FORCASTING INCIDENCE OF WORK LIMITATIONS, DISABILITY INSURANCE RECEIPT, AND MORTALITY IN DYNAMIC SIMULATION MODELS USING SOCIAL SECURITY ADMINISTRATIVE RECORDS: A RESEARCH NOTE

Melissa Favreault*

CRR WP 2002-09
December 2002

Center for Retirement Research at Boston College
550 Fulton Hall
140 Commonwealth Ave.
Chestnut Hill, MA 02467
Tel: 617-552-1762 Fax: 617-552-1750
http://www.bc.edu/crr

*Melissa Favreault is a Research Associate at the Urban Institute. The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium. The opinions and conclusions are solely those of the authors and should not be construed as representing the opinions or policies of SSA, the Internal Revenue Service, or any agency of the Federal Government or of the Center for Retirement Research at Boston College.

About the Center for Retirement Research

The *Center for Retirement Research at Boston College*, part of a consortium that includes a parallel center at the University of Michigan, was established in 1998 through a 5-year grant from the Social Security Administration. The goals of the Center are to promote research on retirement issues, to transmit new findings to the policy community and the public, to help train new scholars, and to broaden access to valuable data sources. Through these initiatives, the Center hopes to forge a strong link between the academic and policy communities around an issue of critical importance to the nation's future.

Center for Retirement Research at Boston College
550 Fulton Hall
140 Commonwealth Ave.
Chestnut Hill, MA 02467
phone: 617-552-1762 fax: 617-552-1750
e-mail: crr@bc.edu
http://www.bc.edu/crr

Affiliated Institutions:

Massachusetts Institute of Technology
Syracuse University
The Brookings Institution
Urban Institute

Introduction

In examining a number of important research questions related to the reform of the Social Security program, it is helpful to understand patterns of participation in the Disability Insurance (DI) program. DI beneficiaries comprise a large fraction, approximately 15 percent, of the pool of workers who receive Social Security benefits (Social Security Administration, 2001: Table 5.A16). They are a particularly vulnerable group in later life, with poverty rates more than twice as high as those for recipients of retirement or survivor benefits from Social Security (Thompson and Smith, 2002: Table A9-13c). Those who receive DI also have very different mortality experiences than those who do not (Zayatz, 1999), so careful modeling of the overlap between mortality and disability is essential when trying to determine the lifetime distributional consequences of Social Security reform. In addition, the larger disabled population, consisting of those who report work limitations but do not necessarily receive DI benefits, is also at higher risk of poverty and death than those who do not report work limitations.[1]

This research note explores these important intersections by presenting estimates from multivariate analyses of self-reported disability status, observed mortality, and reports of DI participation from administrative data. In our analyses, we use data from the 1990 through 1993 panels of the Survey of Income and Program Participation (SIPP) matched to the Social Security Administration's Summary Earnings Records (SER), Master Beneficiary Records (MBR), and Death Master File (from the Social Security number identification file, or Numident). Using administrative data allows us to improve upon several previous studies, as research consistently demonstrates that self-reports of earnings and disability benefit receipt are unreliable. Individuals often round up or down when reporting their earnings, particularly if they are asked about the distant past, and they frequently misreport social insurance and social assistance benefit types because they misunderstand the reasons that they receive benefits (see, for example, Huynh, Rupp, and Sears, 2002).

Background

Disability and Mortality Connections

In a recent paper, Rupp and Davies (2002) use data from the 1984 SIPP matched to administrative records to examine factors associated with death and with disability program dynamics, and to assess the importance of Social Security disability programs (DI and Supplemental Security Income, or SSI) from a life-course perspective. They pay particular attention to the roles of self-reported health and functional limitations, thus addressing the literature on the objectivity/endogeneity of self-reported health status. Our analyses explore similar issues, but with the more limited goal of estimating parameters for use in simulation models.[2] This entails specifying probabilities of work limitations transitions, DI take-up, and

---

[1] For convenience, we use the terms "work limited" and "disabled" interchangeably throughout this report, but recognize the many complexities associated with defining disability (see, for example, Nagi, 1969 or Burkhauser, Houtenville, and Wittenburg, forthcoming).

[2] The specific simulation model for which we are developing these estimates is the Urban Institute's Dynamic Simulation of Income Model (DYNASIM). DYNASIM was first developed in the 1970s, and has recently been updated. The current model is based on a self-weighting sample of over 100,000 persons from the 1990 to 1993 panels of SIPP. Over the years, the model has been used for a wide variety of applications associated with Social

death, preferably in one-year increments, in ways that will replicate joint distributions of characteristics as reliably as possible (subject to the constraint that many predictors of theoretical interest are not available as model covariates).[3]

Work Limitations

Our first goal is to replace the work limitations equations in DYNASIM with ones that are more current. Time series evidence on the incidence of disability among the aged suggests that disability rates are falling slowly, and health is modestly improving (Manton, Corder, and Stallard, 1997; Crimmins, Reynolds, and Saito, 1999). Further, the composition of the disabled population is shifting in important ways. More women now qualify for and receive benefits from the DI program than in the past, higher proportions of disabled workers now receive benefits at younger ages, and beneficiaries' reasons for receiving disability have changed. (For example, more beneficiaries now receive DI because of musculoskeletal impairments.) These trends suggest that the older DYNASIM work limitations model (estimated from PSID data from 1969 through 1972) may no longer produce reliable estimates.

Several recent efforts provide guidance on how to improve this function. Waidmann (2002) models work limitations using a multinomial logit specification and data from the Health and Retirement Study (HRS). His three-category dependent variable includes groups for no work limitation, presence of a work limitation that does not prevent one from working at all, and presence of a work limitation that prevents work. Significant predictors of work limitations category in his model include age, educational attainment, lifetime earnings and earnings pattern (specifically, whether one's earnings are rising), race, sex, and impending mortality. Waidmann finds similar results when he models health status using a binary indicator, differentiating those reporting to be in fair or poor health from all others (those reporting excellent, very good, or good health).

Favreault and Wolf (2002) estimate models of self-reported health status similar to Waidmann's, but for an older population. They use 1990 SIPP topical module data matched to the SER, and find that age, educational attainment, wealth, and impending mortality all significantly predict entries into poor health. Exits from poor health are less patterned in their analyses, but mortality, race, and recent earnings serve as significant predictors.

The principle advantage of the work limitations models that we estimate in this project is that they apply to a broad age range, not just those at midlife and older ages, as do the Waidmann estimates from HRS and the Favreault/Wolf estimates from SIPP. This broader applicability allows researchers to integrate the functions into a full population microsimulation model like DYNASIM.

Security reform. Favreault and Sammartino (2002) and Zedlewski (1990) provide additional information on DYNASIM. While the estimates that we present here are geared for use in DYNASIM, they could be useful for other models as well, including CORSIM/POLISIM or Modeling Income in the Near Term (MINT).

[3] One important aspect of replicating joint distributions is ensuring both cross-sectional and longitudinal validity. Because dynamic microsimulation models generate full life paths (e.g., marital and earnings histories), it is important that individuals' own characteristics are correlated over time. Using longitudinal models and complex error structures are two methods for enhancing longitudinal consistency.

Disability Insurance

Our second goal is to project DI onset given work limitations status. Burtless (1999) has conducted similar analyses of DI participation using the SER matched to the SIPP. His probit models, which he estimates separately by sex, control for age (5-year groupings), self-reported disability status, race/ethnicity (dummies for Hispanic whites and blacks), education (indicators for whether one has less than a high school education, 1 to 3 years of college, a college degree, or more than 5 years of college), and finally 10-year average earnings as a percent of economy wide average earnings (coded categorically with different breaks for men and women).[4] For both men and women, he finds strong effects of self-reported disability, age (with onset probabilities increasing with age), education (with onset probabilities declining with additional schooling), and race (with blacks having higher onset probabilities than whites). The effects of earnings differ somewhat for men and women. Only occupying the lowest earnings category has significant effects for women, decreasing the likelihood they will enter DI. For men, a more complex pattern is evident, with probabilities of onset lowest for very low earners, highest for the reference group of modest earners, and declining with income thereafter.

