Volume Title: Topics in the Economics of Aging

Volume Author/Editor: David A. Wise, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-90298-6

Volume URL: http://www.nber.org/books/wise92-1

Conference Date: April 5-7, 1990

Publication Date: January 1992

Chapter Title: Three Models of Retirement: Computational Complexity versus Predictive Validity

Chapter Author: Robin L. Lumsdaine, James H. Stock, David A. Wise

Chapter URL: http://www.nber.org/chapters/c7097

Chapter pages in book: (p. 21 - 60)

1 Three Models of Retirement
Computational Complexity versus Predictive Validity

Robin L. Lumsdaine, James H. Stock, and David A. Wise

Empirical analysis often raises questions of approximation to underlying individual behavior. Closer approximation may require more complex statistical specifications. On the other hand, more complex specifications may presume computational facility that is beyond the grasp of most real people and therefore less consistent with the actual rules that govern their behavior, even though economic theory may lead analysts to increasingly complex specifications. Thus, the issue is not only whether more complex models are worth the effort but also whether they are better. The answer must necessarily depend on the behavior that the analysis is intended to predict. In this paper, we consider the relation between computational complexity and the predictive validity of three models of retirement behavior.

Retirement has been the subject of a large number of studies over the past decade. Most have emphasized the effect of Social Security provisions on retirement age, but a wide range of methods has been employed. The earlier studies in this time period were based on regression or multinomial logit analysis (see, e.g., Hurd and Boskin 1981). Subsequent analysis relied on nonlinear budget constraint formulations of the retirement decision (see, e.g., Burtless 1986; and Gustman and Steinmeier 1986) and on proportional hazard model formulations (see, e.g., Hausman and Wise 1985). More recently, sev-

Robin L. Lumsdaine is assistant professor of economics at Princeton University. James H. Stock is professor of political economy at the Kennedy School of Government, Harvard University, and a research associate of the National Bureau of Economic Research. David A. Wise is John F. Stambaugh Professor of Political Economy at the John F. Kennedy School of Government, Harvard University, and the area director for health and retirement programs at the National Bureau of Economic Research.

Financial support was provided by the National Institute on Aging (grants R37 AG08146 and T32 AG00186), the Hoover Institution, the National Science Foundation, and the Sloan Foundation. This chapter was written while Lumsdaine was a Ph.D. candidate in the Department of Economics, Harvard University.

eral authors have developed models that focus on forward-looking comparison of the advantages of retirement at alternative ages in the future and on the updating of information as persons age. Although the spirit of these latter models is basically the same, they vary widely in computational complexity. The potential advantages in predictive validity of the computationally more complex versions of these models are the primary motivation for this study, although to broaden the scope of the comparison we consider a much simpler model as well.

We compare the predictive validity of three models of retirement. The first is a simple probit model. The second is the "option value" model developed in Stock and Wise (1990a, 1990b). The third is a stochastic dynamic programming model. We experiment with two versions of this model: one is an adaptation of the extreme value distribution formulation proposed by Berkovec and Stern (1991), and the other is the normal distribution formulation proposed by Daula and Moffitt (1991). A related but still more complex model has been developed by Rust (1989), but we have not attempted to implement his formulation in the analysis in this paper.

The analysis is guided by several key ideas. First, all the models are theoretical abstractions; none of them can be reasonably thought of as "true." The important consideration is which decision rule is the best approximation to the calculations that govern actual individual behavior. In this paper, judgments on which rule is best are based on empirical evidence on the relation between model specification and predictive validity.

Second, the models vary substantially in the computational complexity of the decision rules that they attribute to individual decision makers. The option value and the dynamic programming rules are both intended to capture the same underlying idea, but implementation of dynamic programming rules typically implies considerably more computational complexity than implementation of the option value rule. The option value model makes a simplifying assumption that substantially reduces complexity. The probit model is much simpler than either of these.

Third, although the mathematically correct implementation of some decision rules requires dynamic programming, there is no single dynamic programming rule. The implied computational complexity depends in important ways on specific assumptions, in particular the disturbance term correlation structure. It is easier to incorporate more flexible correlation assumptions in the option value than in the dynamic programming formulations. Thus, for example, the option value specification may be a suboptimal solution to a dynamic programming rule that implies computational complexity difficult to implement even with a computer.

A question of practical importance is therefore whether different decision rules yield significantly different results.

The comparisons in this paper are made by estimating the models on the same data. The data, which pertain to the retirement decisions in a large *For-*

tune 500 firm, have two important advantages for our purposes. First, the retirement decisions can be related to the provisions of the firm's pension plan, and it is therefore possible to simulate the effect of changes in the pension plan provisions. Second, the firm offered an unanticipated "window" plan in one of the years covered by the data.

