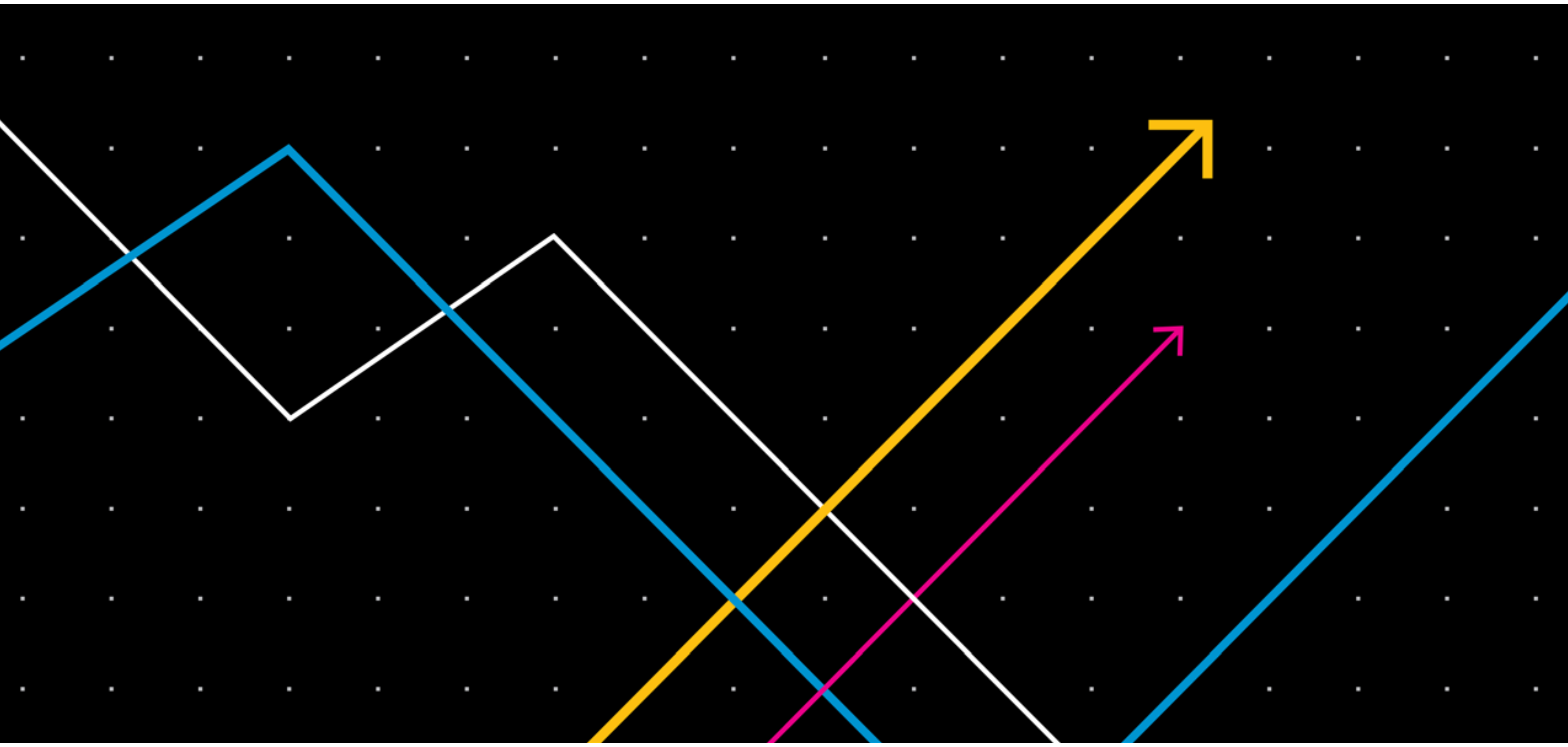


RESEARCH REPORT

Modeling Income in the Near Term 8 and 2014

Primer

Karen E. Smith
URBAN INSTITUTE

Melissa Favreault
URBAN INSTITUTE

April 21, 2019

URBAN INSTITUTE · ELEVATE · THE · DEBATE

ABOUT THE URBAN INSTITUTE

Contents

Acknowledgments

This report was funded by the Social Security Administration. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples.

Diana Tucker was the Social Security Administration contract officer for MINT8. We thank Ms. Tucker and Office of Research, Evaluation, and Statistics (ORES) staff including Mark Sarney, David Shoffner, Patrick Purcell, Glen Springstead, Joni Lavery, Gayle Reznik, Irena Dushi, and Nick Love for their input throughout the contract and for their reviews of the project deliverables. We also thank Thuy Ho for her assistance with disclosure review. Kent Morgan and Michael Morris provided us with updated demographic projections from POLISIM that MINT8 used to generate individuals born after 1979. Antwain Pair provided us invaluable technical support.

Karen Smith was the principal investigator for the MINT8 contract, but MINT8 represents the cummulative contributions of numerous researchers who have worked on MINT since its first implementation in 1999. While not an exhaustive list, Karen Smith, Melissa Favreault, Eric Toder, Gary Burtless, Barry Bosworth, Claudia Sahm, Cori Uccello, Richard Johnson, Jon Bakija, Timothy Waidmann, Kevin Perese, John O'Hare, Caroline Ratcliffe, Barbara Butrica, Douglas Wissoker, Aaron Williams, Stipica Murdrazija, Jillian Berk, Lawrence Thompson, Douglas Wolf, Constantijn Panis, and Lee Lillard made substantive contributions to the model development over the years under the direction of Social Security Administration project director, Howard Iams, who directed all prior versions of MINT.

Abstract

Modeling Income in the Near Term (MINT) is a dynamic microsimulation model developed by the Social Security Administration to facilitate analysis of proposals to change Social Security benefits and payroll taxes. This primer describes MINT's development history. It then details the model's starting sample and the specification of its demographic and economic aging modules, including the calculators that compute various benefits and taxes. It also provides information about previous analyses that have relied on MINT.

Acronyms

ACA	Affordable Care Act
AMT	Alternative Minimum Tax
AWI	Average Wage Index
CBO	Congressional Budget Office
CBOLT	Congressional Budget Office Long-Term dynamic microsimulation
DACS	Disabled Adult Child Status
DER	Detailed Earnings Record
DI	Disability Insurance
DOMA	Defense of Marriage Act
DPE	Division of Policy Evaluation
FICA	Federal Insurance Contributions Act
FPL	Federal Poverty Level
GPO	Government Pension Offset
HRS	Health and Retirement Study
ISM	In-Kind Support and Maintenance
IRS	Internal Revenue Service
LIHEAP	Low Income Home Energy Assistance Program
MBR	Master Beneficiary Record
MEPS	Medical Expenditures Panel Survey
MINT	Modeling Income in the Near Term
NBER	National Bureau of Economic Research
OACT	Office of the Chief Actuary
OASDI	Old-Age, Survivors, and Disability Insurance
OASI	Old-Age and Survivors Insurance
OLS	Ordinary Least Squares
ORES	Office of Research, Evaluation, and Statistics
ORP	Office of Retirement Policy
PBGC	Pension Benefit Guaranty Corporation
PIA	Primary Insurance Amount
PIMS	Pension Insurance Modeling System
PSID	Panel Study of Income Dynamics
RET	Retirement Earnings Test
SCF	Survey of Consumer Finances
SECA	Self-Employment Contributions Act
SER	Summary Earnings Record
SIPP	Survey of Income and Program Participation

SNAP	Supplemental Nutrition Assistance Program
SOI	Statistics of Income
SSA	Social Security Administration
SSI	Supplemental Security Income
SSR	Supplemental Security Record
TANF	Temporary Assistance for Needy Families
TDI	Temporary Disability Insurance
TRIM	Transfer Income Model
UI	Unemployment Insurance
WEP	Windfall Elimination Provision
WIC	Women, Infants, and Children

Overview

This primer describes Modeling Income in the Near Term, Version 8 (MINT8). MINT8 is a tool developed by the Office of Research, Evaluation, and Statistics (ORES) of the Social Security Administration (SSA) for use throughout SSA to analyze the distributional consequences of proposals to modify the Old-Age, Survivors, and Disability Insurance (OASDI) program, better known as Social Security.[1] Contractors from the Urban Institute, the Brookings Institution, and the RAND Corporation have contributed to MINT's development. Other MINT contributors have included Jon Bakija from Williams College, John Coder of Sentier Research, Martin Holmer from Policy Simulation Group, and Douglas Wolf from Syracuse University. Over the years, several advisory boards have also provided input on MINT.[2]

We do not intend to fully document MINT in this primer.[3] Rather, we provide a broad, high-level overview and extensive citations to more comprehensive documents for those seeking additional or more detailed information on model specifications. We also place MINT8 in historical context, detail published analyses that use MINT, and make recommendations for using the model effectively.

MINT8 is based on a micro-level data file of actual and projected individuals born from 1905 to 2068. It starts with a large sample of individuals from the Survey of Income and Program Participation (SIPP) with a rich set of income and demographic characteristics. Individuals in SIPP who were born prior to 1980 are linked to their SSA and other government records on earnings, benefits, and mortality.[4] MINT then projects life outcomes for these core birth cohorts—1905 to 1979—until death or the year 2099. MINT also projects life outcomes for individuals in extended cohorts--born 1980 through 2068--but using a somewhat different projection approach than for the core cohorts. The projections for the extended cohorts use a synthetic population born from 1980 to 2068 generated from the SSA's POLISIM model. MINT selects individuals at age 31 and links them to similar aged SIPP respondents. This link provides a starting point for the extended cohorts that includes the rich histories for individuals in the core cohorts through age 31 including prior earnings, job characteristics, assets, marriage, fertility and immigration histories, and health measures. MINT8 uses various algorithms to project outcomes for each individual from the interview year until death or the year 2099.[5]

While MINT starts with the SIPP data, it relies on other data sources when projecting some outcomes, based on the rationale that any function should use the best available data.[6] These sources include the Health and Retirement Study (HRS), Panel Study of Income Dynamics (PSID), Survey of Consumer Finance (SCF), and Medical Expenditures Panel Survey (MEPS). MINT8 calibrates many key outcomes to the 2018 Social Security Trustees' intermediate assumptions (OASDI Board of Trustees 2018). Calibrated outcomes include future price and wage growth and key Social Security benefit formula parameters.[7] Demographic outcomes in MINT, such as Disability Insurance benefit receipt, life expectancy, fertility, and net immigration, are also tied to the Trustees' intermediate assumptions.

Dynamic Microsimulation

MINT is a microsimulation model. The microsimulation modeling strategy was first conceptualized by Guy Orcutt (1957). Orcutt's vision is that bottom-up representations of economic and demographic processes, combined with detailed representations of program rules, can inform policy by revealing interactions and trends that more aggregate analyses may fail to capture.

Microsimulation models can be either *static* or *dynamic*. Static models typically simulate the immediate effects of a change in law or policy on the current population. Dynamic models simulate the effects of changes in law or policy on a population over time, sometimes over very extended horizons. MINT falls into the latter category. Either type of model may or may not simulate behavioral responses that might result from changes in law or policy.

History of MINT

MINT is based on the "past is prologue" approach Iams and Sandell (1997) outlined. Table 1 provides a succinct history of MINT's development, with references to model documentation and analyses. As the table indicates, MINT1 was produced by analysts from the Urban Institute, the Brookings Institution, and the RAND Corporation and is described in Toder et al. (1999) and Panis and Lillard (1999). Analysts from the Urban Institute again collaborated with Brookings Institution researchers for MINT3, described in Toder et al. (2002). Analysts from the Urban Institute developed subsequent versions of MINT. MINT4, MINT5, MINT6, MINT7, and MINT8 are described, respectively, in Smith et al. (2005), Smith et al. (2007), Smith et al. (2010), Smith and Favreault (2013), and Smith, Williams, and Murdrazija (2019). A number of articles and working papers, including several published in the *Social Security Bulletin*, have also described MINT findings and methods (see last column of table 1).

Each subsequent version of MINT has enhanced the earlier version by adding more recent data, updating assumptions, refining the projection methods, and expanding the suite of modeled outcomes and sample cohorts.

MINT is widely used to provide distributional forecasts of changes to Social Security. For example, MINT projections were used to analyze the distributional impact of reforms included in the Bipartisan Policy Center's Commission on Retirement Security and Personal Savings proposal (Sarney 2017) and for the National Commission on Fiscal Responsibility and Reform report (2010; see f-igure 13, page 55). Analysts from SSA's Office of Research, Evaluation, and Statistics (ORES) further provide the public with results from MINT through SSA's website, for example, which includes population projections (Social Security Administration 2015a) and analyses for prominent provisions to change Social Security (SSA 2011).[8,9] Researchers from various Social Security offices have also used MINT projections in published articles about proposals to change Social Security that have appeared in academic journals and similar venues (for example, Romig and Whitman 2016; Iams 2016; Purcell 2015; Purcell, Iams, and Shoffner 2015; Springstead, Whitman, and Shoffner 2014; Knoll and Olsen 2014; Whitman and Shoffner 2013; Olsen 2012; Whitman, Reznik, and Shoffner 2011; Iams, Reznik, and Tamborini 2010; Olsen and Romig 2013; Reznik, Weaver, and Biggs 2009; Shoffner 2010; and Tamborini and Whitman 2008, 2010; Whitman 2009; Purcell and Shoffner 2004). MINT is also frequently used for policy briefs (for example, Olsen 2008; Sarney 2008, 2010; and Springstead 2010, 2011), and by researchers outside SSA (Wu, Karamcheva, Munnell, and Purcell 2013; Reno and Walker 2011). Researchers from the United States General Accounting Office have also used MINT in several analyses (2001, 2004).