The principle innovations of the DI entry models that we present in this note over the Burtless models are the integration of nativity and marital status. We try to screen for DI eligibility using a PIA calculator that takes into account both the DI quantity and recency of work tests.[5] Burtless does not exclude individuals who are ineligible for DI from his sample.

Benitez-Silva et al. (1999) conducted another informative DI study. Their model replicates the DI application process in great detail. They first model the initial application and then the decision to appeal and the result for applications originally rejected. This is an important way of considering the structure of DI entry, as a large fraction of DI awards are made on appeal. These analysts use HRS data, so the population that they analyze is restricted to older persons (ages 51 and older) with disabilities. Benitez-Silva et al. find a strong association between self-reported work limitations and the results of both initial applications and appeals to SSA. Other significant covariates in their application models include health status indicators, age of application, and gender (with males more likely to apply for DI than females). Instead of excluding ineligibles from their model, they include an indicator of ineligibility for SSI/DI in their application equation. This indicator has the expected negative effect, and the coefficient is statistically significant. There are fewer significant predictors in the Benitez-Silva et al. appeals equations, although the self-report of work limitations is important.

---

[4] The breaks for women are < 15 percent, 15 to 30 percent, 30 to 70 percent (the reference category), 70 to 100 percent, 100 to 130 percent, and 130 percent or more. The breaks for men are the same through 130 percent, and then include three additional categories 130 to 180 percent, 180 to 210 percent, and 210 percent or more rather than just one.

[5] In order to receive benefits from the DI program, an individual needs to be both *fully insured* and *insured in the event of disability*. This means that a person must have accrued at least one quarter of coverage for each year elapsing after 1951 (or the year the worker turns 21, whichever is later) and before the year in which he or she becomes disabled, *and* he or she must also have worked at least half of the quarters during the ten years that preceded the disability (a shorter period for those disabled before age 31). In short, to qualify for DI one must have worked a substantial fraction of one's adulthood, and one must have worked fairly recently. For a discussion of coverage under the DI program, see Mitchell and Phillips (2001).

Kreider and Riphan (2000) also differentiate between DI application and award, and DI appeal and award using HRS data. As Burtless does, they model the process of entering DI separately for men and women. Their model is quite sophisticated, taking into account policy variables like local leniency and acceptance probability that we could not incorporate into our models because DYNASIM does not forecast them.

As we already noted, Rupp and Davies (2002) are also concerned with this problem. Using the 1984 SIPP, they find that baseline characteristics, including self-reports of work limitations and health status, affect disability program participation (for both DI and SSI) over the next 14 years. Like several other researchers, they thus reject the hypothesis that self-reported health status is endogenous. They also find that race, sex, and the square of age affect mortality, and that education, earnings, and marital status significantly affect disability program participation.

Mortality

A third goal for this project is to explore the possibility of developing a parsimonious mortality function that captures important differences between the mortality experiences of persons with and without work limitations, while simultaneously taking into account socioeconomic differentials. We assess the feasibility of using matched SIPP data for this purpose. This is an important objective because analysts have criticized models designed for examining distributional consequences of Social Security reform for inadequate attention to the relationships between disability and mortality (see, for example, Hayward's comments in Cohen et. al, 1999) and lifetime income and mortality (see, for example, Rust's suggestions for the MINT model, 1999).

Since Kitagawa and Hauser (1973)'s seminal work on mortality differentials, a number of researchers have tried to pinpoint the effects of indicators of status, such as income, occupation, and education on mortality. Some recent work, including research by Hurd, McFadden, and Merrill (1999) that relies on AHEAD data, suggests that socioeconomic differentials may decline with age. Other researchers (McCoy, Iams, and Armstrong 1994; Manton, Stallard, and Corder 1997) present evidence that the differential may persist into later life. Many analysts have emphasized the importance of socioeconomic differentials for understanding Social Security redistribution, both under current law and proposed reforms (see, for example, Aaron, 1977, and Garrett, 1995).

Less common is work on program participation and mortality. A recent study by Zayatz (1999) is surely the most comprehensive examination of the connection between DI receipt and mortality. This work draws from aggregate data on a large number of deaths (a 100 percent sample from the Social Security Administration MBR file, representing millions of life-years of exposure). Zayatz uses these data to construct life tables that incorporate not just age of disability onset and sex, but also the duration of disability. He finds that DI beneficiaries face much higher mortality risks than the general population. The analysis also reveals very high death probabilities in the initial years after disability onset, followed by much lower probabilities in later years.

Most analysts who wish to forecast deaths at the micro level do not have access to such rich data as the 100 percent MBR sample. The most accessible mortality data sources often include too few deaths among the disabled population to accurately project the effect of work limitation or DI receipt on mortality probabilities. Appendix Table 1 compares three data sources that analysts could use to estimate microanalytic mortality equations that would incorporate disability differentials: the National Longitudinal Mortality Study (NLMS), the Panel Study of Income Dynamics (PSID), and the SIPP, which we use here.

Panis and Lillard (1999) use the PSID to conduct one recent study of mortality risk that incorporates socioeconomic differentials, but not disability or DI beneficiary status. They find strong mortality effects for education and what they call permanent income. This study also introduces the interesting technique of calibrating mortality equation parameters to aggregate data (in this case, from Vital Statistics), using comparable regressions from aggregate and micro sources.

Favreault (2000) has estimated mortality probabilities using the NLMS. As Appendix Table 1 reveals, the NLMS compares favorably to SIPP and PSID in terms of the number of deaths one can observe. However, the NLMS is dated relative to SIPP, having been fielded over twenty years ago. Further, it is not possible in NLMS to determine interview dates with precision, and thus to estimate time trends, a critical feature for mortality models, given the consistent increase in longevity in recent decades. The NLMS disability indicator and its income measures are more crude than the SIPP indicators, and DI receipt status is not available as a predictor in NLMS. For these reasons, SIPP may have the potential to produce richer estimates for microsimulation than NLMS, despite its smaller sample sizes.

As mentioned earlier, Rupp and Davies (2002) use the 1984 SIPP to model mortality over the 14 years after one's initial interview. They find that the self-reported number of work limitations is an important predictor of death, as are age-squared, self-reported health status, program participation experience, race, and sex.

Models

As the discussion suggests, we examine several distinct dependent variables. First, we consider self-reported disability status. Second, we look at DI take-up given eligibility (in terms of insured status). Third, we consider mortality, specifically whether an individual dies within a 12- or 24-month period following a SIPP interview. Finally, we examine the cross of disability status and DI receipt, generating the following four-category dependent variable:

1. report a work limitation and collect DI;
2. do not report a work limitation but collect DI;
3. report a work limitation but do not collect DI; and finally
4. neither report a work limitation nor collect DI.

These analyses are informative because across many different data sources, large fractions of individuals who report DI benefits do not report having any work limitations. The models try to ascertain the factors associated with these seemingly inconsistent reports.

For the three binary dependent variables (work limitations, DI entry, and death), we employ discrete-time hazard models. We model the *conditional* probability of entering, exiting, or remaining in a state at time *t*, employing characteristics at *t-1* as predictors, such that:

$$\text{Prob}\{\ y_{it} = 1 |\ y_{it-1}\ \} = 1/(1 + e^{-(a\ +\ ßX_{it-1})})$$

where *a* is the intercept term, X is a set of explanatory variables, and ß is their associated coefficients. In the case of work limitations, we model both entries (reporting a work limitation at time *t* given that one did not have a work limitation at *t-1*) and exits (reporting no work limitation at time *t* given that one did have a work limitation at *t-1*). For DI, we model only entries, as DI exits for reasons other than death or administrative conversion to retirement benefits are too rare in the data to permit reliable estimation. Because death is an absorbing state, we model only entries into it (or, put another way, we model exits from being alive).