The principal measure of the predictive validity of the models is how well they predict the effects of the window plan. Like the typical defined benefit pension plan, this firm's plan provides substantial incentives to retire early. In addition, the window plan provided further incentive to retire early. Window plans, which have been offered by many firms in recent years, provide special bonuses to workers in a specific group—often defined by age, occupational group, or even a division within the firm—if the worker retires within a specified period of time, typically a year or less. The window plan allows a unique external test of the predictive validity of the models; it is possible to compare model predictions against actual retirement rates under the window plan.

We begin by obtaining model parameter estimates based on retirement decisions in a year (1980) prior to the window plan. We then use these estimates to predict retirement in a later year (1982) under the window plan. The estimates and predictions are based on male nonmanagerial employees.

A brief description of the firm plan, the special window plan, and the data is presented in section 1.1. A more detailed description, borrowed in large part from Lumsdaine, Stock, and Wise (1991), is provided in appendix A. The models that are compared are explained in section 1.2. The parameter estimates and window plan predictions are presented in section 1.3. Section 1.4 presents simulations of the effects of eliminating the Social Security early retirement option. Conclusions are presented in section 1.5.

1.1 The Data, the Firm Pension Plan, and the Temporary Window

The analysis is based on a random sample of 993 male nonmanagerial office employees at a *Fortune* 500 firm. They were employed at the firm and were at least 50 years old on 1 January 1980, and they had been employed by the firm for at least three years prior to 1980. (The criterion that they be employed three years facilitates the forecasting of future wage earnings on an individual basis.)[1]

The data, obtained from firm records, include the earnings history of each employee from his year of employment, or from 1969 if he was employed before then, to retirement, or to 1983 if he had not retired by then. The data allow determination of whether the employee continued to work at the firm in successive years from 1980 through 1984. The data do not include the employment status of workers who left the firm; some employees probably took another job after departure from this firm. Thus, strictly speaking, the data

1. Employees who died between 1980 and 1982 before retiring were not included in the sample.

pertain to departure from the firm rather than retirement, but because we have no information on postretirement employment, we treat departure as retirement.

The firm's employees are covered by a defined benefit pension plan. The plan provides a substantial incentive for the typical employee to remain in the firm until age 55 and then an additional incentive to leave the firm before age 65. The plan is described in detail in appendix A. It has four key features:

1. The "normal" retirement age is 65.
2. Workers are vested after ten years of service.
3. The early retirement age is 55: a worker who departs before age 55 receives benefits that are reduced actuarially (approximately 7 percent per year) from the normal retirement age benefits, but the benefits of an employee who retires at 55 or later are reduced only about 3 percent per year, thus creating an incentive to stay until 55 and then an incentive to leave the firm.
4. The benefit formula incorporates a Social Security offset—a reduction of firm benefits based on Social Security benefits—but the offset is waived until age 65 for persons who retire at 55 or later, thus creating an additional incentive for workers to retire between 55 and 65.

In addition, an employee accrues a benefit entitlement from Social Security, with early retirement at age 62 and normal retirement at 65.

Particularly important for this study is the firm's 1982 window plan. Under the window plan, the firm offered nonmanagerial office employees a temporary retirement incentive. The window plan applied to employees between 55 and 65 who were vested in the firm's pension plan and to all employees over 65. Employees who retired in 1982 were offered a bonus equivalent to 3–12 months' salary. Although the exact bonus varied by years of service, it was typically largest for employees who were between 58 and 62 years old and smallest for those 55 and 65.[2] Of the 993 employees in our sample, 800 remained in the firm until 1982. The actual 1982 departure rates of these 800 employees are used to assess the out-of-sample predictive validity of the three retirement models.

1.2 The Models

Three retirement models are described, beginning with the "option value" model. The simple probit model is explained next and then the dynamic programming specification.

2. For a detailed description of this window plan and a discussion of the design of efficient window plans, see Lumsdaine, Stock, and Wise (1990).

1.2.1 The Option Value Model

The conceptual model is discussed in some detail in Stock and Wise (1990b). It is described only briefly here. At any given age, it is assumed, on the basis of information available at that age, that an employee compares the expected present value of retiring at that age with the value of retiring at each age in the future through age 70, the mandatory retirement age in this firm. The maximum of the expected present values of retiring at each future age minus the expected present value of immediate retirement is called the option value of postponing retirement. A person who does not retire this year maintains the option of retiring at a more advantageous age later on. If the option value is positive, the person continues to work; otherwise, he retires. With reference to appendix figure 1A.1, for example, at age 50 the employee would compare the value of the retirement benefits that he would receive were he to retire then—approximately \$28,000—with the value of wage earnings and retirement benefits in each future year. The expected present value of retiring at 60 (discounted to age 50), for example, is about \$184,000. This calculation is repeated as the worker ages, using updated predictions of future wage earnings and related pension and Social Security benefits. Future earnings forecasts are based on the individual's past earnings as well as on the earnings of other persons in the firm.[3] The precise model specification follows.