MINT in the Context of Other Microsimulation Models

MINT draws on techniques, data, and parameters that other US-based dynamic microsimulation models have used. For example, the targets for MINT's extended cohorts are derived from the POLISIM model, housed in SSA's Office of the Chief Actuary (OACT). Similarly, MINT's job change model, a component of the pension projection module, uses age-centered regression techniques developed by analysts from the Congressional Budget Office (CBO) (Sabelhaus and Walker 2009). MINT's pension model draws on other, related research (Sabelhaus and Brady 2008) and shares a substantial number of functions with the Dynamic Simulation of Income Model (DYNASIM), developed by the Urban Institute. It also uses parameters from a static microsimulation model, the Urban Institute's Transfer Income Model (TRIM). These synergies help to enhance MINT's validity efficiently, given the considerable development costs for such complex, data-intensive, and detailed models.

Why Start the Simulation in 2004 through 2014?

MINT8 starts with data from the 2004 and 2008 SIPP panels matched to administrative data through 2015. It projects the elements in the administrative data after 2015. With some exceptions, projections for items observed in the SIPP start after the last SIPP-based observation, usually in 2005 or 2009 for the respective panels. One philosophy behind MINT's design is that capturing correlations between complex processes is extremely difficult, and using the most recent data with a minimum of imputation helps to ensure validity.

Because so many life history elements are available in the administrative data and the SIPP topical modules, the amount of baseline imputation required in MINT is relatively modest. The CBO long-term dynamic microsimulation model (CBOLT) starts with a 1 percent sample from the Social Security Administration's Continuous Work History Sample (CWHS) with a historic lifetime earnings since 1951. They augment the CWHS data with information from the CPS and SIPP.

MINT2014 is a parallel version of MINT8 that starts with data from the 2014 SIPP panel matched to administrative data through 2015. The Census Bureau substantially changed the method it used to collect SIPP data for the 2014 panel (National Academies of Sciences, Engineering, and Medicine 2017). Unlike earlier panels that interviewed households every four months to collect data from the prior four months, the 2014 SIPP is asked once a year to collect monthly data for the prior calendar year. In prior SIPP panels, the Census Bureau included topical modules along with the core questions with each interview. The 2014 SIPP expanded the set of core questions compared with earlier panels and dispensed with the topical modules. The 2014 core data did not collect information formerly asked on the Marital History, Retirement and Pension Plan Coverage, Work Disability History, and Functional Limitations and Disability – Adults and Children, topical modules. This information was collected in a telephone follow-up survey (SSA supplement) that had only a 57 percent response rate. The MINT2014 processing included imputing the starting data for SSA supplement non-respondents. Extensive validation and comparisons of MINT8 and MINT2014 found no deficiencies in the 2014 SIPP as a starting sample for MINT after imputing SSA supplement data.

Some other dynamic microsimulation models start in a more distant historical period. For example, POLISIM and its predecessor Corsim started in 1980 and 1960, respectively, the Penn Wharton model starts with 1996 Current Population Survey data matched to 2006 public use Summary Earning Record data with imputed earnings beginning in 1951 (Smetters and Gokhale 2019), and DYNASIM3 starts with a baseline sample set in 1992 with imputed historic earnings beginning in 1951. The rationale for these comparatively early start dates is that they enable developers to validate the projections over the historical period.[10] The Urban Institute updated DYNASIM4 in 2017 to start with 2004 and 2008 SIPP panels with imputed lifetime earnings beginning in 1951 (Smith, Favreault, Cosic, and Johnson 2018). Developers can also align outcomes and determine whether any patterns in the alignment factors indicate changing processes or flawed specification.[11] MINT8 uses only minimal alignment in its aging algorithms. Models that start earlier also allow comparisons of past, current, and future time periods. Individuals who did not survive to provide SIPP interviews, either through death or emigration, are not included in the MINT8 or MINT2014 samples.

Starting Sample and Earnings and Marriage History Data

Starting Samples: Base, Extended Cohort Sample, and Immigrant File

The final MINT8 file contains 336,793 observations (individuals) from four separate groups (table 2):

- 48,479 individuals born before 1980 from the 2004 SIPP panel
- 39,214 individuals born before 1980 from the 2008 SIPP panel
- 210,620 individuals born from 1980 to 2068 at age 31 from POLISIM,[12]
- 38,480 new immigrants who arrived in the United States after age 31 or year 2009 from POLISIM.

The final MINT2014 file has 269,137 observations (individuals) from three separate groups (table 2):

- 44,604 individuals born before 1980 from the 2014 SIPP panel,
- 200,987 individuals born from 1980 to 2068 at age 31 from POLISIM, and
- 23,546 new immigrants who arrive in the United States after age 31 or year 2013 from POLISIM

For core MINT8 cohorts, MINT uses data from a key set of topical modules collected in waves one through seven of the SIPP panels. Only individuals with a positive wave seven longitudinal panel weight are included in the sample.[13] Weights are based on SIPP weights, with small adjustments to account for high levels of mortality in the Disability Insurance (DI) population, especially around the time of first receipt. Without such adjustments, MINT might understate the number of short-duration DI beneficiaries.

For 2014 SIPP SSA supplement nonrespondents ages 15 and older, MINT2014 uses a hotdeck method to impute marriage history data and a minimum distance imputation method to impute pension and retirement plan data. Other than the low response rate for the SSA supplement, the 2014 SIPP largely simplifies the initial data set processing because former topical module data is collected at the same time as the core data, eliminating sample attrition and household composition changes that complicate data processing in the earlier SIPP panels. More information on the imputation methods and validation results are included in Smith, Williams, and Murdrazija (2019).

The unit of analysis for MINT is the individual. However, it tracks marital histories for each person on the file. This enables users to calculate both couple and individual income and assets and to use spouses' characteristics in several of the model's aging functions.

Earnings and Marriage/Fertility Histories

Because Social Security benefits depend both on one's own lifetime earnings and on one's marital experience and spouses' lifetime earnings, MINT's starting sample is enriched with earnings and marital histories. Fortunately, administrative records and SIPP topical module data can provide this information.[14] For individuals with a match to the administrative records, we observe OASDI-covered earnings from 1951 through 2015 and total earnings from 1983 through 2015.[15] About 87 percent of respondents matched to earnings records in the 2004 panel and 93 percent matched in the 2008 panel (table 2).[16] For individuals not matched to the administrative earnings data, MINT uses a hotdecking approach to find a similar person and then uses that person's earnings history as a proxy for the nonmatched SIPP respondent's earnings and benefit history to date. The matching variables include age, sex, DI status, SSI beneficiary indicator, self-reported defined contribution (DC) plan status (yes, no, or missing), average monthly earnings, age of immigration, immigrant source region (native, developed, undeveloped), number of years worked in last 10 years, education, race, and class of worker (government, private, or none).

MINT8 marriage and fertility history data are collected in topical module 2. These data include information on the number of times married, the start and end dates for first and last marriages, the number of children ever born, and the birth dates for the first and last children born.[17] MINT uses the core data to fill in birth dates for observed children in the household. It imputes marriage start and end dates for respondents reporting more than two prior marriages.

Aging Modules and Sequence

Given this enriched starting sample, MINT ages the population, projecting key variables using a variety of techniques, including regression models, statistical matching, and rule-based algorithms. Figure 1 shows a simplified representation of MINT's aging sequence (for a more detailed representation, see Urban Institute 2013b, pages 8–9). Tables 3 through 7 provide summary information about MINT's projection algorithms for demographics (table 3); disability, employment, and earnings (table 4); employer-provided pensions (table 5); wealth (table 6); and, finally, other income sources and expenditures (table 7). We next detail these components.[18]

Family Demographics: Marriage, Divorce, and Spouse Characteristics

The projections of marriage and divorce after baseline are among the first that MINT makes. These projections, based on continuous time hazard models, cover an individual's lifetime. These models take into account age and cohort effects, as well as important socioeconomic differentials (e.g., race, education, income) and important duration effects (e.g., how long one has been married or unmarried) in assigning the likelihood of marrying or divorcing. MINT8 includes an LGB indicator that is set to one for all respondents who reported ever living with a same-sex partner in the SIPP. A subset of never-married individuals is set to LGB status of one to reach a target population of 1.8 percent in each cohort (Smith, Rose, and Cosic 2016). Beginning in 2009, MINT allows LGB individuals to marry and divorce using the same marriage and divorce hazard models as for heterosexual individuals. One limitation is that disability is not incorporated as well as it could be in these functions.

Once MINT has projected that an individual gets married, MINT matches him or her to a spouse. This match is made based on age, projected marriage beginning and end dates, education, race/ethnicity, disability status, and lifetime earnings (Smith, Scheuren, and Berk 2002).[19] An important characteristic of MINT's marriage matching algorithms is that an individual can be matched to multiple spouses. This is known as an "open" marriage market.[20] Essentially, the spouse match provides the economic and demographic characteristics for each spouse over the respondent's lifetime.

Family Demographics: Immigration and Emigration

MINT uses immigrants who were sampled in the SIPP data as "donors" to provide starting characteristics for future immigrants. The numbers of immigrants that enter the model each year are determined by the intermediate targets in the Social Security Trustees' Report by age and sex.[21] We rely on donor characteristics up through the age of arrival in the United States, but after arrival, immigrant outcomes such as earnings, marriage, divorce, mortality, and program participation, evolve in ways that reflect changes in the larger society (for example, mortality improvements, reductions in fertility, and shifts in pension coverage). Many aging algorithms include nativity indicators, as tables 3 through 7 indicate.

One simplifying assumption in MINT is that only immigrants are eligible to emigrate. Immigrants enter the model with a propensity to emigrate (an individual specific error), and this propensity declines with their time in the country.[22]

Individual Demographics: Birth and Death

The SIPP contains detailed data on women's fertility histories from the fertility history topical module. The model completes these histories for those who have not reached the end of their childbearing years prior to the MINT baseline using regressions based on marital status and number of children born. Men inherit the fertility history of their spouses in the years they are married, and out-of-wedlock children (based on self-reports) are imputed using estimates from out-of-wedlock births among women. MINT aligns fertility rates to the 2018 Trustees fertility rate assumptions by age and year. Note that the fertility projections in MINT do not project the future sample population. People born after 1980 are selected from POLISIM. Fertility in MINT is used to determine annual number of children for tax and benefit calculations and as predictors in wealth, homeownership, and coresidency projections.

Death is modeled separately on the basis of age, sex, and disability status. At older ages, separate regression functions for men and women compute death probabilities. Explanatory variables include age, education, race, marital status indicators, and various interactions. At younger ages, death is modeled as part of the earnings and disability simulation, as the next section describes. Death rates are roughly calibrated to match the intermediate assumptions of the OASDI Trustees Report on an age-sex-year basis.

Employment Earnings Disability and Death through Age 54 (Age 67 for DI Beneficiaries)

Because of the close correlation between earnings, death, and DI receipt, MINT models these processes jointly in prime age (through age 54 for the nondisabled and age 67 for those who ever participated in DI) using a hotdeck statistical matching algorithm. As table 4 indicates, key variables in the distance function are disabled adult child status (DACS), age, gender, DI benefit status, number of years worked in the last five years, average earnings in last five years, work status in years four and five of match period, lifetime earnings quintile by cohort and sex, education, race/ethnicity, uncovered worker indicator, self-employed indicator, and SSI receipt. Five-year segments of these outcomes are spliced together, with "recipients" receiving data from individuals who were the age one is now turning in later years of the matched SIPP data. This maintains all the year-to-year relationships among these outcomes within the five-year imputation block. Figure 2 illustrates how the splicing algorithms work in a simplified way.[23] Donors for each cohort include respondents from eight preceding cohorts. This pooling allows MINT to smooth over recession periods and capture recent decedents. A complex calibration process involving alternate donors ensures that these projections meet OACT targets by age, sex, and year.[24]

Over the years of developing MINT, it has become clear that these earnings projections are very sensitive to the launch point (the last few years of observed data before the projection algorithm takes over). Because the last few years for which MINT8 had observed data (2015) was in the immediate aftermath of a severe recession,[25] we "derecessionize" the donor data.[26] This is more consistent with the Trustees' assumptions and most other long-range forecasters, which tend to average over business cycles after the relatively near term.

Health Status and Work Limitations

Beyond DI beneficiary status, MINT includes two health indicators at age 51 and older. The first is self-reported health status, modeled from age 51 to death. While many surveys classify health using a five-point scale (excellent, very good, good, fair, poor), MINT uses a two-outcome scale (fair or poor versus other). MINT's second indicator is self-reported work limitations, which is available from ages 51 through 67, with three possible outcomes: no health condition that limits work, condition that limits but does not prevent work, and condition that prevents work. As with other MINT variables, starting values are observed from the SIPP. Because these elements are projected after mortality and DI receipt, survival and DI status are used as predictors in the equations to ensure proper correlation, along with standard demographic information like education and race (table 3 for work limitations and health, table 4 for disability).