For the cross-classification of work limitations and DI, we employ a multinomial logit specification. Because many of the transition cells have very few observations, we use a cross-sectional (or static) rather than a longitudinal (or dynamic) model. That is, we consider correlates of occupying one of the four states, rather than factors associated with making a transition from one into another. As a result, in this case we use contemporaneous variables (rather than lagged variables) as predictors. The form of the multinomial logit is as follows:

$$\text{Prob}\{\ y_{it} = 1\ \} = 1 / (1 + e^{ß_1 X_{it}} + e^{ß_2 X_{it}} + e^{ß_3 X_{it}})$$
$$\text{Prob}\{\ y_{it} = 2\ \} = e^{(ß_1 X_{it})} / (1 + e^{ß_1 X_{it}} + e^{ß_2 X_{it}} + e^{ß_3 X_{it}})$$
$$\text{Prob}\{\ y_{it} = 3\ \} = e^{(ß_2 X_{it})} / (1 + e^{ß_1 X_{it}} + e^{ß_2 X_{it}} + e^{ß_3 X_{it}})$$
$$\text{Prob}\{\ y_{it} = 4\ \} = e^{(ß_3 X_{it})} / (1 + e^{ß_1 X_{it}} + e^{ß_2 X_{it}} + e^{ß_3 X_{it}})$$

where X again signifies a set of explanatory variables, this time including an intercept, and ß their associated coefficients.

Our ultimate goal in specifying these models is to integrate the best possible predictions of disability and mortality into a dynamic microsimulation model. As a result, not all predictors of theoretical interest are available to us. In particular, the range of contextual variables upon which we can draw is quite limited. For this reason, our models focus on socio-economic differentials in the probability of becoming disabled, entering DI, or dying.

We use many of the same explanatory variables that previous authors have used, including standard demographic and economic variables (age, education, lifetime earnings, race/ethnicity). For example, we use the same breaks for coding earnings as Burtless (1999) employs, though we define longitudinal earnings differently. Additionally, we incorporate indicators for marital status and nativity, found in prior work to be important correlates of disability and mortality experience.

Data

To estimate parameters in the models, we use data from the 1990 through 1993 SIPP waves matched to Social Security Administration earnings, benefit receipt, and mortality records. The exact subset of the SIPP data that we use varies by analysis. For example, in estimating the work limitations transitions, we need to observe whether a person is limited at two points spaced exactly twelve months apart. We thus chose the 1990, 1992, and 1993 waves of the SIPP for this analysis because of the presence of appropriate questions on work limitations in at least two topical modules.

We have more flexibility in the numbers of SIPP panels that we can use when modeling participation in the Social Security Disability Insurance program, because the linked administrative records provide monthly reports of benefit participation. We therefore use data from all four panels (1990 through 1993) when examining DI entry.

The mortality models that we present here are quite simple, and likewise require fewer restrictions than the models of work limits. The Numident data (matched to the SIPP) that we use include reports of deaths through 1999. This allows us to look up to ten years out for some SIPP interviewees. We only look one and two years out, however, because the microsimulation models in which they would be imbedded typically age their populations in single-year increments. We contrast these SIPP models with others we estimate using the National Longitudinal Mortality Study.

Estimates

Work Limitations

The coefficient estimates for entries and exits from work limitations from SIPP are consistent with patterns from the literature. Table 1 presents the logistic coefficients from our detailed work limitations models. We present standard errors in parentheses, and denote statistically significant coefficients with asterisks. (Appendix Table 2 shows coefficients from a more parsimonious version of the model, which replicates the old DYNASIM model.) One can interpret these coefficients as the effects of a one-unit change in the variable on the log-odds of either becoming work limited (in the entry model) or becoming no longer limited (in the exit model).

Coefficients in the models for entry into work limitations (that is, reporting a work limitation this year given that one did not report a work limitation last year) are fairly similar for men and women (the first two columns of the coefficients). For both sexes, the chances of acquiring a work limitation tend to increase with age (though in a non-linear fashion) and decline with education. Race/ethnicity and nativity indicators do not have significant effects on entry into having a work limitation, all else equal, for either men or women.

Table 1. Work Limitations Models from SIPP:
Logistic Coefficients and Standard Errors

	Enter work limit				Exit work limit	
Variables	Men		Women		Pooled	
Intercept	-2.4739***	(0.1428)	-2.9826***	(0.1384)	-1.4274***	(0.1473)
Age dummies (ref: 61-67)						
<=25	-1.9088***	(0.1555)	-1.8260***	(0.1549)	0.6184***	(0.1473)
26-30	-1.2706***	(0.1530)	-1.0364***	(0.1434)	0.3471*	(0.1451)
31-35	-1.0523***	(0.1405)	-0.8025***	(0.1299)	0.0969	(0.1347)
36-40	-0.8259***	(0.1340)	-0.6986***	(0.1289)	-0.0629	(0.1336)
41-45	-0.6273***	(0.1340)	-0.5208***	(0.1271)	-0.2563*	(0.1301)
46-50	-0.5053***	(0.1369)	-0.5192***	(0.1340)	-0.2510*	(0.1278)
51-55	-0.2642	(0.1354)	-0.1597	(0.1282)	-0.3403*	(0.1323)
56-60	0.00116	(0.1340)	-0.1430	(0.1303)	-0.4889***	(0.1313)
Hispanic indicator	0.0991	(0.1180)	-0.0262	(0.1156)	0.1600	(0.1202)
Black indicator	0.0898	(0.1152)	0.0801	(0.1020)	0.0490	(0.1094)
Asian indicator	-0.1428	(0.2260)	0.0402	(0.1966)	-0.0183	(0.2523)
Education dummies (ref: high school grad)						
Less than high school grad	0.2909***	(0.0867)	0.2724**	(0.0832)	-0.1948*	(0.0826)
Any college	-0.3474***	(0.0795)	-0.3411***	(0.0753)	0.0585	(0.0820)
Marital Status indicators (ref: married)						
Never married	0.3496***	(0.1041)	0.5559***	(0.1025)	-0.3026**	(0.1044)
Divorced or separated	0.2596*	(0.1061)	0.4757***	(0.0895)	-0.0864	(0.0933)
Widowed	0.2543	(0.2798)	0.4464***	(0.1301)	0.2928*	(0.1427)
Native born indicator	0.0924*	(0.0834)	0.0845	(0.0817)	-0.3104***	(0.0842)
Male indicator	--		--		0.1206	(0.0699)
Recent (3-year) average earnings / average wage dummies ("RAE")(ref: 0.30 <= RAE < 0.70)						
RAE = 0.00	0.2712*	(0.1153)	0.3959***	(0.0959)	-0.3124**	(0.1056)
0.00 < RAE < 0.15	0.1535	(0.1244)	0.3885***	(0.1010)	-0.1959	(0.1136)
0.15 <= RAE < 0.30	0.3371*	(0.1316)	0.1798	(0.1231)	-0.1075	(0.1370)
0.70 <= RAE < 1.00	-0.3014*	(0.1247)	-0.2690*	(0.1292)	-0.1297	(0.1555)
1.00 <= RAE < 1.30	-0.4292**	(0.1361)	-0.3422*	(0.1566)	0.4611**	(0.1612)
RAE >= 1.30	--		-0.2727	(0.1410)	0.2901*	(0.1400)
1.3 <= RAE < 1.8	-0.6524***	(0.1343)	--		--	
1.8 <= RAE < 2.1	-0.3132	(0.1644)	--		--	
RAE >= 2.1	-0.7625***	(0.1497)	--		--	
N (person years)	28,300		31,623		8,157	
Number of transitions	982		1,076		1,085	
-2 log-likelihood	8,102.255		8,987.388		6,245.057	

Source: Urban Institute estimates from the 1990, 1992, and 1993 SIPP matched to the SSER and MBR
Notes: *** indicates p< 0.001; ** indicates p< 0.01; * indicates p< 0.05; we define recent earnings variables using the average of the past three earnings years divided by the average wage

Marital status appears to be an important predictor of work limitations entry, with never married and divorced or separated adults more likely to enter disability than those who are married.[6] Among women, widows are also more likely to enter work limitations than their married counterparts.