A person at age t who continues to work will earn Y_s in subsequent years s. If the person retires at age r, subsequent retirement benefits will be $B_s(r)$. These benefits will depend on the person's age and years of service at retirement and on his earnings history; thus, they are a function of the retirement age. We suppose that, in deciding whether to retire, the person weighs the indirect utility that will be received from future income. Discounted to age t at the rate β, the value of this future stream of income if retirement is at age r is given by

$$(1) \qquad V_t(r) = \sum_{s=t}^{r-1} \beta^{s-t} U_w(Y_s) + \sum_{s=r}^{S} \beta^{s-t} U_r[B_s(r)],$$

where $U_w(Y_s)$ is the indirect utility of future wage income and $U_r[B_s(r)]$ is the indirect utility of future retirement benefits. It is assumed that the employee will not live past age S.

The gain, evaluated at age t, from postponing retirement until age r is given by

$$(2) \qquad G_t(r) = E_t V_t(r) - E_t V_t(t).$$

Letting r^* be the age that gives the maximum gain, the person will postpone retirement if the option value, $G_t(r^*)$, is positive,

$$(3) \qquad G_t(r^*) = E_t V_t(r^*) - E_t V_t(t) > 0.$$

3. For a description of the earnings forecasts, see Stock and Wise (1990b).

The utilities of future wage and retirement income are parameterized as

$$U_w(Y_s) = Y_s^\gamma + \omega_s, \tag{4a}$$

$$U_r(B_s) = [kB_s(r)]^\gamma + \xi_s, \tag{4b}$$

where ω_s and ξ_s are individual-specific random effects, assumed to follow a Markovian (first-order autoregressive) process

$$\omega_s = \rho\omega_{s-1} + \varepsilon_{\omega s}, \quad E_{s-1}(\varepsilon_{\omega s}) = 0, \tag{5a}$$

$$\xi_s = \rho\xi_{s-1} + \varepsilon_{\xi s}, \quad E_{s-1}(\varepsilon_{\xi s}) = 0. \tag{5b}$$

The parameter k is to recognize that, in considering whether to retire, the utility *associated* with a dollar of income while retired may be different from the utility associated with a dollar of income accompanied by work. Abstracting from the random terms, at any given age s, the ratio of the utility of retirement to the utility of employment is $[k(B_s/Y_s)]^\gamma$.

Given this specification, the function $G_t(r)$ can be decomposed into two components:

$$G_t(r) = g_t(r) + \phi_t(r), \tag{6}$$

where $g_t(r)$ and $\phi_t(r)$ distinguish the terms in $G_t(r)$ containing the random effects, ω and ξ, from the other terms. If whether the person is alive in future years is statistically independent of his earnings stream and the individual effects ω_s and ξ_s, $g_t(r)$ and $\phi_t(r)$ are given by

$$\begin{aligned} g_t(r) = {} & \sum_{s=t}^{r-1} \beta^{s-t}\pi(s|t)E_t(Y_s^\gamma) + \sum_{s=r}^{S} \beta^{s-t}\pi(s|t)\{E_t[kB_s(r)]^\gamma\} \\ & - \sum_{s=t}^{S} \beta^{s-t}\pi(s|t)\{E_t[kB_s(t)]^\gamma\}, \end{aligned} \tag{7a}$$

$$\phi_t(r) = \sum_{s=t}^{r-1} \beta^{s-t}\pi(s|t)E_t(\omega_s - \xi_s), \tag{7b}$$

where $\pi(s|t)$ denotes the probability that the person will be alive in year s, given that he is alive in year t. Given the random Markov assumption, $\phi_t(r)$ can be written as

$$\begin{aligned} \phi_t(r) &= \sum_{s=t}^{r-1} \beta^{s-t}\pi(s|t)\rho^{s-t}(\omega_t - \xi_t) \\ &= K_t(r)\nu_t, \end{aligned} \tag{8}$$

where $K_t(r) = \Sigma_{s=t}^{r-1}(\beta\rho)^{s-t}\pi(s|t)$ and $\nu_t = \omega_t - \xi_t$. The simplification results from the fact that at time t the expected value of $\nu_s = \omega_s - \xi_s$ is $\rho^{s-t}\nu_t$, for all future years s. (The term $K_t[r]$ cumulates the deflators that yield the present value in year t of the future expected values of the random components of utility. The further r is in the future, the larger is $K_t[r]$. That is, the more distant the potential retirement age, the greater the uncertainty about it, yielding a heteroskedastic disturbance term.) $G_t(r)$ may thus be written simply as

$$G_t(r) = g_t(r) + K_t(r)\nu_t. \tag{9}$$

If the employee is to retire in year t, $G_t(r)$ must be less than zero for every potential retirement age r in the future. If $r_t^\dagger$ is the r that yields the maximum value of $g_t(r)/K_t(r)$, the probability of retirement becomes

$$\text{(10)} \qquad \Pr[\text{retire in year } t] = \Pr[g_t(r_t^\dagger)/K_t(r_t^\dagger) < -\nu_t].$$

If retirement in only one year is considered, this expression is all that is needed.