Earnings at Age 55 and Older

Because MINT is focused on projecting retirement income adequacy, developers elected to include a very explicit model of the retirement decision, where retirement is defined as a drop in usual weekly work hours to below 20. The retirement model places special attention on retirement income replacement rates, and contains a lot of information about potential retirement resources and family situation (table 4). For those who choose not to retire in a given year, earnings after the splicing part of the model are projected using age-education fixed-effect models. For those who do retire, MINT uses separate regressions to forecast the probability of work and earnings among workers using covariates that reflect health status and work experience. Age variables in the retirement model are specified as age relative to the Social Security full retirement age (FRA). This allows later cohorts to delay retirement as the FRA increases from 65 for individuals born before 1938 to age 67 for individuals born after 1959.

The process of modeling earnings at older ages also closely accounts for beneficiary status, given policies like the Retirement Earnings Test (RET). To model beneficiary status, MINT includes a set of hazard models for Old-Age and Survivors Insurance (OASI) claiming. These models separately consider spouse-only beneficiaries, high earners, and lower earners. Once an individual has elected to claim benefits, a new set of equations projects employment and earnings based on whether one is a first-year claimant and other factors. A final set of equations models employment and earnings at age 70 and older. The MINT8 claiming equations account for life expectancy and increases in the FRA.

Pensions, Including Job Change and Job Characteristics

MINT tries to capture the very complex and rapidly changing US pension climate. The model starts with observed information on pension coverage from SIPP self reports and from Detailed Earnings Records (DER) on contributions to 401(k) and 403(b)-type pensions ("deferred earnings"). The model represents pensions from defined benefit (DB), defined contribution (DC), cash balance (CB), and combination plans.

Individuals with DB pension coverage are matched to detailed plan provisions from the Pension Insurance Modeling System (PIMS) developed by the Pension Benefit Guaranty Corporation (PBGC). This enables computation of their pension benefits once claimed. Those individuals covered by DC plans can make annual contributions using a two-stage process. A first equation predicts the decision to participate given an offer. A second equation predicts the amount one contributes given participation. Explanatory variables in these equations include information about employer contributions and demographic and economic characteristics (table 5). Contribution amounts include an individual-specific error term calculated from the difference in the actual and predicted contributions from the DER data.

Individuals are assigned an individual-specific risk tolerance based on a multinomial logit model estimated with SCF data (Smith et al. 2010, chapter 5). Given an individual's risk tolerance, retirement account balances are allocated to a mix of stocks and bonds. Individuals with higher tolerance for risk invest a larger share of assets in stocks than individuals with lower tolerance, and asset investments shift more to bonds for all individuals with age.

Using target-date fund prevalence from the Employee Benefits Research Institute (EBRI) (Copeland 2011), MINT8 assigns 50 percent of new workers and 10 percent of SIPP baseline workers with a DC pension to select a target-date fund investment. This assignment method imputes higher initial target-date fund participation for younger workers, lower-tenure workers, and workers with lower account balances than for older, longer tenure, and higher balance account holders, consistent with EBRI tabulations (Copeland 2011, figure 1). A rising share of workers will have target-date funds over time, as workers enter the labor market, change jobs, and increasingly have exposure to target-date fund selection.

MINT8 randomly assigns workers to a target date fund based on the dollar-weighted share of the 40 largest target-date funds according to Morningstar (Morningstar 2012, table 3). MINT8 assumes that target-date fund selection is an absorbing state. Once workers are projected to select a target-date fund, they remain target-date investors until retirement. MINT8 reassigns the specific target-date fund at every job change and assumes that all accumulated DC assets are allocated based on the new target-date fund's asset mix (balances accumulated from prior employment follow the new job's target-date fund asset allocation). MINT8 rebalances target- date fund investments and standard investment portfolios annually.

Stock and bond portfolios earn stochastic rates of return centered around the historic mean stock and bond returns through 2019 and projected average returns thereafter (Ibbotson 2017). MINT8 assumes a 6.5 percent real rate of return after 2019. Bond portfolios include 40 percent long-term government bonds and 60 percent corporate bonds. MINT8 assumes a 3.5 percent real rate of return on corporate bonds and a 3 percent real rate of return on government bonds. Actual annual returns include individual-specific stochastic variance of 0.1728 on stocks and 0.0214 on bonds. MINT8 subtracts 1 percent from stock and bond annual returns to reflect administrative costs. Note that projected retirement incomes are sensitive to rate of return assumptions (Smith 2017).

Individuals in MINT can change jobs. When they do, their job characteristics, including health insurance offer and premiums, union status, employment sector, pension coverage, and pension type—DB, DC, CB, or some combination—may also change according to an elaborate model (table 5). MINT workers may also cash out their DC pension accruals rather than rolling them over upon a job loss or job change, reducing their accumulated wealth. Younger workers, those with lower account balances, and those with job losses are more likely to cash out accumulated balances compared to older workers, those with higher account balances, and workers who seamlessly move from one job to another. When married workers claim their pensions, they decide whether to receive a joint and survivor annuity. For workers selecting a survivor annuity, the survivor receives half of the sponsor's pension benefit. All government DB pensions and a share of private-sector pensions are adjusted annually for cost of living increases.

As part of the projection process, MINT8 models continuing evolution in the pension sector. So even if a worker does not change jobs, his or her pension can still change. MINT8 models freeze in DB plans, plan conversions, and the proliferation of target date funds as options for DC plans. MINT8 assumes that between 2008 and 2016, 49 percent of private-sector nonunion DB pensions, 20 percent of private-sector union DB pensions, and 57 percent of state and local pensions will freeze. Among frozen DB plans, MINT8 assumes all state and local plans and 27 percent of private sector plans allow existing participants to continue to accrue benefits (soft freeze), and 73 percent of private sector plans cease accruals for all participants (hard freeze).

MINT8 also projects work expenses in the job change model. It starts with self-reported commute mode and work expenses. Then with each job change, MINT8 projects commute mode (drive or not drive to work). Then given commute mode, MINT8 projects work expenses as a share of total earnings (Smith, Williams, and Murdrazija 2019).

Wealth

As with pensions, MINT8 wealth projections begin with self-reports from the relevant SIPP topical modules. Because of known deficiencies in SIPP's wealth data, MINT8 calibrates the initial wealth distribution to data from the SCF. MINT projects housing wealth separately from nonhousing wealth. The latter, which we sometimes label as financial assets, includes vehicles, other real estate, and farm and business equity; stock, mutual fund, and bond values; checking, saving, money market, and certificate of deposit account balances; and value of other assets, less unsecured debt.

MINT projects wealth in several phases we can characterize simply as *build up* and *spenddown*, though the latter term is somewhat of a misnomer because a share of the elderly continues to accumulate wealth in retirement. The present value of lifetime earnings is a key predictor of wealth, which is modeled separately for unmarried and married people (table 6).

Smith, Michelmore, and Toder (2008) provide an extensive evaluation of wealth projections in an earlier version of MINT. They observe that wealth estimates in the available nationally representative data sources differ significantly from one another, making it challenging to determine precisely how well the model was performing. MINT's asset projections generally align with the target SCF historical series.

One challenge for analysts is how to convert wealth values into income streams. MINT enables the user to choose among several options, depending upon one's interest. For example, one can assume a rate of return on projected wealth. Alternatively, MINT includes two separate annuity factors that allow users to convert assets into income flows. One annuity factor is based on unisex age and cohort mortality rates from the 2018 Trustees assumptions. The second annuity factor varies by age, sex, cohort, education, and race. The annuity factors assume a 50 percent joint and survivor annuity using a 3 percent real return on assets. MINT also imputes taxable interest, dividend, rental income, and capital gains as a function of accumulated assets based on a statistical match to the Statistics of Income (SOI) data (Smith et al. 2007, chapter 5). It also calculates annual taxable withdrawals from retirement account balances including statutory minimum distribution requirements after age 70.

Homeownership, Home Equity, and Mortgage Status

MINT8 homeownership, home equity, and mortgage status begin with self-reports from the relevant SIPP topical modules. It projects annual home purchase hazards for renters and home sale hazards for homeowners. Among homeowners, MINT8 projects home equity and mortgage status (table 6). Home purchase and home sale models are based on logistic regression models estimated using PSID data from 1968–1994 (Toder et al 2002). Home equity models are based on random-effects models estimated on PSID and HRS data (Toder et al 2002;Smith et al 2007). Have-mortgage models are based on random-effects logistic models estimated on 1968–2011 PSID data (Smith, Williams, and Murdrazija 2019).

Transfer Income, Including SSI and Noncash Transfers

To better capture total incomes, and thus poverty status, MINT began to include SSI benefits starting in MINT3. In MINT6, the model started to include other cash transfers, including means-tested sources like Temporary Assistance for Needy Families (TANF) and general assistance and non-means-tested sources like workers compensation, veterans compensation, Unemployment Insurance (UI), state-level Temporary Disability Insurance (TDI), severance payments, employer/union temporary sickness payments, own sickness, accident, and employer disability payments. MINT7 also added noncash transfers, including food assistance like the Supplemental Nutrition Assistance Program (SNAP), Women, Infants, and Children (WIC), and heating and rental assistance to the model.

While these transfer projections are stylized relative to a full-scale static microsimulation model that would model most of the rules for these programs directly, like TRIM, they offer a great improvement over excluding this information. This is especially true when examining the well-being of working-age individuals and DI beneficiaries, given that many of these sources, like TANF, tend to be concentrated outside the retiree population that was MINT's original focus.

The SSI module is the most directly rule based of these functions. The module begins by applying an eligibility screen that mimics SSI law to all individuals in the model who are age- or disability- eligible.[27] Literature shows that not all who are eligible for SSI participate in the program. So MINT uses logistic regression equations to select those most likely to participate in SSI (generally, those who can expect relatively high benefits). We describe the computation of SSI benefit levels below.

The other cash and noncash transfers are all projected using regression-based strategies, rather than program rules. Most use a two-stage process, considering first whether income is present and then, if so, the projected value of the transfer (table 7). The equations use explanatory variables that correlate highly with eligibility criteria for these programs (assets, income, earnings changes, health status, presence of young children, poverty status, etc.).

Living Arrangements and Income of Coresidents

Because moving in with relatives or friends is a common strategy for avoiding poverty, thus an important resource to take into account, MINT models living arrangements and the income of one's coresidents. The model uses separate algorithms for those ages 25 to 61 and 62 and older to model the decision to coreside (table 3). Predictors include the standard from the literature, including number of children, detailed marital history information, impending mortality, other demographics, and SSI receipt—to take into account SSI regulations on in-kind support and maintenance (ISM).

MINT starts with observed SIPP coresidency status. After baseline, it imputes the characteristics of the persons with whom one coresides using a statistical match to historical SIPP family members of coresidents. In each age range, individuals and couples who are selected to coreside are classified into 1 of 16 recipient groups based on marital status, homeowner status, presence of children, and nativity. MINT selects an appropriate donor family among all families in the analogous donor group based on the wage-adjusted per capita income of the respondent.

Expenditures

The federal government's official poverty measure compares family money income to a series of thresholds that were originally computed in the 1960s as multiples of the cost of a nutritionally adequate diet. The thresholds vary by family size and whether the householder is under or over age 65. In recent years, the Census Bureau has developed a supplemental poverty measure (SPM) that takes into account certain kinds of noncash income and certain expenses in addition to food, including taxes, work expenses, child care expenses, child support paid, and out-of-pocket expenses for medical care (Short 2012).

To help compute supplemental poverty, MINT7 added projections of both premium and nonpremium out-of-pocket medical expenditures. MINT first assigns health insurance coverage to one of eight types: employer-provided health insurance; privately purchased nongroup insurance; Medicare with employer coverage; Medicare with a gap policy; Medicare only; Medicaid only; Medicare Medicaid dual; and uninsured. Coverage depends on age, family income, and access to employer benefits (own or spouse). Premiums vary depending on the type of coverage.

For calculating out-of-pocket nonpremium expenses, MINT uses a two-stage process. It first determines whether an individual or family has these expenses and then projects the amount given presence of expenses (table 7). Both stages use information on the individual's type of health insurance coverage. Implementing these equations required building a detailed representation of insurance status, including a representation of how the Affordable Care Act (ACA) is expected to change coverage as it phases in including a state-specific vector to store the year each state expanded (or did not expand) Medicare coverage through the ACA. As with other functions, a wide array of demographic and economic predictors helps to explain these expenditures.

Beginning with MINT8, MINT now projects child care expenses, and child support payments to enhance MINT8 supplemental poverty calculations. Each year, MINT8 projects whether working women with children under age 16 pay child care expenses during work hours. For women with child care expenses, MINT8 projects the child care expense as a share of earnings (Smith, Williams, and Murdrazija 2019). Similarly, it projects whether individuals with a child living in a different household pay child support. For individuals projected to pay child support, MINT projects the amount of child support as a share of earnings (Smith, Williams, and Murdrazija 2019).