The pattern in association between work limitations and earnings is important, generally showing a decline in the chances of becoming work limited with an increase in recent earnings (defined as the average of earnings divided by the average wage in the three years preceding the interview), though there is evidence of non-linearities.[7] For both men and women, coefficients on the earnings variable are positive before the reference group (of between 0.3 and 0.7 times the average wage) and negative after it, but not all coefficients are statistically significant.[8]

The model for exit from work limitations (that is, reporting no work limitation this year even though one reported a work limitation last year), reveals, as one would expect, basically the opposite patterns as the entry models. In this case, we pool observations for both men and women, and include an indicator for whether one is a male in the equation. This indicator has a negative (but not statistically significant) coefficient, suggesting that men may be less likely than women to recover from work limitations, but that we cannot say with confidence that this effect differs from zero. The chance of exiting from having a work limitation declines with age (with the exception of the reference group). It increases with education and recent earnings (again, defined based on the past three years, and using the same reference category).[9] Never married adults are less likely than married adults to recover once they have reported a work limitation, though being divorced or separated does not appear to significantly affect disability exit. Being widowed may actually increase probability of reporting exit from work limitations. Native-born adults are less likely to exit disability than immigrants, all else equal.

DI Entry and Receipt

Table 2 presents estimates from the SIPP models of DI entry. In this case, the dependent variable is the probability of entry into DI given that one was not receiving DI last year, *conditioned on eligibility.* We present separate models for men and women, and also a pooled version with an indicator for being male. As in the previous table, we present the standard errors in parentheses and use asterisks to denote statistically significant effects.

---

[6] Hypothesizing that the effects of marital status may vary by age, we estimated the models using interactions between age and marital status. In some specifications, being younger and never married had negative but only marginally significant effects on entering work limitations. These results are available upon request.
[7] We experimented with alternative specifications of earnings, including a linear measure with a squared term and averaging over longer intervals. The specification we present here provides a better fit than the alternatives.
[8] Once again speculating that the effects of these variables (earnings) may vary by age, we estimated the models using interactions between having zero earnings and age. In some specifications, being younger and not having worked recently had negative but only marginally significant effects on entry into work limitations. These results are also available upon request.
[9] To try to enhance intertemporal continuity, we added additional variables for number of years elapsed since one had had covered earnings (coded linearly, categorically, and with varied topcodes) in some specifications of the model. These results are likewise available upon request.

Table 2. Disability Insurance Models from SIPP:
Logistic Coefficients and Standard Errors

	Enter DI given Eligibility		
Variables	Men	Women	Pooled
Intercept	-7.4296*** (0.8193)	-7.5309*** (0.8789)	-7.1818*** (0.5749)
Age dummies (ref: 61-67)			
<=25	0.2315 (0.5369)	-0.9986 (0.7200)	-0.4199 (0.4057)
26-30	0.1060 (0.3939)	0.5197 (0.4835)	0.1309 (0.2901)
31-35	0.1658 (0.3123)	1.0348* (0.4313)	0.4557 (0.2417)
36-40	0.3512 (0.2888)	0.4224 (0.4697)	0.3375 (0.2442)
41-45	0.2616 (0.2965)	1.2560** (0.4195)	0.6330** (0.2314)
46-50	0.7596** (0.2721)	1.7327*** (0.4081)	1.1041*** (0.2187)
51-55	1.1313*** (0.2546)	2.0471*** (0.3976)	1.4481*** (0.2087)
56-60	1.5169*** (0.2407)	2.0736*** (0.3935)	1.6846*** (0.2023)
Non Hispanic black indicator	0.7076*** (0.1824)	-0.3219 (0.2549)	0.2835 (0.1470)
Education dummies (ref: high school grad)			
Less than high school grad	0.4987*** (0.1364)	0.6162*** (0.1565)	0.5723*** (0.1025)
Marital Status indicators (ref: married or divorced)			
Never married	0.2866 (0.2003)	0.5810*** (0.2073)	0.3940** (0.1419)
Widowed	0.9758** (0.3692)	0.4199 (0.2629)	0.5001* (0.2117)
Homeowner indicator	0.1185 (0.1410)	-0.3811** (0.1451)	-0.1078 (0.0999)
Work limited at t-1 indicator	1.7703*** (0.1306)	2.0170*** (0.1450)	1.8791*** (0.0970)
Indicator of impending mortality (within 3 years)	2.1699*** (0.1999)	2.0018*** (0.2929)	2.1105*** (0.1640)
Number of years worked	0.0143 (0.1063)	0.0361 (0.1238)	0.0039 (0.0793)
Number of years worked^2	0.0031 (0.0040)	-0.0004 (0.0050)	0.0019 (0.0031)
Native born indicator	-0.1204 (0.2298)	-0.0475 (0.2369)	-0.0793 (0.1647)
Male indicator	--	--	0.0822 (0.0964)
Recent (3-year) average earnings / average wage dummies ("RAE")(ref: 0.30 <= RAE < 0.70)			
RAE = 0.00	-2.7516*** (0.7270)	-1.1562** (0.3921)	-1.7726*** (0.3374)
0.00 < RAE < 0.15	-0.4904* (0.2331)	-0.3067 (0.2264)	-0.4091** (0.1622)
0.15 <= RAE < 0.30	0.0052 (0.2219)	0.0593 (0.2029)	0.0424 (0.1493)
0.70 <= RAE < 1.00	-0.1752 (0.1877)	-0.1745 (0.2034)	-0.1728 (0.1373)
1.00 <= RAE < 1.30	-0.6650** (0.2336)	-0.2244 (0.2532)	-0.4750** (0.1716)
RAE >= 1.30	--	-1.0790*** (0.3178)	-0.6739*** (0.1453)
1.3 <= RAE < 1.8	-0.4199* (0.2056)	--	--
1.8 <= RAE < 2.1	-0.7502* (0.3102)	--	--
RAE >= 2.1	-0.8767*** (0.2567)	--	--
N (person years)	86,232	77,233	163,465
Number of entries	289	224	513
-2 log-likelihood	3,302.065	2,562.760	5,925.494

Source: Urban Institute estimates from the 1990 through 1993 SIPP matched to the SSER and MBR

Notes: *** indicates $p < 0.001$; ** indicates $p < 0.01$; * indicates $p < 0.05$; we define recent earnings variables using the average of the past three earnings years divided by the average wage

In all three models, DI entry is significantly associated with age, with older persons more likely to enter than younger ones, though this plateaus with the reference group (persons ages 61 to 67). This reduction for the oldest group may be due to the availability of retired worker benefits at age 62 (leading some people to classify themselves as not working because of retirement rather than disability). For men, significant increases in entry probabilities begin at ages 46 to 50, while for women they begin earlier. Having reported a work limitation last year, not surprisingly, has a very large effect on the probability of entering DI in all three equations. (Recall that there is a 5-month waiting period for DI entitlement, so lags between disability/work limitation onset and DI receipt should not be unusual.) The only coefficients that rival work limitations in size are those for impending mortality, which are also strongly, positively associated with DI take-up for all three groups, and, among men, the coefficient for not having any earnings in the last three years, which is strongly negatively associated with entry.[10]

Having less than a high school education is positively associated with claiming DI benefits. Race, specifically being non-Hispanic black, has significant effects on DI claiming among men but not among women. Once again, marital status appears to have some important effects, with being never married significantly, positively associated with DI take-up in both the pooled model and the model for women, and being widowed significantly, positively associated with DI take-up in both the pooled model and the model for men.

Recent average earnings (defined over the three-year interval preceding the interview) have somewhat surprising effects in these models. As already noted, not having any earnings has very large, negative, statistically significant effects on DI take-up for men, and more modest, but still relatively large, negative, and statistically significant effects, for women. The earnings pattern otherwise suggests an inverted U-shaped non-linearity. That is, the probability of taking up DI increases until the group just before the reference group (those with modest average earnings of between 15 and 30 percent of the average wage), and then decreases steadily again at higher levels of earnings. This resembles the Burtless findings, and we would expect some differences because of the differing treatments of eligibility. We do not find effects of homeownership or of the number of years worked in covered employment net of these controls for earnings and other factors in any of the models.