More generally, retirement decisions may be considered over two or more consecutive years. In this case, the retirement probabilities are simply an extension of equation (10). The probability that a person who is employed at age t will retire at age $\tau > t$ is given by

$$\text{(11)} \qquad \begin{aligned} \Pr[\text{retire in year } \tau] = \Pr[&g_t(r_t^\dagger)/K_t(r_t^\dagger) > -\nu_t, \ldots, \\ &g_{\tau-1}(r_{\tau-1}^\dagger)/K_{\tau-1}(r_{\tau-1}^\dagger) > -\nu_{\tau-1}, \\ &g_\tau(r_\tau^\dagger)/K_\tau(r_\tau^\dagger) < -\nu_\tau]. \end{aligned}$$

The probability that the person does not retire during the period covered by the data is given by

$$\text{(12)} \qquad \begin{aligned} \Pr[\text{do not retire by year } T] = \Pr[&g_t(r_t^\dagger)/K_t(r_t^\dagger) > -\nu_t, \ldots, \\ &g_{T-1}(r_{T-1}^\dagger)/K_{T-1}(r_{T-1}^\dagger) > -\nu_{T-1}, \\ &g_T(r_T^\dagger)/K_T(r_T^\dagger) > -\nu_T], \end{aligned}$$

where T is the final period in the data set. This is a multinomial discrete choice probability with dependent error terms ν_s.

Finally, we assume that ν_s follows a Gaussian Markov process, with

$$\text{(13)} \qquad \nu_s = \rho\nu_{s-1} + \varepsilon_s, \quad \varepsilon_s \ \textit{i.i.d.}\ N(0, \sigma_\varepsilon^2),$$

where the initial value, ν_t, is *i.i.d.* $N(0, \sigma^2)$ and is independent of ε_s. The covariance between ν_τ and $\nu_{\tau+1}$ is $\rho\text{var}(\nu_\tau)$, and the variance of ν_τ for $\tau > t$ is $[\rho^{2(\tau-t)}]\sigma^2 + (\Sigma_{j=0}^{\tau-t-1}\rho^{2j})\sigma_\varepsilon^2$.

The estimates in this paper are based on retirement decisions in only one year, and the random terms in equation (5) are assumed to follow a random walk, with $\rho = 1$. In this case, the covariance between ν_τ and $\nu_{\tau+1}$ is $\text{var}(\nu_\tau)$, and the variance of ν_τ for $\tau \geq t$ is $\sigma^2 + (\tau - t)\sigma_\varepsilon^2$. Prior estimates show that one- and multiple-year estimates are very similar.[4]

1.2.2 The Probit Model

The option value model proposes that a person will continue to work if the option value of postponing retirement—given by $G_t(r^*) = E_tV_t(r^*) - E_tV_t(t)$ in equation (3)—is greater than zero. In that model, the option value is determined by estimation. That is, the observed retirement decisions are described in terms of $\Pr[G_t(r^*) > 0]$, which in turn is described by a particular parame-

4. Estimates based on several consecutive years and with ρ estimated are shown in Stock and Wise (1990b). These generalizations have little effect on the estimates.

terization of $V_t(r)$. The maximum likelihood estimation procedure determines these parameters—γ, k, β, and σ (and σ_ε if two or more consecutive years are used in estimation). Thus, one can think of this procedure as estimating the option value on the basis of how employees value future income and leisure.

An alternative approach is to specify retirement in terms of the gain from continuing to work but to calculate the gain on the basis of an assumed valuation of income (determined by γ and k) and an assumed discount rate (β) instead of estimating them. Assuming that retirement depends on this *calculated* option value as well as other unobserved determinants of retirement, a standard specification of retirement is

$$\text{(14)} \qquad \Pr[\text{retire in year } t] = \Pr[\delta_0 + \delta_1 \hat{G}_t(r^*) + \varepsilon > 0],$$

where $\hat{G}_t(r^*)$ is the option value calculated under the presumed parameter values, and assuming the random components of $G_t(r)$ ($\phi_t[r]$ in [6] and [7b]) are all zero. This is a probit formulation, assuming that ε has a normal distribution.

In this case, the effect of the assumed gain from retirement is estimated by the parameter δ_1. This formulation is the closest probit counterpart to the option value model. In addition to this specification, several others are also estimated. The alternative specifications predict retirement on the basis of Social Security (SS) benefits, pension benefits, the present value of SS benefits (SS wealth), the present value of pension benefits (pension wealth), the change in the present value of SS benefits from working another year (SS accrual), the change in the present value of pension benefits from working another year (pension accrual), predicted earnings in the next year, and age.