OASDI and SSI Benefit Calculators

OASDI: The MINT Social Security benefit calculator was developed by analysts in SSA's ORES. This calculator contains extremely detailed rules for computing benefits, taking into account one's own and one's spouse's lifetime earnings; timing of retirement, disability, and death; and other key benefit determinants.[28] It also includes many nuanced aspects of the benefit calculation, including the RET, Government Pension Offset , and Windfall Elimination Provision.

MINT8 projects current law scheduled Social Security benefits. The long-term projections do not reduce benefits when OASI or DI trust fund balances are depleted, although the strength of MINT is that it allows the user to examine the distribution of income under alternate Social Security policy including payable options (Social Security Administration 2015b).

SSI: MINT's SSI calculator similarly mimics rules and regulations from SSI law. It includes stylized state supplements, with parameters derived from TRIM, which in turn relies on SSA's publication *State Assistance Programs for SSI Recipients.* MINT uses baseline state information from the SIPP (or from the donor SIPP record in the case of postbaseline immigrants and the extended cohorts) in the assignment of state supplements. The model does not project state to state migration.

Income and Payroll Tax Calculators

MINT's income tax calculator is adapted from one Jon Bakija developed. To obtain quantities that MINT does not project, such as capital gains and charitable deductions for itemizers, MINT statistically matches members of the sample to tax units on an Internal Revenue Service SOI public use file, which has information on key fields from the federal income tax form.

MINT8 uses current law federal tax parameters through December 2019 and state tax law through 2013. MINT8 assumes current-law tax assumptions through 2099 including the provisions passed in the 2012 American Taxpayer Relief Act (ATRA) and 2017 Tax Cut and Jobs Act (TCJA). The TCJA modified the tax threshold inflation provisions to increase with the chained Consumer Price Index (CPI) instead of CPI. MINT8 uses the TCJA assumptions through 2099. MINT8 calculates the Medicare surtax beginning in 2013. The Medicare surtax is a 3.8 percent surtax on married filers with investment income above $250,000 and single filers with investment income above $200,000 and a 0.9 percent surtax on earned income above the same thresholds.

MINT makes many assumptions about the personal income tax code to compute tax liabilities into the future. Users should be cognizant of the uncertainty in tax law over extended periods and thus the stylized nature of these assumptions.[29] But the ability to understand how net income will change offsets these limitations for some analyses.

Calculations of payroll taxes are straightforward given earnings, coverage, and the schedule of rates and the taxable maximum.[30] The calculated payroll tax on the final file includes only the worker's share calculated at the family level, but users can compute alternatives that look at individuals and/or include both the employer and employee share.

MINT imputes taxable interest, taxable divided income, rental income, and long-term and other capital gains as part of its income tax calculation. This is done through a statistical match to the 2009 SOI file linked to the 2008 SIPP data. Detailed methodology for the 2001 SOI tax donor file is included in Smith et al 2007. The statistical match minimizes a distance function that includes age, race, tax unit wage and salary and self-employment earnings, log of financial assets, home equity, and Social Security and DB pension income. The match is done separately by age (under 65 and 65+) and filing status (joint filers, single filers, and head of household filers).

Total Income Calculations, Including Poverty and Supplemental Poverty

Once all the MINT projections are complete, users can combine income and expenses to compute alternative measures of economic well-being, including how family incomes relate to the federal poverty level and the SPM. Users can also evaluate how changes to Social Security and SSI benefits influence these measures or how further changes to economic structures and outcomes, like the pension landscape and the stock market, will affect them. With every model simulation, MINT automatically generates a large number of analytic tables that compute total incomes and poverty rates, both at points in time and at various ages (for example, 62 and 67). Many of the MINT products cited in table 1 present these projections.

MINT8 expanded the modeled population by adding individuals born before 1926 that were excluded in earlier versions. MINT8 backcasts earnings before age 31 for the extended cohorts. MINT8 is nearly representative of the US population from 2008 to 2068, butexcludes people who die or emigrate before age 31 and stops adding individuals born after 2068. Population and income values are censored before age 31 and after 2068.

Recommendations for Use, Including Policy Analyses

When using MINT for policy analysis, it is important to adhere to guidelines used with other survey data. For example, projections for smaller population or beneficiary groups are less reliable than those for larger ones, so users will want to take care about drawing inferences from a small number of unweighted cases. Similarly, users may want to examine percentiles rather than (or in addition to) means when examining quantities that are highly skewed in the United States and thus in MINT.[31] Earnings, wealth, asset income, and medical expenses are the most prominent examples. Given the enormous number of assumptions incorporated into MINT's long-range projections, users may wish to avoid the appearance of excessive precision when presenting certain estimates, especially when using nominal dollars. For example, analysts might consider rounding certain quantities in situations when one is not examining changes. Such situations would include characterizations of the future wealth distribution.

Because MINT projects all the way through, and even beyond, the 75-year projection horizon employed for the Trustees Report, users may wish to avoid focusing on a single point in time when examining a proposal to change Social Security that has a time path that changes greatly over time (for example, benefit reductions are highly backloaded). Similarly, users may wish to avoid focusing exclusively on very distant projection years when this is not necessary (for example, if a proposal phases in and effects stabilize relatively quickly), given that MINT's great strength is its observed data on real individuals. New retirees in 2070, for example, are just children today, so their future in MINT is entirely simulated. With long-run projections, the issue of whether to present estimates in nominal, price-indexed, or wage-indexed terms also arises. One solution is to present multiple estimates, or when presenting just one, to explain how using other measures does or does not change the story (Butrica, Iams, and Smith 2003).

Another issue is how to compare alternative changes to the OASDI program that achieve different cost savings or rely on a different mix of costs savings between the payroll tax and benefit sides of the program.[32] Lifetime measures of benefits and taxes, and measures that relate these quantities to one another, can help display how changes affect both beneficiaries and taxpayers.

Behavioral Responses

An occasionally controversial aspect of using a model like MINT is determining how to account for changes in behavior that might result from changes to policy parameters (say, the payroll tax rate or the early retirement age). Researchers differ in estimates and expectations of how large responses are likely to be. Historical data from which to estimate potential effects and analytic techniques may be insufficient; there are often few natural experiments for policy changes of significant magnitude under analogous economic and demographic circumstances for retirees and the disabled.

MINT has a limited capacity to account for policy changes that are captured in the estimated model parameters. For example, an increase in the full retirement age will have some modest effects on earnings and OASDI claiming.[33] Users uncomfortable with the default assumptions can impose an alternative response that is reasonable given expert judgment. In such circumstances, it is often helpful to test sensitivity of outcomes over a range of options (i.e., best guess, high, and low) derived from the best literature on the policy or process. Larger changes to Social Security, taxes, or other benefits are likely to lead to more sizable behavioral responses, and thus more likely to call for sensitivity analyses, than more modest changes.

User Tools

Because many analysts use MINT, we have developed a number of tools that permit users to easily compare outcomes across model runs, cohorts, time, and other characteristics. Similarly, MINT's highly parameterized source code facilitates developing and processing simulations with alternative assumptions about Social Security law or core processes (like mortality or wage growth).

MINT8 includes detailed validation tools that compare MINT8 and MINT2014 projections with numerous tables included in the 2017 Annual Statistical Supplement to the Social Security Bulletin (Social Security Administration 2017), published Internal Revenue Service tables (Internal Revenue Service 2018a; 2018b), and published OASDI Board of Trustees (2018) historic and projected tables. The validation tools also compare MINT8 and MINT2014 with tabulated demographic and income data from the annual Current Population Surveys and annual American Community Surveys. The result of this validation shows that MINT8 projections align closely with administrative income, tax, and benefit data and with demographic and income tabulations from the survey data. In cases where we know the survey data has shown to be deficient (Bee and Mitchell 2017), MINT more closely resembles the administrative data than the survey data. But general patterns by sex, age, race, education, and nativity in MINT follow patterns in the survey data.

Conclusions

MINT is a large, complex model that has been under development since 1999 and is now used extensively by SSA analysts. We have tried to provide readers with a brief overview of the model. We recommend that users needing more detailed information consult the documentation identified in tables 1 and 3 through 7.

Note

[1] Throughout this report, we use the terms OASDI and Social Security interchangeably. When we wish to refer to a subset of Social Security, like Old-Age and Survivors Insurance (OASI) or Disability Insurance (DI), we do so explicitly.

[2] Advisors from outside SSA have included Christopher Bone, Richard Burkhauser, Alan Gustman, Mark Hayward, Kathleen McGarry, Olivia Mitchell, John Rust, John Sabelhaus, and Finis Welch. While these advisors provided valuable advice, they are not responsible for ultimate choices about model specification.

[3] For dataset documentation, see Smith, Williams, and Murdrazija (2019a).

[4] The administrative files MINT uses include the following:

- -The Detailed Earnings Record (DER) and Summary Earnings Record (SER), which provide information on earnings, Social Security coverage, and contributions to deferred earnings plans, like 401(k) and 403(b) plans, plus limited demographic information, like date of birth;
- -The Supplemental Security Record (SSR), which provides information on SSI receipt;
- -The Master Beneficiary Record (MBR), which provides information on timing, level, and type of benefits received from the Old-Age, Survivors, and Disability Insurance program (OASDI); and
- -The Numident, which provides information on date of death and place of birth and year of entry plus legal status at entry for the foreign born.

[5] While developing MINT, we have experimented with different ages for donors for the extended cohorts. The advantage to using an earlier age is that MINT can readily forecast DI receipt at younger ages. The advantage to using a later age is that formal education is less likely to be incomplete (be "right censored" in statistical terms), an important limitation given education's central role in predicting earnings and given that MINT does not currently model education.

[6] While we use SIPP for estimating parameters in many MINT functions, in some cases the SIPP data are too limited to enable estimation. For example, the relatively short SIPP panels may not be adequate for estimating random effects models, and information about defined benefit pension accruals is insufficient for developing a detailed retirement decision model. Some interesting outcome data are only available in a single topical module of the SIPP, making it impossible to estimate transitions or other types of dynamics.

[7] While the Trustees' assumptions inform MINT employment rates, the model does not directly calibrate to them.

[8] The group historically has not provided cost analyses. The Office of the Chief Actuary provides such estimates on its web site, and the Congressional Budget Office periodically releases similar projections (CBO 2012).

[9] These distributional analyses use a wide range of cross-sectional outcome measures, including medians of both individual Social Security benefits and household total income. MINT tables present the percentages of beneficiaries changes affect, while special tabulations provide projections of the sizes of changes among those a policy change affects. Analysts often juxtapose the cross-sectional simulation results at three points in time: 2030, 2050, and 2070. Various tables classify individuals by gender, education, household income, lifetime earnings, payable lifetime earnings, marital status, race, and type of benefit (retired worker, survivor, disabled worker, spouse, dually entitled survivor, dually entitled spouse). Some MINT analyses include projections of poverty status.

[10] To some degree, they also reflect aging of the models themselves. For example, DYNASIM3's starting sample was among the most recent available at the time DYNASIM3 was developed. Irregular updating may reflect responses to developer priorities and resource constraints.

[11] For more information about alignment in dynamic microsimulation models, see, for example, Klevmarken (1998) and Neufeld (2000).

[12] We elected to use POLISIM because the model is calibrated to intermediate assumptions from the Social Security Trustees Report (OASDI Board of Trustees 2018) and has information about individuals' education, nativity (and

region of origin for the foreign born), race, and marital status. We use these characteristics to select an appropriate mix of records from the SIPP sample for the later cohorts. For information about an earlier version of POLISIM, see Favreault and Smith (2007).

[13] MINT uses data from the following topical modules: employment history; marital history; fertility history; migration history; disability history; health conditions and work limitations; retirement and pension plan coverage; assets and liabilities; annual income and retirement accounts; and employer-provided health benefits.

[14] MINT ignores administrative data when the difference in SIPP self-reported birth year and administrative birth year is more than five years. MINT treats these records as nonmatched cases and imputes the administrative data values for these respondents.

[15] While total earnings are available from as early as 1978, the data are not high quality until about 1983.

[16] Table 2 also reports match rates to Numident, which are comparable to the SER match rates, and to benefit records for OASDI and SSI (the MBR and SSR, respectively). The rates for the benefit records are much lower than for the earnings and mortality records because they may not be established until one makes a claim for benefits. Several researchers have examined match representativeness for SIPP (for example, Davis and Mazumder 2011).

[17] Since the 2001 panel, some of these data fields have been restricted, but SSA has obtained the required permissions to use these detailed data.

[18] As we detail projection algorithms, we will highlight a few interesting aspects of each. We cannot detail all completely. Please see the tables and referenced documents for more complete descriptions.

[19] Match weights are not empirically derived, but rather assumed.

[20] A closed marriage market is one in which there is exactly one spouse for every married person in the simulation. If there are insufficient numbers of partners for the sample of individuals selected to marry in a given year, then the person does not marry that year, but can instead re-enter the marriage pool in subsequent years.

[21] MINT uses 2018 Trustees historic and projected number of gross legal and other than legal number of immigrants that enter the United States each year. MINT8 uses the emigration hazard from Dowhan and Duleep (2002) to impute emigration by age at entry and source region. It uses US Department of Homeland Security Yearbook data (2016, 2012a, 2012b) to impute source region and legal status to the target population.

[22] MINT immigrants leave as individuals, rather than as family units.