Mortality

Tables 3 and 4 provide the logistic coefficients from the mortality models, again with standard errors in parentheses and asterisks denoting the statistically significant effects. As in the models for entry into having a work limitation and into DI, we stratify the sample by gender. We examine two different versions of the dependent variable: one in which a respondent dies within a year of the SIPP interview (Table 3), and a second in which the respondent dies within two years (Table 4).[11] We limit the sample to persons between the ages of 16 and 67 at baseline.

[10] The inclusion of the impending mortality indicator may appear counterintuitive, given that one normally assumes that disability predicts mortality, not the reverse. We change the causal sequencing in this case in order to be consistent with some microsimulation models, which for technical reasons process decisions about death prior to decisions about DI take-up. Correlating the outcomes is more important than using an intuitive processing order.

[11] These models (the one- and two-year versions) have different strengths and weaknesses. The further after the baseline interview that one looks, the more deaths that one observes, and hence the more reliable the coefficients on fixed (or reasonably fixed) characteristics, like age, ethnicity/race, and in most cases education. However, looking

Table 3. Mortality Models for Ages 16 though 67 from SIPP and NLMS: Logistic Coefficients and Standard Errors for Single-Year

	SIPP		NLMS	
Variables	Men	Women	Men	Women
Intercept	-5.1803***(0.2787)	-5.6813***(0.3268)	-4.3599***(0.2017)	-4.8158***(0.2268)
Age dummies (ref: 61-67)				
<=25	-2.2566***(0.2749)	-2.4540***(0.4337)	-2.6814***(0.1289)	-2.9992***(0.2122)
26-30	-2.6255***(0.4238)	-1.5465***(0.3631)	-2.8136***(0.1849)	-2.6386***(0.2404)
31-35	-1.5375***(0.2506)	-2.7836***(0.5941)	-2.7295***(0.1850)	-2.4092***(0.2247)
36-40	-1.7790***(0.2820)	-1.9383***(0.4021)	-2.2591***(0.1628)	-1.8814***(0.1949)
41-45	-1.8244***(0.2889)	-1.0103***(0.2704)	-2.2303***(0.1719)	-1.8727***(0.2072)
46-50	-0.9083***(0.2082)	-0.9985***(0.2809)	-1.4111***(0.1207)	-1.5062***(0.1769)
51-55	-0.7881***(0.2068)	-0.7000** (0.2505)	-1.2171***(0.1064)	-0.8958***(0.1312)
56-60	-0.4915** (0.1839)	-0.9037***(0.2649)	-0.6307***(0.0843)	-0.4543***(0.1093)
Hispanic indicator	0.0007 (0.2414)	-0.6374 (0.3558)	-0.3229 (0.1895)	-0.3648 (0.2446)
Black indicator	0.3454 (0.1862)	0.2693 (0.2166)	0.2225* (0.0977)	0.3388** (0.1159)
Education dummies (ref: high school grad)				
Less than high school grad	0.7981***(0.1354)	0.8815***(0.1754)	0.0973 (0.0756)	0.2509* (0.1015)
Any college	-0.0917 (0.1921)	-0.0994 (0.2528)	-0.3322***(0.0798)	-0.1634 (0.1086)
Native born indicator	0.1887 (0.2433)	-0.1586 (0.2777)	0.3942* (0.1776)	0.1093 (0.1893)
Indicator work limited at t-1	1.3050***(0.1276)	1.6495***(0.1673)	1.4695***(0.0919)	2.2060***(0.1323)
Time trend (year-1990)	-0.2590***(0.0534)	-0.2425***(0.0684)	--	--
Missing data indicators (Ref: nonmissing)				
Hispanicity	--	--	0.2985 (0.1925)	-0.2148 (0.1103)
Nativity	--	--	-0.0055 (0.0779)	0.1686 (0.2846)
Education	--	--	0.6145 (0.7280)	1.1455 (1.0426)
Employment status	--	--	-1.3682 (0.8217)	-12.1523 (220.7)
Social Security Number	--	--	0.2223* (0.0988)	0.1047 (0.1310)
N (Total person years)	101,941	111,494	202,651	218,278
Number of deaths	283	170	1,120	635
-2 log-likelihood	3,402.143	2,179.474	12,165.656	7,688.579

Source: Urban Institute estimates from the 1990 through 1993 SIPP matched to Numident and NLMS

---

further after the baseline does introduce possible biases, including, for example, emigration bias (e.g., immigrants who return to their home country to die, and who thus whose deaths do not appear in vital registration data in the United States). Also, coefficients on characteristics which are likely to change, for example marital status or work limitations, become less meaningful given the complex lag structure.

Notes: *** indicates $p < 0.001$; ** indicates $p < 0.01$; * indicates $p < 0.05$

Table 4. Mortality Models for Ages 16 though 67 from SIPP and NLMS: Logistic Coefficients and Standard Errors for Two-Year

	SIPP		NLMS	
Variables	Men	Women	Men	Women
Intercept	-4.4499***(0.2088)	-4.5525***(0.2016)	-3.6444***(0.1399)	-4.2564***(0.1652)
Age dummies (ref: 61-67)				
<=25	-2.3205***(0.1730)	-2.7164***(0.2823)	-2.8182***(0.0984)	-3.0696***(0.1598)
26-30	-2.5910***(0.2538)	-2.4622***(0.3146)	-2.6344***(0.1164)	-2.9064***(0.1861)
31-35	-2.0339***(0.1891)	-1.8934***(0.2307)	-2.9236***(0.1365)	-2.4680***(0.1567)
36-40	-1.6817***(0.1665)	-2.2638***(0.2722)	-2.5166***(0.1240)	-2.3350***(0.1630)
41-45	-1.9531***(0.1884)	-1.2331***(0.1757)	-2.1166***(0.1121)	-1.6940***(0.1322)
46-50	-1.2435***(0.1486)	-1.1480***(0.1797)	-1.5498***(0.0888)	-1.7907***(0.1397)
51-55	-0.9366***(0.1392)	-0.9880***(0.1716)	-1.2749***(0.0762)	-0.9606***(0.0945)
56-60	-0.5310***(0.1207)	-0.6347***(0.1487)	-0.6739***(0.0597)	-0.5082***(0.0782)
Hispanic indicator	-0.2502 (0.1849)	-0.2998 (0.2019)	-0.2274 (0.1267)	-0.4650* (0.1819)
Black indicator	0.3626** (0.1235)	0.2058 (0.1437)	0.2519***(0.0688)	0.4051***(0.0817)
Education dummies (ref: high school grad)				
Less than high school grad	0.7017***(0.0907)	0.7667***(0.1120)	0.1210* (0.0533)	0.2798***(0.0724)
Any college	0.0305 (0.1188)	0.1387 (0.1452)	-0.2651***(0.0547)	-0.0711 (0.0751)
Native born indicator	0.5492** (0.1905)	-0.1225 (0.1722)	0.4262***(0.1239)	0.1276 (0.1380)
Indicator work Limited at t-1	1.0711***(0.0854)	1.4650***(0.1039)	1.4470***(0.0685)	1.9593***(0.1072)
Time trend (year-1990)	-0.2461***(0.0354)	-0.2782***(0.0434)	--	--
Missing data indicators (Ref: nonmissing)				
Hispanicity	–	--	0.1342 (0.1474)	0.0405 (0.2144)
Nativity	–	--	0.0346 (0.0540)	-0.0145 (0.0735)
Education	–	--	-0.3948 (0.4981)	0.8335 (0.7792)
Employment status	--	--	0.1535 (0.4630)	-0.9735 (0.9561)
Social Security number	--	--	0.1760** (0.0678)	0.2078* (0.0964)
N (Total person years)	101,941	111,494	202,651	218,278
Number of deaths	650	432	2,329	1,274
-2 log-likelihood	6,842.965	4,814.142	21,911.828	13,724.854

Source: Urban Institute estimates from the 1990 through 1993 SIPP matched to Numident

Notes: *** indicates p< 0.001; ** indicates p< 0.01; * indicates p< 0.05

To verify the validity of these models, we contrast the SIPP coefficients with coefficients from analogous models that we estimate using the NLMS.