1.2.3 The Stochastic Dynamic Programming Model

The key simplifying assumption in the Stock-Wise option value model is that the retirement decision is based on the maximum of the expected present values of future utilities if retirement occurs now versus each of the potential future ages. The stochastic dynamic programming rule considers instead the expected value of the maximum of current versus future options. The expected value of the maximum of a series of random variables will be greater than the maximum of the expected values. Thus, to the extent that this difference is large, the Stock-Wise option value rule underestimates the value of postponing retirement. And to the extent that the dynamic programming rule is more consistent with individual decisions than the option value rule, the Stock-Wise rule may undervalue individual assessment of future retirement options. Thus, we consider a model that rests on the dynamic programming rule.

As emphasized above, it is important to understand that there is no single dynamic programming model. Because the dynamic programming decision rule evaluates the maximum of future disturbance terms, its implementation depends in important ways on the error structure that is assumed. Like other users of this model, we assume an error structure—and thus a behavioral rule—that simplifies the dynamic programming calculation. In particular, although the option value model allows correlated disturbances, the random disturbances in the dynamic programming model are assumed to be uncorrelated, except for a random individual effect that is used in some specifications. Thus, the two models are not exactly comparable. Whether one rule is a better approximation to reality than the other may depend not only on the basic idea but also on its precise implementation.

In fact, we implement two versions of the dynamic programming model. In the first model, disturbance terms are assumed to follow an extreme value distribution. This model is adopted from Berkovec and Stern (1991), with two modifications. First, Berkovec and Stern consider three outcomes (full-time work, part-time work, and retirement), whereas we consider only two (full-time work and retirement, the only states for which we have data). Second, the way that we account for individual-specific effects differs from Berkovec and Stern's formulation.

In the second dynamic programming model, the disturbances are assumed to be normally distributed. This formulation is adopted from Daula and Moffitt's (1991) dynamic programming model of retention in the military. Our model generalizes their specification by allowing for additive individual-specific disturbances and by specifying retirement in terms of a parameterized utility function. With the additional assumption that the unobserved individual-specific effects are normally distributed across employees, the error structure in this dynamic programming specification is similar to the structure in the option value model. In both cases, future errors are normally distributed with nonzero covariances. In the option value model, the covariance structure derives from the random walk assumption; in the dynamic programming model, the covariances derive from a components-of-variance structure, with an individual-specific effect.

A more general dynamic programming model of retirement has been developed by Rust (1989). Unfortunately, comparison with his model is beyond the scope of this study. He assumes that an employee optimizes jointly over both age of retirement and future consumption. By admitting continuous and discrete choice variables, his model poses substantially greater numerical complexity than the ones we implement.

In most respects, our dynamic programming model is analogous to the option value model. As in that model, at age t an individual is assumed to derive utility $U_w(Y_t) + \varepsilon_{1t}$ from earned income or $U_r[B_t(s)] + \varepsilon_{2t}$ from retirement benefits, where s is the retirement age. The disturbances ε_{1t} and ε_{2t} are random perturbations to these age-specific utilities. Unlike the additive disturbances in the option value model, these additive disturbances in the dynamic programming model are assumed to be independent. Future income and retirement benefits are assumed to be nonrandom; there are no errors in forecasting future wage earnings or retirement benefits.

Individuals will presumably have different preferences for employment versus retirement. Variation in preferences is allowed for in the extreme value distribution version of our model by including individual-specific effects in $U_r(\cdot)$ and $U_w(\cdot)$. They are assumed to be fixed for each person, but they vary randomly from person to person. Berkovec and Stern modeled these individual-specific effects as additional additive errors. In the extreme value distribution version of our model, they enter multiplicatively. In the normal distribution version of our model, the random effects enter additively, as explained below.

The Model

The dynamic programming model is based on the recursive representation of the value function. At the beginning of year t, the individual has two choices: retire now and derive utility from future retirement benefits, or work for the year and derive utility from income while working during the year and retaining the option to choose the best of retirement or work in the next year. Thus, the value function W_t at time t is defined as

$$W_t = \max\left[E_t[U_w(Y_t) + \varepsilon_{1t} + \beta W_{t+1}],\ E_t\left(\sum_{\tau=t}^{S}\beta^{\tau-t}\{U_r[B_\tau(t)] + \varepsilon_{2\tau}\}\right)\right], \tag{15}$$

with

$$W_{t+1} = \max\left[E_{t+1}[U_w(Y_{t+1}) + \varepsilon_{1t+1} + \beta W_{t+2}],\ E_{t+1}\left(\sum_{\tau=t+1}^{S}\beta^{\tau-t-1}\{U_r[B_\tau(t+1)] + \varepsilon_{2\tau}\}\right)\right],$$

where β is the discount factor and, as in the option value model, S is the year beyond which the person will not live.