[23] In reality, the algorithm is more complex, searching throughout the donor file for an exact match on these variables, and then incrementally relaxing the match criteria until a satisfactory match is found. The match imputes earnings, disability status, and mortality information.

[24] The technique MINT8's statistical match uses to perform the calibration relies on assigning several potential donors for each recipient and performing multiple rounds of projections. If a Trustees' target is not reached in the first round of projections, then the program swaps donors to align the projections to the target disability prevalence and mortality rates (Toder et al. 2002). For example, if DI prevalence is too low, the program loops through the individuals who were not selected to become disabled and swaps a requisite number of those who have an alternate donor who is disabled to disabled status. Sometimes this requires multiple rounds, given that individuals in MINT do not have equal weights, so swapping an individual with a high weight can lead MINT to overshoot or undershoot the desired disability rate for a given age-sex group.

[25] The National Bureau of Economic Research dates the great recession as having lasted from December 2007 through June 2009. High unemployment often endures long after the recession officially ends.

[26] The derecession process leaves the historic earnings record unaltered but replaces the donor record with an updated future that has higher employment rates. Specifically, we count up the number of years between 2006 and 2010 with zero earnings each donor record has. For 80 percent of cases with zero years, we make changes as follows: for donors with one year of zero earnings, we replace the zero year with the average of the four nonzero years. For donors with two years of zero earnings, we replace a randomly selected zero year with the average of the three nonzero earnings.

For donors with three years of zero earnings, we replace a selected zero year with the average of the two nonzero earnings. We made no changes to donor earnings for disabled donors and donors ages 60 and older.

[27] Part of this screen requires computing potential Social Security benefits for eligible individuals not currently receiving OASDI benefits. SSI's status as "program of last resort" implies that individuals must apply for all other forms of support for which they are eligible, including Social Security, before receiving SSI.

[28] MINT tracks multiple spouses, enabling the calculator to compare benefits for those with multiple entitlements (e.g., because a previous marriage ended in divorce).

[29] We hold the Social Security taxation thresholds at their current law values, because Congress chose intentionally not to index them for inflation when enacting these rules in 1983 and 1993 in order to increase over time the share of Social Security benefits subject to tax. Similarly, the Medicare surtax thresholds are not indexed for inflation. After 2019, we index exemptions and bracket widths of both the regular income tax and the alternative minimum tax using chained CPI.

[30] The Trustees' file, used in many important MINT calculations, includes the historical and projected OASI, DI, and Hospital Insurance tax rates.

[31] Alternative or supplemental approaches include examining means excluding the top cases—for example, the top 5 percent of asset income holders—or calculating means by income percentile.

[32] Favreault and Steuerle (2012), for example, compare alternative counterfactuals.

[33] Typically, changes to Social Security benefits are modeled as a postprocess without rerunning the model or changing any behaviors.

References

Bee, Adam, and Joshua Mitchell. 2017. "Do Older Americans Have More Income Than We Think?" SESHD Working Paper #2017–39. Suitland, Maryland: Census Bureau.

Butrica, Barbara A., and Howard M. Iams. 2000. *Divorced Women at Retirement: Projections of Economic Well-Being in the Near Future*. Social Security Bulletin 63(3): 3–12. Washington, DC: Social Security Administration.

Butrica, Barbara A., and Howard M. Iams. 1999. *Projecting Retirement Income of Future Retirees with Panel Data: Results from the Modeling Income in the Near Term (MINT) Project.* Social Security Bulletin 62(4): 3–8. Washington, DC: Social Security Administration.

Butrica, Barbara A., Howard M. Iams, and Steven H. Sandell. 1999. *Using Data on Couples to Project the Distributional Effects of Social Security Policy Changes.* Social Security Bulletin 62(3): 20–27. Washington, DC: Social Security Administration.

Butrica, Barbara A., Howard M. Iams, and Karen E. Smith. 2003. "It's All Relative: Understanding the Retirement Prospects of Baby-Boomers." Working Paper 2003–21. Chestnut Hill, MA: Center for Retirement Research at Boston College.

Butrica, Barbara A., Howard M. Iams, and Karen E. Smith. 2004. "The Changing Impact of Social Security on Retirement Income in the United States." Social Security Bulletin 65(3): 1–13. Washington, DC: Social Security Administration.

Butrica, Barbara A., Howard M. Iams, and Karen E. Smith. 2007. "Understanding Baby Boomers' Retirement Prospects." In *Redefining Retirement: How Will Boomers Fare?* edited by Brigitte Madrian, Olivia S. Mitchell and Beth J. Soldo. New York: Oxford University Press.

Butrica, Barbara A., Howard M. Iams, and Karen E. Smith. 2013. "Poverty Status of Older Americans: How You Measure It Matters." Washington, DC: Urban Institute.

Butrica, Barbara A., Howard M. Iams, Karen E. Smith, and Eric J. Toder. 2009. T*he Disappearing Defined Benefit Pension and Its Potential Impact on the Retirement Incomes of Boomers.* Social Security Bulletin 69(3): 1–27. Washington, DC: Social Security Administration.

Butrica, Barbara A., and Karen E. Smith. 2012a. *The Impact of Changes in Couples' Earnings on Married Women's Social Security Benefits.* Social Security Bulletin 72(1): 1–10. Washington, DC: Social Security Administration.

Butrica, Barbara A., and Karen E. Smith. 2012b. *Racial and Ethnic Differences in the Retirement Prospects of Divorced Women in the Baby Boom and Generation X Cohorts.* Social Security Bulletin 72(1): 23–22. Washington, DC: Social Security Administration.

Butrica, Barbara A., and Karen E. Smith. 2012c. "The Retirement Prospects of Divorced Women." Social Security Bulletin 72(1): 11–22. Washington, DC: Social Security Administration.

Butrica, Barbara A., Karen E. Smith, and Howard M. Iams. 2012. *This is Not Your Parent's Retirement*. Social Security Bulletin 72(1): 37–56. Washington, DC: Social Security Administration.

Butrica, Barbara A., Karen E. Smith, and Eric Toder. 2002. "Projecting Poverty Rates in 2020 for the 62 and Older Population: What Changes Can We Expect and Why?" Working Paper 2002–03. Chestnut Hill, MA: Center for Retirement Research at Boston College.

Butrica, Barbara A., Karen E. Smith, and Eric Toder. 2010. "What the 2008 Stock Market Crash Means for Retirement Security." *Journal of Aging & Social Policy* 22(4): 339–59.

Congressional Budget Office (CBO). 2012. *The 2012 Long-Term Projections for Social Security: Additional Information. October*. Washington, DC: CBO.

Congressional Budget Office (CBO). 2018. An *Overview of CBOLT: The Congressional Budget Office Long-Term Model*. Washington, DC: CBO.

Copeland, Craig. 2011. “Target-Date Fund Use in 401(k) Plans and the Persistence of Their Use, 2007–2009.” Washington, DC: Employee Benefits Research Institute.

Davis, Jonathan and Bhashkar Mazumder. 2011. “An Analysis of Sample Selection and the Reliability of Using Short-term Earnings Averages in SIPP-SSA Matched Data.” Working Paper No. CES-WP 11–39. Washington, DC: US Census Bureau Center for Economic Studies.

Dowhan, Dan and Harriet Duleep. 2002. *Projecting Immigrants for MINT*. Social Security Bulletin Vol. 68(1). Washington, DC: Social Security Administration.

Favreault, Melissa M., and Owen G. Haaga. 2013. “Validating Longitudinal Earnings in Dynamic Microsimulation Models: The Role of Outliers.” Working Paper 2013–19. Chestnut Hill, MA: Center for Retirement Research at Boston College.

Favreault, Melissa M., and Karen E. Smith. 2007. A Guide to the POLISIM Model. Washington, DC: Urban Institute.

Favreault, Melissa M., and Karen E. Smith. 2013. “Projections of the Characteristics of Disabled Workers.” Under review, Social Security Administration.

Favreault, Melissa M., and C. Eugene Steuerle. 2012. “Measuring Social Security Proposals by More than Solvency: Impacts on Poverty, Progressivity, Horizontal Equity, and Work Incentives.” Working Paper 2012–15. Chestnut Hill, MA: Center for Retirement Research at Boston College.

Favreault, Melissa M., and Douglas A. Wolf. 2004. “Living Arrangements and Supplemental Security Income Receipt among the Aged.” Working Paper 2004-03. Chestnut Hill, MA: Center for Retirement Research at Boston College.

Iams, Howard M., Gayle L. Reznik, and Christopher R. Tamborini. 2009. *Earnings Sharing in Social Security: Projected Impacts of Alternative Proposals Using the MINT Model*. Social Security Bulletin 69(1): 1–17. Washington, DC: Social Security Administration.

Iams, Howard M., Gayle L. Reznik, and Christopher R. Tamborini. 2010. “Earnings Sharing in the U.S. Social Security System: A Microsimulation Analysis of Future Female Retirees.” *The Gerontologist* 50(4): 495–508.

Iams, Howard M. 2016. *Married Women’s Projected Retirement Benefits: An Update.* Social Security Bulletin 76 (2). Washington, DC: Social Security Administration.

Iams, Howard M., and Steven H. Sandell. 1997. *Projecting Social Security Earnings: Past is Prologue. Social Security Bulletin* 60(2): 3–16. Washington, DC: Social Security Administration.

Ibbotson, Roger, Roger Grabowski, James Harrington, and Carla Nunes. 2017. *2017 Stocks, Bonds, Bills, and Inflation (SBBI) Yearbook.* Chicago IL: Morningstar Inc.

Internal Revenue Service. 2018a. “SOI Tax Stats - Individual Information Return Form W-2 Statistics.” Washington, DC.: Internal Revenue Service.

Internal Revenue Service. 2018b. *SOI Tax Stats - Individual Income Tax Returns Publication 1304 (Complete Report).* Washington, DC.: Internal Revenue Service.

Kaiser Family Foundation. 2012. *Employer Health Benefits, 2012*. Washington, DC: Kaiser Family Foundation and Health Research and Education Trust.

Klevmarken, N. Anders. 1998. “Statistical Inference in Micro Simulation Models: Incorporating External Information.” Uppsala, Sweden: Uppsala University Department of Economics Working Paper Series.

Knoll, Melissa, and Anya Olsen. 2014. *Incentivizing Delayed Claiming of Social Security Retirement Benefits Before Reaching the Full Retirement Age*. Social Security Bulletin 74 (No. 4). Washington, DC: Social Security Administration.

Moore, James H. Jr., and Leslie A. Muller. 2002. “An Analysis of Lump-Sum Pension Distribution Recipients.” Monthly Labor Review May: 29–46. Washington, DC: Office of Policy, Division of Policy Evaluation at the Social Security Administration.

Morningstar. 2012. “Target-Date Series Research Paper: 2012 Industry Survey.” Chicago, IL: Morningstar, Inc.

National Academies of Sciences, Engineering, and Medicine. 2017. *The 2014 Redesign of the Survey of Income and Program Participation: An Assessment*. Washington, DC: The National Academies Press.

National Commission on Fiscal Responsibility and Reform. 2010. *The Moment of Truth: Report of the National Commission on Fiscal Responsibility and Reform*. Washington, DC: The White House.

Neufeld, Christopher. 2000. "Alignment and Variance Reduction in DYNACAN." In *Microsimulation in Government Policy and Forecasting*, edited by A. Gupta and V. Kapur. Amsterdam: Elsevier.

OASDI Board of Trustees. 2012. *2012 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds*. Washington, DC.

OASDI Board of Trustees. 2018. *2018 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds*. Washington, DC.

Olsen, Anya, and Kathleen Romig. 2013. *Modeling Behavioral Responses to Removing the Retirement Earnings Test*. Social Security Bulletin 73(1): 39–58. Washington, DC: Social Security Administration.

Olsen, Anya. 2008. "Distributional Effects of Reducing the Cost-of-Living Adjustments." Washington, DC: Social Security Administration.

Olsen, Anya. 2012. *Mind the Gap: The Distributional Effects of Raising the Early Eligibility Age and Full Retirement Age.* Social Security Bulletin 72 (No. 4). Washington, DC: Social Security Administration.

Orcutt, Guy H. 1957. "A New Type of Socio-Economic System." *Review of Economics and Statistics* 39:116–23.

Panis, Constantijn, and Lee Lillard. 1999. *Near Term Model Development. Final Report, SSA Contract No. 600-96-27335.* Santa Monica, CA: RAND.

Purcell, Patrick, and David Shoffner. 2004. "The Effects of Alternative Demographic and Economic Assumptions on MINT Simulations: A Sensitivity Analysis." Research and Statistics Note No. 2014-03. Washington, DC: Social Security Administration.

Purcell, Patrick, Howard M. Iams, and David Shoffner. 2015. *Education, Earnings Inequality, and Future Social Security Benefits: A Microsimulation Analysis*. Social Security Bulletin 75 (2). Washington, DC: Social Security Administration

Purcell, Patrick. 2015. "Income Taxes on Social Security Benefits." *Social Security Issue Paper*, no. 2015-02 (December). https://www.ssa.gov/policy/docs/issuepapers/ip2015-02.html.

Rendall, Michael. 2005. "POLISIM Equation Documentation." Washington, DC: Urban Institute.