As expected, age is a primary determinant of one's probability of dying in the coming year (or two) in both SIPP and NLMS. (We code age categorically for illustrative purposes, with alternative specifications, like age and age squared, available upon request.) The results further reveal a very strong effect of work limitations on mortality regardless of sex, data set, or the time interval over which mortality probabilities are estimated. This is consistent with the findings in Davies and Rupp (2002). The estimates also reveal a substantial effect of having less than a high school education on mortality risk in nearly all specifications. In the SIPP models, we can exploit variation in year of interview to estimate a time trend, and we find a significant decline in the probability of dying for each year into the future.

Race and nativity are significantly associated with mortality risk for men in the two-year SIPP mortality probabilities and in all the NLMS probabilities (for both men and women and one- and two-year versions). Namely, blacks have higher death probabilities than persons of other races, and persons born in the U.S. have higher probabilities than those born abroad. The literature documents the former relationship well, and the latter relationship is likely due to the selectivity of immigrants. Hispanicity has significant effects in one NLMS model, reducing women's probability of death.[12] Perhaps there are no significant race or nativity effects for women in SIPP and for men in the one-year estimates in SIPP because there are substantially fewer cases of death (compared to the NLMS or the two-year male SIPP file).

We included recent individual earnings measures in several of the preliminary models from SIPP (not shown), but did not find significant effects. In NLMS, in contrast, we did find significant relationships between death and family earnings, defined both using thresholds and using quintiles. This finding is interesting, given that NLMS measures income much less well than SIPP (especially when the latter survey is matched to the administrative earnings and benefit records), and suggests that future work in this area should focus on family rather than individual income or earnings.

Examining the NLMS and SIPP coefficients side-by-side (in Tables 3 and 4) allows us to compare the relative merits of the two data sources for developing mortality models for persons in these age ranges (16 to 67). (Note that because the NLMS equations include a number of controls for missing data and do not incorporate a time trend, the SIPP and NLMS equations are not strictly comparable.) There is a striking qualitative similarity between the SIPP and NLMS equations that is reassuring. The signs of the coefficients are almost always consistent. Although the absolute magnitudes of coefficients sometimes differ (in part because of the time trend and intercept/slope differences), the relative magnitudes are similar.

---

[12] The literature on Hispanic mortality is more ambiguous. See, for example, Hayward and Heron (1999), Sorlie et al. (1992; 1993), and Rosenberg et al. (1999). The latter study is very useful, providing quantitative estimates of certain biases that affect estimation of death probabilities by race.

Joint Work Limitations and DI Receipt

Table 5 displays results from the analyses of the intersection of work limitations and DI participation. For this model, we pool observations for men and women and include an indicator for sex. In all cases, one can interpret coefficients as the effects of a one-unit change in the variable on the log-odds of occupying the category in question, being on DI and work limited, being on DI and not work limited, or being work limited but not on DI, compared to the base case of neither reporting a work limitation nor collecting DI benefits.

Consistent with the work limitations and DI entry models just reported, these analyses show clearly the importance of age, marital status, and socioeconomic standing in determining one's probability of being disabled, whether this is defined as self-reporting a work limitation, receiving DI benefits, or both. Impending death has a large effect on the probability of being in a disability category, again suggesting that neither defining oneself as having a disability nor having the government define oneself this way is strictly subjective or simply a rationalization for not working or obtaining benefits.

Specifically, having less than a high school education is positively associated with occupying each of the three disability categories, while having some college or being a college graduate or higher is negatively associated with it. Men are more likely then women to occupy one of the disability categories, all else equal. Once again, marital status appears to have important associations to disability, with never married, divorced or separated, and widowed people more likely to occupy one of the three disability categories compared to the reference group. As in the previous functions, the association does not necessarily imply a causal link. It could be documenting a selection effect (e.g., persons with work limitations are less likely to marry or stay married) rather than demonstrating a protective effect for marriage. Age tends to accelerate the probability of being disabled by one of these three definitions, with the exception of the reference category.

Race and ethnicity have little effect on occupying a disability category once we include all these controls, with the exception of being black (which is positively associated with occupying the anomalous DI-no work limitation category compared to the base category of being neither work limited nor on DI). Nativity is positively associated with occupying the DI and work limit and just work limit categories, but not with the intermediate category of DI and no reported work limit. Earnings have more linear effects on one's joint work limitations-DI category than they did in the DI entry models alone (which had contained the DI eligibility screen). In both categories in which a person is collecting DI (with or without reporting a work limitation), the probability of occupying the category decreases monotonically with recent average earnings (over the prior three years). In the category with no DI receipt, there is a monotonic decline if one excludes the category for no earnings over the past three years.

Table 5. Joint Work Limitations-DI Receipt Models from SIPP: Logistic Coefficients and Standard Errors

Variables	Work Limited and Receiving DI *Versus* Neither Work Limited nor Receiving DI	Not Work Limited but Receiving DI *Versus* Neither Work Limited nor Receiving DI	Work Limited but Not Receiving DI *Versus* Neither Work Limited nor Receiving DI
Intercept	-7.8970*** (0.3616)	-5.6575*** (0.4666)	-3.3425*** (0.1512)
Age dummies (ref: 61-67)			
<=25	-1.6616*** (0.2390)	-2.0784*** (0.3168)	-1.1425*** (0.0994)
26-30	-1.0378*** (0.1552)	-1.6345*** (0.2281)	-0.8153*** (0.0761)
31-35	-0.8296*** (0.1309)	-1.2590*** (0.1974)	-0.5098*** (0.0653)
36-40	-0.4216*** (0.1175)	-0.6251*** (0.1710)	-0.4006*** (0.0635)
41-45	-0.1444 (0.1123)	-0.3229* (0.1646)	-0.1762** (0.0629)
46-50	0.0831 (0.1103)	-0.2947 (0.1706)	-0.1014 (0.0645)
51-55	0.4635*** (0.1066)	0.2098 (0.1554)	0.0472 (0.0661)
56-60	0.9352*** (0.0994)	0.2958 (0.1552)	0.1890** (0.0673)
Hispanic indicator	-0.0150 (0.1126)	-0.2925 (0.1760)	-0.1131 (0.0598)
Black indicator	-0.1287 (0.0978)	0.3373** (0.1278)	-0.0482 (0.0528)
Indicator work limited t-1	5.4937*** (0.0756)	2.9341*** (0.0940)	4.7131*** (0.0316)
Education dummies (ref: high school grad)			
Less than high school grad	0.4659*** (0.0696)	0.2562* (0.1050)	0.1687*** (0.0413)
Some college	-0.3067*** (0.0792)	-0.4099*** (0.1221)	-0.0604 (0.0372)
College graduate	-0.6313*** (0.0966)	-0.5619*** (0.1357)	-0.3558*** (0.0426)
Marital Status indicators (ref: married)			
Never married	0.9505*** (0.0888)	0.6695*** (0.1311)	0.2806*** (0.0448)
Divorced or separated	0.4807*** (0.0761)	0.2829* (0.1217)	0.1906*** (0.0428)
Widowed	0.3277* (0.1422)	0.2900 (0.2089)	0.2968*** (0.0869)
Homeownership indicator	0.0577 (0.0629)	-0.1824 (0.0941)	-0.0994** (0.0325)
Native born indicator	0.4433*** (0.1240)	-0.1395 (0.1587)	0.3005*** (0.0608)
Male indicator	0.3351*** (0.0589)	0.5809*** (0.0904)	0.1189*** (0.0305)
Indicator of impending death (within 3 years)	1.6083*** (0.1737)	2.1490*** (0.1836)	0.6456*** (0.1431)
Number of earnings years	0.1409** (0.0482)	0.1369* (0.0644)	-0.0017 (0.0192)
Number of earnings years^2	-0.0033 (0.0019)	-0.0067** (0.0026)	0.0003 (0.0008)
Recent (3-year) average earnings / average wage dummies ("RAE")(ref: 0.30 <= RAE < 0.70)			
RAE = 0.00	1.6334*** (0.1032)	1.3586*** (0.1458)	0.0696 (0.0726)
0.00 < RAE < 0.15	0.9209*** (0.0894)	0.8498*** (0.1328)	0.1680*** (0.0503)
0.15 <= RAE < 0.30	0.5453*** (0.0989)	0.2696 (0.1574)	0.1609** (0.0517)
0.70 <= RAE < 1.00	-0.5379*** (0.1138)	-0.5854*** (0.1726)	-0.1725*** (0.0484)
1.00 <= RAE < 1.30	-0.5636*** (0.1297)	-0.3749* (0.1800)	-0.3437*** (0.0563)
RAE >= 1.30	-0.9247*** (0.1150)	-0.9327*** (0.1689)	-0.4225*** (0.0478)
-2 log-likelihood		56,895.796	