Because the errors ε_{it} are assumed to be i.i.d., $E_t\varepsilon_{it+\tau} = 0$ for $\tau > 0$. In addition, in computing expected values, each future utility must be discounted by the probability of realizing it, that is, by the probability of surviving to year τ given that the worker is alive in year t, $\pi(\tau|t)$. With these considerations, the expression (15) can be written as

$$\begin{aligned} W_t &= \max(\bar{W}_{1t} + \varepsilon_{1t}, \bar{W}_{2t} + \varepsilon_{2t}), \text{ where} \\ \bar{W}_{1t} &= U_w(Y_t) + \beta\pi(t+1|t)E_tW_{t+1}, \\ \bar{W}_{2t} &= \sum_{\tau=t}^{S}\beta^{\tau-t}\pi(\tau|t)U_r[B_\tau(t)]. \end{aligned} \tag{16}$$

The worker chooses to retire in year t if $\bar{W}_{1t} + \varepsilon_{1t} < \bar{W}_{2t} + \varepsilon_{2t}$; otherwise, he continues working. The probability that the individual retires is $\Pr(\bar{W}_{1t} + \varepsilon_{1t} < \bar{W}_{2t} + \varepsilon_{2t})$. If a person works until the mandatory retirement age (70), he retires and receives expected utility $\bar{W}_{2t_{70}}$.

Recursions and Computation

With a suitable assumption on the distribution of the errors ε_{it}, the expression (16) provides the basis for a computable recursion for the nonstochastic terms $\bar{W}_{it}$ in the value function. The extreme value and normal distribution versions of the model are considered in turn.

Extreme Value Errors. Following Berkovec and Stern (1991), the ε_{it} are assumed to be i.i.d. draws from an extreme value distribution with scale parameter σ. Then, for the years preceding mandatory retirement, these assumptions together with equation (16) imply that

$$
\begin{aligned}
& E_t W_{t+1}/\sigma \equiv \mu_{t+1} \\
(17) \quad & = \gamma_e + \ln[\exp(\bar{W}_{1t+1}/\sigma) + \exp(\bar{W}_{2t+1}/\sigma)] \\
& = \gamma_e + \ln\{\exp[U_w(Y_{t+1})/\sigma]\exp[\beta\pi(t+2|t+1)\mu_{t+2}] \\
& \quad + \exp(\bar{W}_{2t+1}/\sigma)\},
\end{aligned}
$$

where γ_e is Euler's constant. Thus, (17) can be solved by backward recursion, with the terminal value coming from the terminal condition that $\mu_{t_{70}} = \bar{W}_{2t_{70}}$.

The extreme value distributional assumption provides a closed form expression for the probability of retirement in year t:

$$
\begin{aligned}
(18) \quad \Pr[\text{retire in year } t] & = \Pr[\bar{W}_{1t} + \varepsilon_{1t} < \bar{W}_{2t} + \varepsilon_{2t}] \\
& = \exp(\bar{W}_{2t}/\sigma)/[\exp(\bar{W}_{1t}/\sigma) + \exp(\bar{W}_{2t}/\sigma)].
\end{aligned}
$$

Gaussian Errors. Following Daula and Moffitt (1991), the ε_{it} are assumed to be independent draws from an $N(0, \sigma^2)$ distribution. The Gaussian assumption provides a simple expression for the probability of retiring:

$$
\begin{aligned}
(19) \quad \Pr[\text{retire in year } t] & = \Pr[(\varepsilon_{1t} - \varepsilon_{2t})/\sqrt{2}\sigma \\
& < (\bar{W}_{2t} - \bar{W}_{1t})/\sqrt{2}\sigma] = \Phi(a_t),
\end{aligned}
$$

where $a_t = (\bar{W}_{2t} - \bar{W}_{1t})/\sqrt{2}\sigma$. Then the recursion (16) becomes:

$$
\begin{aligned}
(20) \quad E_t W_{t+1}/\sigma \equiv \mu_{t+1} & = (\bar{W}_{1t+1}/\sigma)[1 - \Phi(a_{t+1})] \\
& + (\bar{W}_{2t+1}/\sigma)\Phi(a_{t+1}) + \sqrt{2}\phi(a_{t+1}),
\end{aligned}
$$

where $\phi(\cdot)$ denotes the standard normal density and $\Phi(\cdot)$ denotes the cumulative normal distribution function. As in (19), $\Phi(a_t)$ is the probability that the person retires in year t and receives utility $\bar{W}_{2t}$ plus utility from $E(\varepsilon_{2t} \mid \varepsilon_{1t} - \varepsilon_{2t} < \bar{W}_{2t} - \bar{W}_{1t})$. The latter term, plus a comparable term when the person continues to work, yields the last term in equation (20).