Rendall, Michael. 2003. POLISIM 4.1 Equation Appendix. Ithaca, NY: Strategic Forecasting.

Reno, Virginia P., and Elisa A. Walker. 2011. *How Would Seniors Fare—by Age, Gender, Race and Ethnicity, and Income—Under the Bowles-Simpson Social Security Proposals by 2070?* Washington, DC: National Academy of Social Insurance.

Reznik, Gayle L., David A. Weaver, and Andrew G. Biggs. 2009. "Social Security and Marginal Returns to Work Near Retirement." Issue Paper No. 2009–02. Washington, DC: Social Security Administration, Office of Retirement and Disability Policy.

Romig, Kathleen, and Kevin Whitman. 2016. "Distributional Effects of Applying Social Security Taxes to Employer-Sponsored Health Insurance Premiums." *Social Security Policy Brief*, no. 2016–01 (August).

Sabelhaus, John, and Peter Brady. 2008. "*The Role of Defined Contribution Accounts in Future Retirement Security: A Dynamic Microsimulation Approach*." Washington, DC: Investment Company Institute. [Unpublished manuscript.]

Sabelhaus, John, and Lina Walker. 2009. "Econometric Flexibility in Microsimulation: An Age-Centered Approach." *International Journal of Microsimulation* 2(2): 1–14.

Sarney, Mark A. 2008. "Distributional Effects of Increasing the Benefit Computation Period." Policy Brief 2008–02. Washington, DC: Social Security Administration.

Sarney, Mark A. 2010. "Distributional Effects of Price Indexing Social Security Benefits." Social Security Administration Policy Brief No. 2010–03, November. Washington, DC: Social Security Administration.

Sarney, Mark A. 2017. "A Field Guide to Social Security Distributional Analysis." *Research and Statistics*, Note No. 2017–02 (December). Washington, DC: Social Security Administration.

Shoffner, Dave. 2010. "Distributional Effects of Raising the Social Security Payroll Tax." Social Security Administration Policy Brief No. 2010–01 (released April). Washington, DC: Social Security Administration.

Short, Kathleen. 2012. *The Research Supplemental Poverty Measure: 2011. Current Population Report P60–244*. Washington, DC: US Census Bureau.

Smetters, Kent, and Jagadeesh Gokhale. 2019. "Penn Wharton Budget Model." Philadelphia, PA: Penn Wharton University of Pennsylvania.

Smith, Karen E., David Cashin, and Melissa Favreault. 2005. "Modeling Income in the Near Term 4." Washington, DC: Urban Institute.

Smith, Karen E., Melissa M. Favreault, Caroline Ratcliffe, Barbara Butrica, Eric Toder, and Jon Bakija. 2007. *Modeling Income in the Near Term 5*. Washington, DC: Urban Institute.

Smith, Karen, Katherine Michelmore, and Eric Toder. 2008. *Comparisons of MINT 2003 and 2004 Projections with Survey Data*. Washington, DC: Urban Institute.

Smith Karen E., Melissa Favreault, Barbara Butrica, and Philip Issa. 2010. *Modeling Income in the Near Term 6*. Washington, DC: Urban Institute.

Smith, Karen E. and Melissa M. Favreault. 2013. *Modeling Income in the Near Term 7*. Washington, DC: Urban Institute.

Smith. Karen E. 2017. *Analyzing Participant Behavior in DC Plans: Employee Contribution Levels and the Timing of Retirement.* Washington, DC. Urban Institute. Project report to Department of Labor.

Smith, Karen E., Stephen Rose, Damir Cosic. 2016. *How Might Legal Recognition of Same-Sex Marriage Affect Retirement Incomes and Federal Programs?* Washington, DC: Urban Institute. Project report for the Department of Labor.

Smith, Karen E, Melissa Favreault, Damir Cosic, and Richard Johnson. 2018. *DYNASIM4 Validation Report: Comparing DYNASIM4 Outcomes with Administrative Data, Government Forecasts, and Scholarly Literature*. Washington DC. Urban Institute, project report for Department of Labor.

Smith, Karen E., Aaron Williams, and Stipica Murdrazija. 2019. *Modeling Income in the Near Term 8: Final Report.* Washington, DC: Urban Institute.

Smith, Karen E. and Aaron Williams. 2019. *Modeling Income in the Near Term 8: Dataset Documentation*. Washington, DC: Urban Institute.

Smith, Karen E., Fritz Scheuren, and Jillian Berk. 2002. "Adding Historical Earnings to the Survey of Income and Program Participation (SIPP)." Proceedings of the 2001 Joint Statistical Meetings [CD-ROM].

Social Security Administration. 2011. "Policy Option Projections." Washington, DC: Social Security Administration Office of Retirement Policy.

Social Security Administration. 2015a. "Population Projections." Washington, DC: Social Security Administration Office of Retirement Policy.

Social Security Administration. 2015b. "Program Explainers: Scheduled vs. Payable Benefits." Washington, DC: Social Security Administration. https://www.ssa.gov/policy/docs/program-explainers/scheduled-payable.html.

Social Security Administration. 2017. "Annual Statistical Supplement to the Social Security Bulletin, 2016." 12–11700. Washington, DC: Social Security Administration.

Springstead, Glenn R. 2010. "Distributional Effects of Reducing the Social Security Benefit Formula." Policy Brief 2010–02. Washington, DC: Social Security Administration.

Springstead, Glenn R. 2011. "Distributional Effects of Accelerating and Extending the Increase in the Full Retirement Age." Policy Brief 2011-01. Washington, DC: Social Security Administration.

Springstead, Glenn, Kevin Whitman, and David Shoffner. 2014. “Proposed Revisions to the Special Minimum Benefit for Low Lifetime Earners.” Social Security Policy Brief, no. No. 2014-01 (September). Washington, DC: Social Security Administration.

Tamborini, Christopher R. 2008. *The Never-Married in Old Age: Projections and Concerns for the Near Future*. Social Security Bulletin 67(2): 25-40. Washington, DC: Social Security Administration.

Tamborini, Christopher R., and Kevin Whitman. 2010. “Lowering Social Security’s Duration-of-Marriage Requirement: Distributional Effects for Future Female Retirees.” *Journal of Women & Aging* 22(3): 184–203.

Toder, Eric, Cori Uccello, John O’Hare, Melissa Favreault, Caroline Ratcliffe, Karen Smith, Gary Burtless, and Barry Bosworth. 1999. *Modeling Income in the Near Term– Projections of Retirement Income through 2020 for the 1931–60 Birth Cohorts.* Washington, DC: Urban Institute.

Toder, Eric, Larry Thompson, Melissa Favreault, Richard Johnson, Kevin Perese, Caroline Ratcliffe, Karen Smith, Cori Uccello, Timothy Waidmann, Jillian Berk, Romina Woldemariam, Gary Burtless, Claudia Sahm, and Douglas Wolf. 2002. *Modeling Income in the Near Term–Revised Projections of Retirement Income through 2020 for the 1931–1960 Birth Cohorts.* Washington, DC: Urban Institute.

United States Department of Homeland Security. 2016. *Yearbook of Immigration Statistics: 2015*. Washington, D.C: U.S. Department of Homeland Security, Office of Immigration Statistics.

United States Department of Homeland Security. 2012a. *2011 Yearbook of Immigration Statistics.* Alexandria, VA: US Department of Homeland Security, Office of Immigration Statistics.

United States Department of Homeland Security. 2012b. *Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2011*. Alexandria, VA: US Department of Homeland Security, Office of Immigration Statistics.

US General Accounting Office (GAO). 2001. *Social Security: Program’s Role in Helping Ensure Income Adequacy*. GAO–02–62. Washington, DC: GAO.

US General Accounting Office (GAO). 2004. *Social Security: Distribution of Benefits and Taxes Relative to Earnings Level.* GAO–04–747. Washington, DC: GAO.

Urban Institute. 2013a. “Modeling Income in the Near Term, Version 7.2, Dataset Documentation.” Dated July 12. Washington, DC: Urban Institute.

Urban Institute. 2013b. “Modeling Income in the Near Term, Version 7.2, Source Code Documentation.” Dated July 12. Washington, DC: Urban Institute.

Whitman, Kevin. 2009. “Distributional Effects of Raising the Social Security Taxable Maximum.” *Social Security Policy Brief*, no. 2009-01 (July). Washington, DC: Social Security Administration.

Whitman, Kevin, Gayle Reznik, and David Shoffner. 2011. *Who Never Receives Social Security Benefits?* Social Security Bulletin 71 (2). Washington, DC: Social Security Administration.

Whitman, Kevin, and David Shoffner. 2013. “The Projected Effects of Social Security Benefit Increase Options for Older Beneficiaries.” *Social Security Policy Brief*, no. 2013–01 (October). Washington, DC: Social Security Administration.

Wu, April Yanyuan, Nadia Karamcheva, Alicia H Munnell, and Patrick Purcell. 2013. *How Do Trends in Women’s Labor Force Activity and Marriage Patterns Affect Social Security Replacement Rates?* Social Security Bulletin 73 (4).

Tables and Figures

TABLE 1

MINT History

Version	SIPP data used	Birth cohorts	Last year of admin data used	Trustees' assump-tions	Innovations/updates from prior version (beyond data and trustees' assumption updates)	Projection horizon	Technical documents	Selected related publications
MINT1	1990–93	Core: 1931–60	1996	N/A (see insert to report)	N/A	2020 focus, but extends to 2027a	Toder et al. (1999); Panis and Lillard (1999)	Butrica and Iams (1999, 2000); Butrica, Iams, and Sandell (1999)
MINT3	1990–93	Core: 1931–60	SER/ MBR:1999 SSR: 1997 Numident: 1998	2001; subsequent update to 2004	Added iterative (dynamically recursive), year-by-year processing, including retirement model (based on premium value); added work limitations, health status, living arrangements, and SSI.	2020 focus, but extends to 2027a	Toder et al. (2002)	Butrica, Iams, and Smith (2003, 2004, 2007); Butrica, Smith, and Toder (2002); Favreault and Wolf (2004)
MINT4	1996	Core: 1926–72; extended: 1973–2017	SER:2000 SSR:1998 Numident/ MBR:2002	2004	Added DER data and modified many functions to accommodate uncapped earnings; separate self-employment from wage and salary earnings.	2099	Smith et al. (2005)	N/A
MINT5	1990–93, 1996	Core: 1926–75; extended: 1976–2018	2004	2006; subsequent update to 2008	Added fertility history. Substantially revised many algorithms. Estimated poverty measure.	2099	Smith et al. (2007)	Butrica et al. (2009); Smith, Michelmore, and Toder (2008).
MINT6	2001, 2004	Core: 1926–75; extended: 1976–2070	2007 updated to 2009	2009; subsequent updates to 2010 and 2011	Added younger disabled workers further into simulation period; added transfer income (means-tested and non-means tested); replaced dated job change model for pensions with SIPP-based models; developed more sophisticated ways to treat extended cohorts (e.g., directly model pensions rather than assign from donor file); modeled immigrants more directly.	2099	Smith et al. (2010)	Butrica and Smith (2012a, 2012b, 2012c); Butrica, Smith, and Iams (2012)

Version	SIPP data used	Birth cohorts	Last year of admin data used	Trustees' assump-tions	Innovations/updates from prior version (beyond data and trustees' assumption updates)	Projection horizon	Technical documents	Selected related publications
MINT7	2004, 2008	Core: 1926–79; extended: 1980–2068	2010	2012	Added health insurance coverage and out-of-pocket medical expenditures; added non-cash transfers. Estimated supplemental poverty measure.	2099	Smith and Favreault (2013); Urban Institute (2013a,b)	Butrica, Iams, and Smith (2013); Favreault and Haaga (2013); Favreault and Smith (2013)
MINT8	2004, 2008,	Core: 1905–79; extended 1980–2068	2015	2018	Added projections of work expenses, child care expenses, child care payments, and homeowner with a mortgage to enhance projections of supplemental poverty; updated estimates of Social Security claiming and retirement; updated Trustees assumptions; expanded validation suite.	2099	Smith and Williams (2019)	Smith, Williams, and Murdrazija (2019)
MINT2014	2014	Core: 1905–83; extended 1984–2068	2015	2018	Added projections of work expenses, child care expenses, child care payments, and homeowner with a mortgage to enhance projections of supplemental poverty; updated estimates of Social Security claiming and retirement; updated Trustees assumptions; expanded validation suite.	2099	Smith and Williams (2019)	Smith, Williams, and Murdrazija (2019)

Notes: Additional policy simulations using MINT are described on page 11.

a. 2027 is year 1960 cohort turns 67.