Source: Urban Institute estimates from the 1990 through 1993 SIPP matched to the SSER, MBR, and Numident

Notes: *** indicates p< 0.001; ** indicates p< 0.01; * indicates p< 0.05; we define recent earnings variables using the average of the past three earnings years divided by the average wage

The models also reveal that these age and socioeconomic differentials appear to relate to the *severity* of the disability. One might consider that the two categories that include DI beneficiaries reflect more serious conditions than the category that includes individuals self-reporting work limitations but not DI receipt, given the stringent requirements for qualifying for DI benefits.[13] The impending mortality indicator is much larger in the two DI categories, compared to the work limitations only category. Socio-economic differentials are larger and more important in the categories for the stricter government disability definitions. For example, being in the lowest earnings category is positively, significantly associated with occupying both DI categories, but does not have a significant relationship to self-reported work limits alone.

Conclusions and Recommendations

Consistent with prior research, we have found strong linkages between work limitations self-reports, entry onto DI program roles, and mortality experience. If microsimulation models are to accurately capture the distributional consequences of Social Security reform, they must retain these important relationships.

How microsimulation model developers can best do this remains an open question. Indeed, the models that we have presented may raise as many questions as they answer. One specific question that looms large is whether developers should endeavor to model disability, work limitations, and mortality independently, for example through the separate equations that we present in Tables 1 through 4, or jointly, for example through a single system of equations similar to the one we present in Table 5.[14] Our joint model of work limitation and DI receipt suggests that at least two of these outcomes can be successfully integrated in a cross-sectional model. Unfortunately, though, the SIPP data cannot support this type of model in a longitudinal context, and a longitudinal specification is highly preferable in the dynamic microsimulation setting.[15]

In the recent update of the MINT model to MINT3, Burtless and Sahm (2002) contend with a similar problem by jointly modeling mortality, earnings, and Disability Insurance receipt using a splicing method. This method employs observed (i.e., historical) joint distributions of these trajectories for use as donors for projecting future trajectories among those whose experiences are censored. This approach has the advantage of being able to replicate patterns in rare occurrences and trajectories (for example, cases of recovery and re-entry to DI). A potential disadvantage to the method is that it can incorporate cohort effects only to the extent that compositional changes in the population (e.g., educational attainment and race/ethnicity)

---

[13] The Social Security Act defines a person as disabled if he or she is unable to engage in substantial gainful activity due to a permanent mental or physical impairment that should last at least twelve continuous months and/or end in death. In January 2001, evidence of substantial gainful activity is average monthly earnings of 740 dollars/month for a nonblind beneficiary or 1240 dollars monthly for a blind beneficiary.

[14] In theory, one could model all three outcomes simultaneously using a state space with four origins (the four states for our multinomial logit) and five destinations (the preceding four, plus death). This requires estimation of 20 transition probabilities or, assuming that exiting from DI does not occur for reasons other than death, 16 transition probabilities.

[15] Integrating terms for unobserved heterogeneity is also desirable in many dynamic microsimulation contexts. Unfortunately, data limitations prevented such a specification for these models.

generate these differences. This should be less of a problem the shorter one's projection time horizon.

Assuming that a series of independent equations is the most appropriate approach for a model with a long time horizon (like DYNASIM), we close by evaluating the specific models that we have estimated. The work limitations models from SIPP are the most clearly appropriate for integration into a dynamic microsimulation model as they stand. They are longitudinal, rely on sufficient numbers of observations, and show patterns consistent with the prior literature. The models of DI take-up are also suited fairly well to being integrated into a larger model, though the absence of an exit equation is a limitation.

While the mortality models that we present are consistent across data sources (NLMS and SIPP) and consistent with prior literature, they do not compare favorably with the duration-specific life tables in Zayatz (1999) in terms of their capability of capturing the impacts of DI. Thus even though the SIPP models show promise, we recommend continuing to rely on adaptations of the Zayatz lifetables for modeling the mortality of the disabled in DYNASIM. Updating the SIPP models to include data from the 1996 (and even 1984) SIPP, however, would substantially boost both the number of observations and our ability to reliably estimate a time trend.

Sources cited:

Aaron, Henry J. with the assistance of Philip Spevak. 1977. "Demographic Effects on the Equity of Social Security Benefits." In Martin S. Feldstein and Robert P. Inman, Editors. *The Economics of Public Services: Proceedings of a Conference Held by the International Economic Association at Turin, Italy*. London: MacMillan.

Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, John Rust, and Sofia Sheidvasser. 1999. "An Empirical Analysis of the Social Security Disability Application, Appeal, and Award Process." *Labour Economics* 6(2): 147-178.

Burkhauser, Richard V., Andrew J. Houtenville, and David C. Wittenburg. Forthcoming. "A User Guide to Current Statistics on the Employment of People With Disabilities" in *What is Causing the Decline in the Employment of People with Disabilities: A Policy Mystery,* Edited by David C. Stapleton and Richard V. Burkhauser.

Burtless, Gary. 1999. "Estimation and Projection of Lifetime Earnings." In Toder et al. (below).

Burtless, Gary and Claudia Sahm. 2002. "Projections of Earnings and Prevalence of Disability Entitlement." In Toder et al. (below).

Cohen, Lee et al. 1999. *CORSIM Evaluation [Draft]*. Washington, DC: Office of Research, Evaluation and Statistics, Social Security Administration.

Crimmins, Eileen M., Sandra L. Reynolds, and Yasuhiko Saito. 1999. "Trends in Health and Ability to Work among the Older Working-Age Population." *Journal of Gerontology: Social Sciences* 54B No. 1 (January): S31-S40.

Favreault, Melissa. 2000. Unpublished DYNASIM mortality estimates.

Favreault, Melissa and Douglas Wolf. 2002. "Supplemental Security Income and Living Arrangements." In Toder et al. (below).

Favreault, Melissa and Frank Sammartino. 2002. "Impact of Social Security Reform on Low-Income and Older Women." Washington, DC: The Urban Institute [Project Report for AARP, available at URL: http://research.aarp.org/econ/2002_11_ssreform.html (accessed September, 2002)].

Garrett, Daniel M. 1995. "The Effects of Differential Mortality Rates on the Progressivity of Social Security." *Economic Inquiry* 33(3): 457-475.

Hayward, Mark D. and Melonie Heron. 1999. "Racial Inequality in Active Life Among Adult Americans." *Demography* 36(1): 77-91.

Hurd, Michael D. and Daniel McFadden, and Angela Merrill. 1999. "Predictors of Mortality Among the Elderly." Cambridge, MA: National Bureau of Economic Research [Working Paper W7440].

Huynh, Minh, Kalman Rupp, and James Sears. 2002. "The Assessment of Survey of Income and Program Participation (SIPP) Benefit Data using Longitudinal Administrative Records." SIPP Working Paper No. 238. Washington, DC: U.S. Census Bureau.

Johnson, Jon, Richard Wertheimer, and Sheila R. Zedlewski. 1983. *The Dynamic Simulation of Income Model (DYNASIM): Volume 1 The Family and Earnings History Model*. Project Report No. 1434-03. Washington, DC: The Urban Institute.

Kitagawa, Evelyn M. and Philip M. Hauser. 1973. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Cambridge, MA: Harvard University Press.

Kreider, Brent and Regina T. Riphan. 2000. "Explaining Applications to the U.S. Disability System." *Journal of Human Resources* 35(1): 82-115.

Manton, Kenneth G., Larry Corder, and Eric Stallard. 1997. "Chronic Disability Tends in Elderly United States Populations: 1982-1994." *Proceedings of the National Academy of Sciences 94* No. 6 (March): 2593-8.

