Policy Input**	Tax revenue	Expenditures by program	Deficits and debt	Social Security and Medicare
Tax Policy				
Pension				
General				
Social Security				
Retirement age		x		x
Benefit structure		x		x
Indexation		x		x
Payroll tax		x		x
Trust fund investment		x		x
Individual accounts		x		x
Funding and Guarantees				
PBGC premium				
Funding rules				
Pension Regulation and Policy				
ERISA/IRS				
Employer plans				
Pension and saving incentives/mandates				

Blank cell indicates that the effects of the policy issue or input on that outcome cannot be simulated in this model.