TABLE 2

MINT8 and MINT2014 Starting Sample

	2004 SIPP	2008 SIPP	2014 SIPP
Number of observations	48,479	39,214	44,604
Extended cohorts	210,620		200,987
New immigrants	38,480		23,546
	Match rates (unweighted/weighted):		
Summary Earnings Records	0.857/0.837	0.932/0.914	.918/.916
Numident	0.862/0.842	0.938/0.919	.931/.929
Master Beneficiary Record	0.517/0.493	0.561/0.502	.489/.445
Supplemental Security Record	0.125/0.120	0.138/0.132	.931/.929
	Topical module data:		
Marital, migration, fertility, and disability history	topical module 2	topical module 2	Core, SSA supplement
Medical expenses and health care utilization	topical modules 3 and 7	topical modules 4 and 7	Core
Retirement and pension plan coverage	topical module 7	topical module 3	Core, SSA supplement
Assets and liabilities	topical modules 3 and 6	topical modules 4 and 7	Core
Employer-provided health insurance, work history	topical module 5	topical module 6	Core
Annual income and retirement accounts	topical module 7	topical module 5	Core
Functional limitations/disabilities	topical module 5	topical module 6	Core
Child care expenses and child support payments	topical modules 3 and 6	topical modules 4, 7, 10	Core

Notes: The 2004 and 2008 SIPP panels contains four rotation groups. The survey starts in a different month depending on rotation group. Correspondingly, topical modules are fielded in different months depending on rotation group. The 2014 SIPP includes one interview to collect monthly data for calendar year 2013 and one SSA supplement (the final 2014 SIPP will include 4 annual interviews).

TABLE 3

MINT Summary Specification Table: Demographics, Including Health and Work Limitations

Process		Data	Form and predictors	For more information
Birth		PSID/NLSY	Self-reported (observed) SIPP fertility through SIPP panel. Completed fertility using nine separate logistic equations by marital status, parity. Predictors: age splines, duration since last birth, education, employment status, log of earnings, race, ethnicity, own mother was a teen mom indicator, own mother's education. Adjust post-hoc for DI status.	Rendall (2005; 2003); Table 2–17, Smith et al. (2010)
Death (see also Table 4: death, earnings, and disability)		SIPP 2001 to 2004, matched to Numident calibrated to OACT	Administrative data through 2010. Splicing method before age 67. Continuous time hazard for ages > 67: age splines, education, disability status, marital status, race, birth year, calendar year, permanent income.	Table 2–5, Smith et al. (2010)
LGB status		SIPP 2004, 2008, 2004–2013 ACS	LGB is set to 1 for all individuals in the SIPP data who report ever living with a same-sex unmarried partner. Other never-married individuals are assigned LGB status of 1 to reach a target rate of 1.8 percent in each cohort.	Smith, Rose, Cosic (2016)
DACS		SIPP 2004, 2008	DACS to true for individuals age 15 and older reporting a) being mentally retarded, b) having a developmental disability, c) being blind, or d) having never worked for pay because of own disability. The age-specific prevalence rate is about 1.5 percent.	
Living arrangements		SIPP, 2001, 2004, and 2008	Separate logistic regressions for initialization and continuation for ages < 62 and ≥ 62. Predictors: age, gender, education, race/ethnicity, number of children born, household income, marital status, SSI eligibility and participation, remarriage indicator, health, home ownership, nativity, mortality, interactions.	Re-estimated under MINT7: Tables A2.15 and A2.16, Smith and Favreault (2013)
Characteristics of coresidents		SIPP 2004–08 donors	Statistical match by group (defined on the basis of age, marital status, homeowner status, nativity, and childbearing history) from SIPP-based donor file based on respondent's per capita income.	Page 6–7, Smith et al. (2010)
Health status (fair or poor)		Ages 51–67: HRS 1992–2010; Ages 68+: SIPP/ SER/Numident 2008	Logistic, separate initialization and then by lagged status (excellent or fair-poor): age, sex, education, race, ethnicity, DI receipt, survival; At older ages, entry is separate by gender, and includes wealth/home ownership, nativity, marital status, and lagged earnings.	Re-estimated under MINT7: Tables A2.3, A2.4, A2.7, and A2.8, Smith and Favreault (2013)
Work limitations ages 51 to 67		HRS 1992–2010	Multinomial logit: age, sex, education, health, race/ethnicity, DI receipt, survival (in initialization).	Re-estimated under MINT7: Tables A2.5

Process		Data	Form and predictors	For more information
				and A2.6, Smith and Favreault (2013)
Marriage		SIPP, 2001, 2004, and 2008	SIPP marriage history. Continuous time hazard model, separate by gender and race (black, nonblack). Age splines, calendar time, duration unmarried, number of previous marriages, race/ ethnicity, education, immigrant status, widowed indicator, permanent income.	Re-estimated under MINT7: Table A2.1, Smith and Favreault (2013)
Spouse characteristics		PSID and SIPP	Probability tables derived from the PSID and SIPP. Characteristics include race, Hispanicity, spouse age difference by sex and marriage number.	Panis and Lillard (1999)
Spouse match (i.e., pointer to a specific person)		Assumption	Minimum distance matching function, variables in function include birth year, education, race, Hispanicity, disability start date, marriage begin and end dates, DI status, marriage termination type.	MINT7 update
Divorce		SIPP, 2001, 2004, and 2008	SIPP marriage history. Continuous time hazard model, separate by gender and race (black, nonblack). Age splines, duration married, calendar time, number of previous marriages, race, nativity, Hispanicity, education.	Re-estimated under MINT7: Table A2.2, Smith and Favreault (2013)
Immigration		SIPP calibrated to OACT and Homeland Security	Cloning method to impute new immigrants based on recent immigrants in the SIPP data. Gross age, sex, and legal status target population from OACT 2018. Source region and legal status shares from US Department of Homeland Security (2016, 2012a, 2012b).	Smith et al. (2010), chapter 2, section VI
Emigration		SIPP	Restricted to immigrants (i.e., the native born do not emigrate). Hazard function based on age, source region, time in the US, individual-specific permanent error term.	Dowhan and Duleep (2002)
Institutionalization at ages 62 and older		SIPP, 1990–93	Logistic regression. Age, marital status, education, race, nativity, homeowner status, health status, indicator dies within next two years.	Toder et al. (2002), chapter 7 (Table 7–2)
Link POLISIM target file to MINT donors for extended cohorts (characteristics and earnings before age 31)		POLISIM 2018 and MINT8	Minimum distance function statistical match. Match variables include race, education, marital status, immigration age, foreign born status, immigrant source region (developed, undeveloped).	Smith et al. (2007), chapter 5

Notes: SIPP matched data refers to SIPP matched to SER/DER/MBR Numident.

TABLE 2

MINT Summary Specification Table: Employment, Earnings, Disability Insurance

Process	Data	Form and predictors	For more information
DI receipt, earnings, and employment status for 1951–2010 for all ages	SIPP matched data	Observed from matched earnings records. For non-match cases, use hotdeck statistical match. Match variables include age, gender, death indicator, DI status, SSI status, report making a DC contribution on the SIPP, mean monthly earnings (7 categories), immigration age, immigrant source region, earnings status, education, race/ethnicity, class of worker.	Smith et al. (2010), chapter 4, section II
Years 2016+:			
Earnings, disability, and death through age 67 (later processes overwrite never-DI beneficiaries' post-age 54 earnings)	SIPP calibrated to OACT	"Splice" 5-year segments using statistical matching algorithm (hotdeck); variables in hotdeck match include DACS, age, gender, DI benefit indicator, number of years worked out of the last five, average earnings in last five years, work in year 5 of the match period, work in year 4 of match period, lifetime earnings quintile by cohort and sex, education, race/ethnicity, uncovered worker indicator, self-employed indicator, SSI receipt.	Smith et al. (2010), chapter 4 Toder et al. (2002), chapter 2
Earnings ages 55 through "retirement" for never-disabled	SIPP (2004 and 2008) matched to DER	"Trajectory method": standard age-earnings profile, separate by sex and education group, from fixed effects model. Predictors: age, cohort for women, 0.3 percent of high-earnings observations are capped (caps differ by education group). Capped earnings are reapplied after regression prediction is solved.	Re-estimated under MINT7: Tables A2.9 and A2.10, Smith and Favreault (2013)
"Retirement"	HRS matched data	Separate models by marital status and sex: replacement rate from Social Security, pension accruals, permanent earnings, age, education, health/work limitations status, nativity, self-employment, spouse characteristics (age, permanent income, pension characteristics) for married people, financial assets.	Re-estimated for MINT8: Smith, Williams, Murdrazija (2019), chapter 2.
Earnings for retirees, ages 55 to 61	HRS matched data through wave 13 (2016)	Separate entry and exit models: age, education, gender, lifetime earnings, work limitations, ethnicity/race, wealth (housing and financial).	Re-estimated under MINT7: Table A2.11, Smith and Favreault (2013)
Employment and earnings for OASI beneficiaries ages 60 to 69	SIPP (2004 and 2008) matched to DER	Four logistic participation models (separate entry and exit models for claiming age and subsequent ages) and five separate ordinary least squares (OLS) models for earnings for similar groups. Age, education, gender, health status, marital status, recent earnings, lagged employment/employment duration, lifetime earnings, incentives in OASI (non-contributory, dual entitlement), pension indicators.	Partially re-estimated under MINT7: Table A2.13, Smith and Favreault (2013)
Employment and earnings at ages 70 and older	SIPP (2004 and 2008) matched to DER	Employment modeled using separate equations based on work status last period. Age, education, gender, health status, wealth, lagged employment, duration of employment, recent earnings, lifetime earnings.	Re-estimated under MINT7: Table A2.14, Smith and Favreault (2013)

TABLE 5

MINT Summary Specification Table: Pensions (Including Job Characteristics)

Process	Data	Form and predictors	For more information
Pension coverage for job changers	SIPP 2001–04	Logistic regression: age, gender, education, ethnicity, nativity, earnings, job sector, OASDI coverage share, union status.	Tables 5–4 and 5–5, Smith et al. (2010)
Job change	SIPP 2001–04	Age-centered logistic regression: age, gender, education, Hispanicity, nativity, number of children, job tenure and job tenure squared, job sector, earnings and earnings change, union status, OASDI coverage status.	Table 5–1, Smith et al. (2010)
Job sector	SIPP 2001–04	Multinomial logit: age, gender, education, race, lagged job sector, earnings and earnings change, union status, region (DC indicator), OASDI coverage status and share.	Table 5–2, Smith et al. (2010)
Union status	SIPP 2001–04	Logistic regression: lagged union status, gender, education, race/ethnicity, earnings, job sector, OASDI coverage share.	Table 5-3, Smith et al. (2010)
Risk tolerance	SCF 1998–2007	Multinomial logit: age, education, marital status.	Table 5–8, Smith et al. (2010)
Contributions to DC pensions	1996 SIPP/DER match	Logit for whether contributes conditional on offer, tobit for amount contributed. Predictors: age, age squared, gender, marital status, number of dependents, jointly offered or frozen plan, contribute two years ago, own earnings/average wage, job tenure (1, 2, 3–4), spouse earnings/average wage, self-employment status. Increase in participation probability for new workers in plans with automatic enrollment. Amount given participation: race, employer contribution match, job tenure (≤ 1 year, 5+ years), homeownership status, jointly offer or frozen plan.	Tables 8–3 and 8–4, Smith et al. (2007)
Election of single life pension	HRS 1992–2000	Probit by gender: pension wealth, non-pension wealth, marriage duration, health status, marriage duration, race/ethnicity, education.	Table 8-11, Smith et al. (2007)
Decision to save lump sum distribution	SCF	Look-up table by age (8 groups) and size of distribution (4 groups). Probability of taking a lump sum distribution decreases with age and higher account balances.	Adapted from Moore and Muller (2002)
Work expenses	2004 and 2008 SIPP	Two-part model. First predict commute mode (drive alone to work, do not drive alone to work) with logistic regression estimates. The second model assigns work expenses as a share of earnings based on commute mode, earnings, and a random number to assign the expense distribution within earnings category. Commute mode predictors: earnings, sex, age, age squared, metro status, union status, education, race, employment sector.	Smith, Williams, and Murdrazija (2019) chapter 1.
Employer sponsored health insurance offer	2004 and 2008 SIPP	Logistic regression. Predictors: firmsize, region, employment sector, sex, education, mean recent earnings, post ACA year.	Smith and Favreault (2013) appendix table A1. 1

Employer sponsored health insurance premiums		For firms that offer health insurance, MINT imputes the full health insurance premium and the worker's share of the premium for both a family and a single plan. We assign the mean premiums based on a worker's geographic region and apply a distribution factor (percent of the mean) to calculate the health insurance premium (single and family plan) using data from Kaiser Family Foundation (2012).	Kaiser Family Foundation (2012) and Smith and Favreault (2013)