McCoy, John L., Howard M. Iams, and Timothy Armstrong. 1994. "The Hazard of Mortality Among Aging Retired and Disabled-Worker Men: A Comparative Sociodemographic and Health Status Analysis." *Social Security Bulletin* 57(3): 76-87.

Mitchell, Olivia S. and John W.R. Phillips. 2001. "Eligibility for Social Security Disability Insurance." Working Paper 2001-011. Ann Arbor: University of Michigan Retirement Research Center.

Nagi, S. Z. 1969. "Congruency in Medical and Self-Assessment of Disability." *Industrial Medicine* 38: 74-83.

Panis, Constantijn and Lee Lillard. 1999. "Demographic Projections." Paper presented at "New Developments in Retirement Research," First Annual Joint Conference for the Retirement Research Consortium, May 20-21, Washington, DC.

Rosenberg, Harry M., Jeffrey D. Maurer, Paul D. Sorlie, Normal L. Johnson, Marian F. MacDorman, Donna L. Hoyert, James F. Spitler, and Chester Scott. 1999. "Quality of Death Rates by Race and Hispanic Origin: A Summary of Current Research." *Vital and Health Statistics* 2(128). Hyattsville, MD: National Center for Health Statistics.

Rupp, Kalman and Paul S. Davies. 2002. "A Long-Term View of Health Status, Disabilities, Mortality, and Participation in the Social Security Administration's Disability Programs." Paper presented at the Population Association of America Annual Meeting (May).

Rust, John. 1999. "Comments on the completed MINT project," Correspondence to Howard Iams, Office of Research, Evaluation and Statistics, Social Security Administration, dated July.

Social Security Administration, Office of Research, Evaluation and Statistics. 2001. *Annual Statistical Supplement, 2001* to the *Social Security Bulletin*. Washington, DC: Author.

Sorlie, Paul D., Eric Backlund, Norman J. Johnson, and Eugene Rogot. 1993. "Mortality by Hispanic Status in the United States." *JAMA* 270(20): 2464-2468.

Sorlie, Paul D., Eugene Rogot, and Norman J. Johnson. 1992. "Validity of Demographic Characteristics on the Death Certificate." *Epidemiology* 3(2): 181-184.

Thompson, Lawrence and Karen Smith. 2002. "Summary of Model Results." In Toder et al. (below).

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: The Urban Institute [Project Report].

Toder, Eric, Lawrence 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 1931-60 Birth Cohorts.* Washington, DC: The Urban Institute [Project Report for the Social Security Administration].

Waidmann, Timothy. 2002. "Options for Projecting Job Demands and Health Status." In Toder et al. (above).

Zayatz, Tim. 1999. *Social Security Disability Insurance Program Worker Experience*. Actuarial Study No. 114, June. Baltimore: Office of the Chief Actuary of the Social Security Administration.

Zedlewski, Sheila R. 1990. "The Development of the Dynamic Simulation of Income Model (DYNASIM)." In Gordon H. Lewis and Richard C. Michel, Editors. *Microsimulation Techniques for Tax and Transfer Analysis*. Washington, DC: Urban Institute Press.

Appendix Table 1. Data Options for Modeling Socioeconomic Differentials in Mortality in Dynamic Microsimulation Models

Source	Number of Events	Period	Comments on strengths and weaknesses
NLMS	42,919 deaths on 3.68 million person years at risk; using a single year framework, which allows one to use time-varying (lagged) independent variables but which precludes estimating a time trend at the micro-level, 4,079 deaths over age 24 plus 125 under 25, on a total of 616,500 person years	1979-1981 to base year+9	Although NLMS tracks mortality for nine years after initial (CPS) survey, it captures the independent variables, including family income and disability, which it only crudely represents, at just one point in time (the time of survey); there may be less measurement error in identifying whether a death has occurred than in panel data, since statistical matches to administrative data are attempted; emigration could be a problem, particularly for Hispanics and the foreign born (one cannot distinguish emigration from remaining alive); difficult to accurately integrate period effects in a model based on these data because of confidentiality restrictions (users cannot distinguish 79 to 81 interviews); can only use cross-sectional income rather than permanent income; age is topcoded at 98 (may affect coefficient estimates for the highest age ranges); using a single year, data are somewhat dated (though calibration using Vital Statistics regression can help with this).
PSID	All: 2,213 deaths; with crude DI info: 1,355 deaths (total of 194 DI deaths); with more precise DI info: 693	1969-1993	Better at measuring time-varying independent variables (like earnings and/or permanent income), and capturing a time trend than NLMS (and, to a lesser extent, SIPP); has a limited capacity for DI interaction; some deaths may be mislabeled as attrition; sample is less representative than NLMS or SIPP, with particular trouble with immigrants.
SIPP	2,775 deaths reported on SIPP (probably more with numident)	1990-1996	Less capacity to capture long-term time-trend than in PSID, but has the most recent time pattern of the three sources; with link to administrative records, may be the only way to capture DI receipt and permanent income accurately; Numident link should lead to better classification of attrition and death than other surveys allow, but some matches are poor (e.g., multiple death dates are present or death occurs before survey in Numident); age is topcoded at 85, but using administrative records can uncap this.

Notes: Event counts are not restricted by age.

Appendix Table 2. Old Specification of the DYNASIM Disability Model: Contrasting PSID Coefficients with SIPP Replication

	Disability Entry		Disability Exit	
Variable	PSID version	SIPP version	PSID version	SIPP version
Intercept	-2.979**	-3.7594**	-0.510**	-1.7331**
Age < 35	Reference	Reference	0.639**	0.3613**
Age 35-44	0.344	0.3693**	Reference	Reference
Age 45-54	0.833**	0.6649**	-0.417**	-0.1166
Age 55-64	1.153**	1.1593**	-0.647**	-0.3195**
Age 65+	1.521**	1.2511**	-0.516**	0.4747**
Nonwhites	0.301**	0.2393**	--	--
Married Female	-0.306*	-0.2148**	--	--
Female	--	--	0.271**	-0.1835**
Nonwhite Female	--	--	-0.364**	0.0477
Education < 12 Years	--	--	-0.201*	-0.2693**

Sources: PSID version: Based on 1969 through 1972 data, as reported in Johnson, Wertheimer, and Zedlewski (1983); SIPP version: Author's calculations from 1990, 1992, and 1993 SIPP

Notes: ** indicates p < 0.01; * indicates p < 0.05

RECENT WORKING PAPERS FROM THE CENTER FOR RETIREMENT RESEARCH AT BOSTON COLLEGE

Job Search Behavior at the End of the Life Cycle
Hugo Benitez-Silva, December 2002

Forecasting Incidence of Work Limitations, Disability Insurance Receipt, And Mortality in Dynamic Simulation Models Using Social Security Administrative Records: A Research Note
Melissa Favreault, December 2002

Medicare, Retirement Costs, and Labor Supply at Older Ages
Richard W. Johnson, December 2002

The Rise in Disability Recipiency and the Decline in Unemployment
David H. Autor and Mark G. Duggan, September 2002

Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors
Brent Krieder and John V. Pepper, September 2002

The Effect of Unplanned Changes in Marital and Disability Status: Interrupted Trajectories and Labor Force Participation
John B. Williamson and Tay K. McNamara, September 2002

Health Insurance and the Disability Insurance Application Decision
Jonathan Gruber and Jeffrey Kubik, September 2002

Projecting Poverty Rates in 2020 for the 62 and Older Population: What Changes Can We Expect and Why?
Barbara A. Butrica, Karen Smith, and Eric Toder, September 2002

Social Security and the Private Pension System: The Significance of Integrated Plans
Pamela Perun, July 2002

Pension Reform in the Presence of Financial Market Risk
Barry Bosworth and Gary Burtless, July 2002

Why Some Workers Remain in the Labor Force Beyond the Typical Age of Retirement
John B. Williamson and Tay K. McNamara, November 2001

All working papers are available on the Center for Retirement Research website (http://www.bc.edu/crr) and can be requested by e-mail (crr@bc.edu) or phone (617-552-1762).