Individual-Specific Effects

Individual-specific terms are modeled as random effects but are assumed to be fixed over time for a given individual. They enter the two versions of the dynamic programming models in different ways. Each is discussed in turn.

Extreme Value Errors. Single year utilities are

$$U_w(Y_t) = Y_t^\gamma, \tag{21a}$$

$$U_r[B_t(s)] = [\eta k B_t(s)]^\gamma, \tag{21b}$$

where ηk is constant over time for the same person but random across individuals. Specifically, it is assumed that η is a lognormal random variable with mean one and scale parameter λ: $\eta = \exp(\lambda z + \frac{1}{2}\lambda^2)$, where z is i.i.d. $N(0, 1)$. A larger λ implies greater variability among employee tastes for retirement versus work; when $\lambda = 0$, there is no variation, and all employees have the same taste.

Normal Errors. In this case, the unobserved individual components are assumed to enter additively, with

$$U_w(Y_t) = Y_t^\gamma + \zeta, \tag{22a}$$

$$U_r[B_t(s)] = [kB_t(s)]^\gamma, \tag{22b}$$

where γ and k are nonrandom parameters, as above, but ζ is a random additive taste for work, assumed to be distributed $N(0, \lambda^2)$. When $\lambda = 0$, there is no taste variation.

To summarize, the dynamic programming models are given by the general recursion equation (15). It is implemented as shown in equation (17) under the assumption that the ε_{it} are i.i.d. extreme value and as shown in equation (20) under the assumption that ε_{it} are i.i.d. normal. The retirement probabilities are computed according to equations (18) and (19), respectively. The fixed effects specifications are given by equations (21) and (22). The unknown parameters to be estimated are (γ, k, β, σ, λ). Because of the different distributional assumptions, the scale parameter σ is *not* comparable across option value or dynamic programming models, and λ is *not* comparable across the two dynamic programming models.

1.3 Results

The option value and the dynamic programming specifications yield quite similar results, and both provide rather good predictions of retirement behavior under the window plan. The probit specifications yield very poor predictions of retirement under the window plan, although some specifications fit the sample data well. The parameter estimates are discussed first, together with standard measures of fit. We then graphically describe the correspondence between predicted versus actual retirement behavior, with emphasis on out-of-sample predictions of retirement under the 1982 window plan.

1.3.1 Parameter Estimates

The Probit Model

The parameter estimates for several probit specifications are shown in table 1.1. The variables are defined as follows:

Option value: $G_t(r^*)$ calculated as described in section 1.2.1 with $\gamma \equiv 1$, $k \equiv 1$, and $\beta \equiv .95$.
Age: Age in years.
Income: The predicted wage earnings in the following year if the person continues to work.
SS pv (present value): The predicted present value of entitlement to future SS benefits, were the person to retire at the beginning of the year, SS wealth.
Pension pv (present value): The predicted present value of entitlement to future firm pension benefits, were the person to retire at the beginning of the year, pension wealth.
SS accrual: The predicted change in the present value of entitlement to future SS benefits, were the person to continue to work for another year.
Pension acc (accrual): The predicted change in the present value of entitlement to future firm pension benefits, were the person to continue to work for another year.

The parameter estimates are with respect to the probability that a person will *retire*. Thus, the negative option value coefficient in specification 1 indicates that, the greater the option value of continuing to work, the less likely the person is to retire. To interpret this specification, recall that the principal difference between this probit specification and the option value model is the use of assumed parameter values to calculate the option value variable used in the probit model. If this probit specification were estimated using the optimized option value model parameters discussed below (see table 1.2), and if the intercept were forced to be zero, then the probit model would essentially reproduce the option value model, except for the heteroskedastic disturbance term incorporated in the option value model.

The addition of age (specification 2) substantially improves the model fit, but, as is shown in the graphic comparison below, this specification has little behavioral relevance.

Specifications 3–9 are intended to parallel the specification used by Hausman and Wise (1985) in their proportional hazard model of retirement. The probit model is a one-period counterpart to the Hausman and Wise analysis that followed older workers for ten years, covering five two-year periods. Their analysis relied solely on SS wealth and SS accrual (plus other personal attributes), however; they had no firm pension data. Specification 8 shows that