TABLE 6

MINT Summary Specification Table: Wealth

Process	Data	Form and predictors	For more information
Home purchase (among renters)	PSID 1968–94	Annual logistic regression hazard. Predictors: present value of lifetime earnings/cohort average (husband + wife), present value of lifetime earnings/cohort average squared (husband + wife), number of years with earnings above the taxable maximum (husband + wife), current year earnings/average wage (capped at 2.46) for husband and wife, single dummy, married dummy, black dummy, number of children under age 18, first child born indicator, self-employed dummy, number of years divorced.	Toder et al. (2002), chapter 6
Home sale (among homeowners)	PSID 1968–94	Annual logistic regression hazard. Predictors: present value of lifetime earnings divided by the cohort average (husband + wife), present value of lifetime earnings/cohort average squared (husband + wife), husband current year earnings/average wage (capped at 2.46), wife current year earnings/average wage (capped at 2.46), age*Hispanic, age*self-employed, age splines, married dummy, widowed dummy, single female dummy, divorced duration, first child born dummy, number of children less than 18.	Toder et al. (2002), chapter 6
Housing wealth ages 25 to 49, dependent variable= ln(home equity/average wage)	PSID 1968–94	Separate random effects models for unmarried and married homeowners. Individual-specific permanent error term imputed from PSID. Predictors: own present value of earnings/cohort specific average, former spouse present value of earnings/cohort specific average, average capped earnings in last five years, number of years of earnings above taxable maximum, age splines, age interactions (black, Hispanic, college, never married).	Toder et al. (2002), chapter 6
Housing wealth ages 50 to retirement, dependent variable =ln(home equity/ average wage); implemented as a change in assets from the prior year	HRS 1992–2004; calibrated to SCF	Separate random effects models for single and married homeowners. Predictors: age, age interactions (health, pension, self-employment, education, number of children ever born, race, male headed family indicator), number of years of earnings above the taxable maximum, present value of lifetime earnings/cohort average, spouse education.	Smith et al. (2007), chapter 3
Have mortgage	PSID 1968–2011	Separate random-effects logistic models for singles and married homeowners. Predicters: age, age interactions (health, pension, self-employment, education, number of children ever born, race, male headed family indicator), number of years of earnings above the taxable maximum, present value of lifetime earnings/cohort average, spouse education.	Smith, Williams, and Murdrazija (2019) chapter 1.
Nonhousing wealth to age 49, dependent variable=ln(home equity/average wage + 0.2)	PSID 1984–94; initial SIPP data calibrated to SCF	Separate random effects models for single and married. Individual-specific permanent error term imputed from PSID. Predictors: present value of per capita shared earnings/cohort specific average, mean earnings/ average wage in past five years capped at	Toder et al. (2002), chapter 6

Process	Data	Form and predictors	For more information
		2.46*average wage, age splines, cohort dummies, number of years of earnings above taxable maximum, age interactions (education, race, homeowner, marital status; for married, wife self- employed, wife education, husband race, husband education, husband self-employment), widowed and divorce intercepts.	
Nonhousing wealth age 50 to retirement, dependent variable= ln(financial assets/average wage + 0.02); implemented as a change in assets from the prior year	HRS; initial SIPP data calibrated to SCF	Separate random effects models for single and married. Predictors: age and age interactions (homeowner and renter dummies, number of children ever born, health, DI beneficiary, race, ethnicity, male dummy, DB and DC pension indicators, self-employment, widowed dummy, number of years ever married, own and spouse education, own and spouse self-employed), present value of earnings/cohort average, number of years of earnings over taxable maximum.	Smith et al. (2007), chapter 3
Non-housing wealth from retirement to death, dependent variable=ln(financial + retirement account assets/average wage + .02); implemented as a change in assets from the prior year	1984–93 SIPP data linked to SER	Separate OLS models for single and married. Predictors: homeowner, race, pension income receipt, average earnings/average wage age 50–60 > 1.2, widowed dummy, die within two years, husband or wife dies within two years.	Toder et al. (1999), chapter 7

Notes: All functions use initial values from SIPP self-reports where available.

TABLE 7

MINT Summary Specification Table: Other Income Sources and Health Coverage and Expenditures

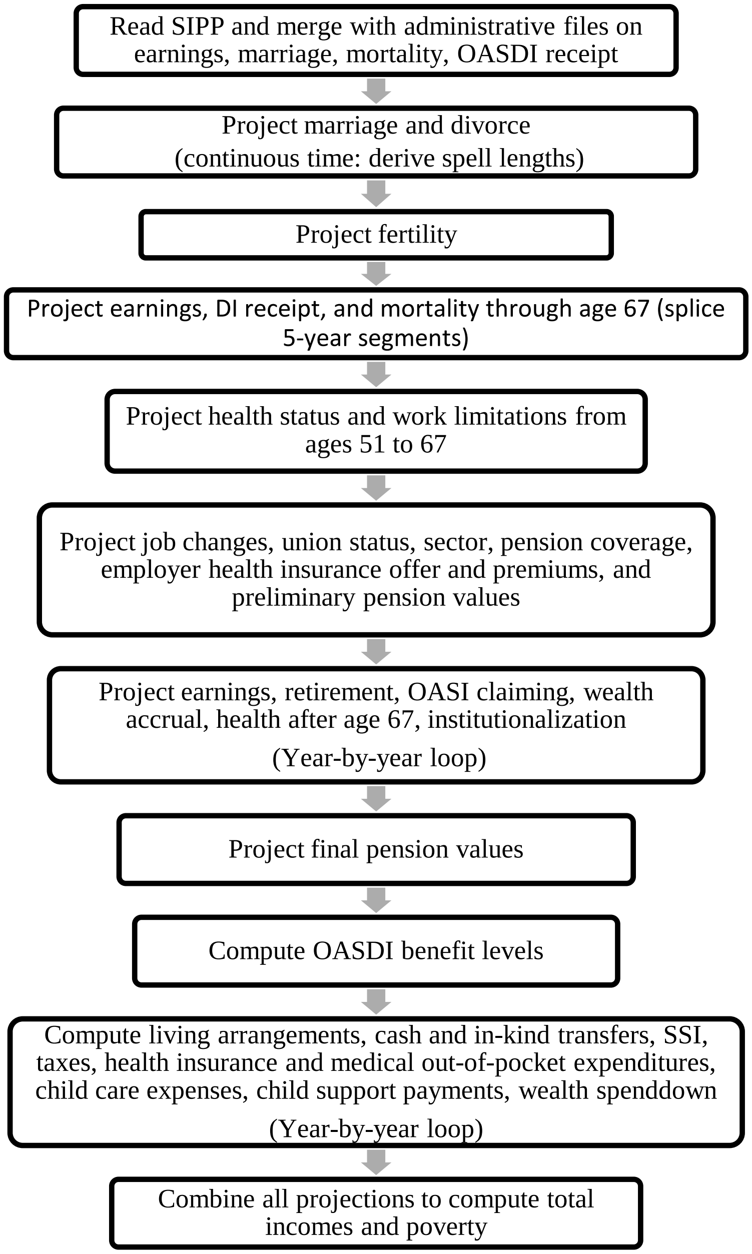
Process	Data	Form and predictors	For more information
OASI claiming	SIPP	Three separate hazard models for spouse only, low earners and high earners (defined based on relationship of lagged earnings to the exempt amount). Predictors: age, retirement indicator, spouse's claiming, pension coverage indicators, health status, marital status, wealth, lagged earnings, high earnings, lifetime earnings (PIA and present value of lifetime earnings). Updated in MINT8 to use age relative to the FRA.	2 equations re-estimated under MINT7: Table A2.12, Smith and Favreault (2013); spouses table 4-4 in Smith et al. (2010)
SSI participation among eligible individuals	SIPP	Separate models for new entrants and continuing beneficiaries in 3 age ranges: 25–61, 62–64, and 65 plus. Predictors: age, sex, education, nativity, marital status, expected federal benefit, Southern indicator, income, work limits, home ownership, years since last earned, total earnings years, lagged status, interactions with lag status. Uses TRIM state supplement data.	Young ages: Table 6-3, Smith et al. (2010); Older ages: unpublished tables
Means-tested transfers (TANF, general assistance)	SIPP (2001 and 2004 matched to DER and Numident)	Separate models for presence (logistic) and amount (OLS). Predictors: age, sex, education, marital status, number of children, homeowner status, survival, type and composition of income, state dummies, earnings and earnings changes.	Table 6-5, Smith et al. (2010)
Non-means-tested transfers (workers' comp, UI, temporary disability)	SIPP (2001 and 2004 matched to DER and Numident)	Separate models for presence (logistic) and amount (OLS). Predictors: age, sex, education, marital status, number of children, homeowner status, survival, type and composition of income, dummy for states with TDI, earnings and earnings changes, lagged SSI, health status, wealth.	Table 6-6, Smith et al. (2010); Presence re-estimated under MINT7: Table A2.17, Smith and Favreault (2013)
Noncash transfers (housing assistance, Low Income Home Energy Assistance Program, WIC, SNAP)	SIPP 2001–08	Separate models for presence (logistic, both initial and subsequent), and amount (OLS). Predictors: age, age squared, education, race, Hispanicity, nativity, marital status, health status, homeowner status, financial assets/average wage, year dummies, earnings and earnings changes, recent employment, family income/poverty (5 groups), dummy for states with TDI, other state dummies, metro status, SSI and OASDI, means-tested benefits, number and ages of children. When projecting subsequent receipt: lagged receipt status and lagged household benefit amount/average wage.	New in MINT7: Tables A1.7 and A1.8, Smith and Favreault (2013)
Health insurance coverage offered by employer	2004, 2008 SIPP	Employer health insurance offer: logistic regression among workers. Predictors: firm size (8 groups), region (4 groups), employer sector (4 groups), education, union, average earnings/average wage in last 3 years * sex, year ≥ 2014.	New in MINT7: Table A1.1, Smith and Favreault (2013)
Medigap purchase among Medicare beneficiaries	MEPS 2007–11	Logistic: age, race, ethnicity, ln of household income/average earnings.	New in MINT7: Table A1.2, Smith and Favreault (2013)

Process	Data	Form and predictors	For more information
Out-of-pocket medical expenses	MEPS 2007–11	Logistic for presence, OLS for ln amount, separate by married/unmarried. Age, education, ethnicity/race, detailed marital status, number of children, homeownership, wealth, detailed health insurance status indicators, SSI and OASDI, earnings and earnings changes, metropolitan status indicator, institutionalization indicator, state indicators.	New in MINT7: Tables A1.3 through A1.6, Smith and Favreault (2013)
Child support paid	2004 and 2008 SIPP	Two-part model. First predict if a parent with children under 18 living with a parent in another household pays child support. The second model predicts the child support payment amount for parents paying child support. Participation predictors include log of AWI adjusted total income, age, age squared, sex, education, marital status, and number of absent children.	Smith, Williams, and Murdrazija (2019) chapter 1.
Child care expenses	2004 and 2008 SIPP	Two-part model. First predict if workers with children under age 16 pay for child care with a logistic regression. The second model assigns child care amount for workers with expenses. Participation predictors include age, age squared, education, number of children under age 18, age of youngest child, age of youngest child squared, and married status (married or single). Amount predictors include age, age squared, education, number of children under age 18, log earnings AWI adjusted), age of youngest child, and married status.	Smith, Williams, and Murdrazija (2019) chapter 1.

Notes*:* All functions use initial values from SIPP self-reports or administrative records for starting values where available.

FIGURE 1

Stylized Representation of the MINT Processing Sequence



Notes: This process is stylized. For a more complete representation with references to key file, see Urban Institute (2013b), pages 8 and 9.

FIGURE 2

MINT8 Splicing of Earnings, Mortality, and Disability

DONOR FILE

	Age 50	Age 51	Age 52	Age 53	Age 54
Donor 1	$15,000, no DI, alive,	$16,000, no DI, alive	$16,000, no DI, alive	$17,000, no DI, alive	$18,000, no DI, alive
Donor 2	$50,000, no DI, alive	$52,000, no DI, alive	$54,000, no DI, alive	$55,000, no DI, alive	$54,000, no DI, alive
Donor 3	$30,000, no DI, alive	$31,000, no DI, alive	$6,000, enter DI, alive	$0, receive DI, alive	$0, receive DI, die

RECIPIENT FILE

	Age 45	Age 46	Age 47	Age 48	Age 49
Recipient 1	$14,000, no DI, alive	$15,000, no DI, alive	$14,000, no DI, alive	$15,000, no DI, alive	$15,000, no DI, alive
Recipient 2	$49,000, no DI, alive	$48,000, no DI, alive	$49,000, no DI, alive	$50,000, no DI, alive	$51,000, no DI, alive
Recipient 3	$28,000, no DI, alive	$29,000, no DI, alive	$31,000, no DI, alive	$30,000, no DI, alive	$32,000, no DI, alive

Recipient 1 matches to donor 1

Receipient 2 matches to donor 2

Recipient 3 matches to donor 3

RESULTING "SPLICED" FILE

	Age 45	Age 46	Age 47	Age 48	Age 49	Age 50	Age 51	Age 52	Age 53	Age 54
Recipient 1	$14,000, no DI, alive	$15,000, no DI, alive	$14,000, no DI, alive	$15,000, no DI, alive	$15,000, no DI, alive	$15,000, no DI, alive	$16,000, no DI, alive	$16,000, no DI, alive	$17,000, no DI, alive	$18,000, no DI, alive
Recipient 2	$49,000, no DI, alive	$48,000, no DI, alive	$49,000, no DI, alive	$50,000, no DI, alive	$51,000, no DI, alive	$50,000, no DI, alive	$52,000, no DI, alive	$54,000, no DI, alive	$55,000, no DI, alive	$54,000, no DI, alive
Recipient 3	$28,000, no DI, alive	$29,000, no DI, alive	$31,000, no DI, alive	$30,000, no DI, alive	$32,000, no DI, alive	$30,000, no DI, alive	$31,000, no DI, alive	$6,000, enter DI, alive	$0, receive DI, alive	$0, receive DI, die

Statement of Independence

500 L'Enfant Plaza SW
Washington, DC 20024

www.urban.org