Table 1.1 **Probit Parameter Estimates**

	Specification									
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Constant	−.38	−7.18	−1.00	−.82	−1.10	−.61	−.76	−.93	−.71	−1.83
	(.11)	(1.09)	(.10)	(.09)	(.10)	(.11)	(.10)	(.48)	(.12)	(.24)
Option value	−.68	−.30								
	(.09)	(.09)								
Age		.11								
		(.02)								
Income			−.70	−5.11	−5.07	−1.71	−1.81	−2.66	−3.21	−.94
			(.28)	(.70)	(.70)	(.34)	(.33)	(.79)	(.76)	(.31)
SS pv				.69				.90		2.79
				(.08)				(1.09)		(.71)
Pension pv				1.39				.32	.53	
				(.18)				(.26)	(.24)	
SS + Pen pv					1.36					
					(.17)					
SS accrual						−26.47		−21.43		−27.54
						(2.44)		(8.64)		(5.68)
Pension acc						−10.65		−8.86	−7.59	
						(1.18)		(1.73)	(1.58)	
SS + Pen acc							−10.69			
							(1.11)			
Summary statistics:										
$-\ln \mathcal{L}$	299.22	277.75	339.69	298.52	298.38	282.62	284.22	281.38	284.85	329.98
χ^2 sample	59.1	35.5	179.5	68.6	65.3	29.1	31.1	28.2	38.2	145.9
χ^2 window	180.3	108.2	512.2[a]	191.2	164.9	76.4	75.8	67.5	57.3	229.7[a]

Note: Estimation is by maximum likelihood. All monetary values are in \$100,000 (1980 dollars). The χ^2 sample statistic is the chi-square statistic relative to the predicted vs. the actual number of retirements by age in the estimation sample; the χ^2 window statistic is the corresponding statistic for predicted vs. actual retirement under the window plan. Standard errors are in parentheses.

[a]The window plan bonus is treated as a one-time addition to income.

Table 1.2 **Parameter Estimates for the Option Value and the Dynamic Programming Models**

	Option Value Models		Dynamic Programming Models					
			Extreme Value			Normal		
Parameter	(1)	(2)	(1)	(2)	(3)	(4)	(5)	(6)
γ	1.00[a]	.612	1.00[a]	1.018	1.187	1.00[a]	1.187	1.109
		(.072)		(.045)	(.215)		(.110)	(.275)
k	1.902	1.477	1.864	1.881	1.411	2.592	2.975	2.974
	(.192)	(.445)	(.144)	(.185)	(.307)	(.100)	(.039)	(.374)
β	.855	.895	.618	.620	.583	.899	.916	.920
	(.046)	(.083)	(.048)	(.063)	(.105)	(.017)	(.013)	(.023)
σ	.168	.109	.306	.302	.392	.224	.202	.168
	(.016)	(.046)	(.037)	(.036)	(.090)	(.021)	(.022)	(.023)
λ			.00[a]	.00[a]	.407	.00[a]	.00[a]	.183
					(.138)			(.243)
Summary statistics:								
$-\ln \mathcal{L}$	294.59	280.32	279.60	279.57	277.25	277.24	276.49	276.17
χ^2 sample	36.5	53.5	38.9	38.2	36.2	45.0	40.7	41.5
χ^2 window	43.9	37.5	32.4	33.5	33.4	29.9	25.0	24.3

Note: Estimation is by maximum likelihood. The option value model is described in sec. 1.2.1, and the stochastic dynamic programming model is described in sec. 1.2.3. All monetary values are in $100,000 (1980 dollars). See the note to table 1.1.

[a]Parameter value imposed.

both SS and pension *accrual* are associated with continued employment, but the estimated coefficients would suggest substantial difference in the magnitude of the effects; the SS accrual coefficient is two and a half times as large as the pension coefficient (−21.43 vs. −8.64). (When the SS and the pension wealth and accrual variables are combined [specifications 5 and 7], however, the estimated effects are much closer to the pension than the SS effects.) Neither the SS nor the pension wealth coefficient is significantly different from zero, although both are positive.

The exclusion of the SS variables has little effect on the estimated effects of pension wealth and accrual (specification 9 vs. 8), but the exclusion of the pension variables has a substantial effect on the estimated SS effects (specification 10 vs. 8). This suggests that other estimates of the effects of SS on retirement, such as those in Hausman and Wise, may be biased because they do not control for firm pension benefits. Hausman and Wise, for example, find a strong estimated effect of both SS present value and SS accrual, but they do not have data on the corresponding pension values. In addition, the χ^2 sample statistics show that the specifications with the pension variables fit the sample data much better than the specification with only SS variables (specifications 8 and 9 vs. 10). And with only SS variables the effect of the window plan cannot be predicted, except by assuming that the effect of pension accrual or wealth is the same as the corresponding SS effect. Specification 8 shows that this is far from accurate in this case.

Higher expected wage earnings prolong labor force participation, according to these results.

Likelihood values and two χ^2 statistics are shown at the bottom of table 1.2. Aside from the specification that explicitly includes age, the highest likelihood value is obtained using expected wage earnings for the coming year and SS and pension wealth and accruals (specification 8). The sample χ^2 statistic compares predicted versus actual departure rates by age on the basis of the 1980 data used in the estimation. The window χ^2 statistic compares predicted versus actual departure rates by age under the 1982 window plan.

The Option Value Model

Parameter estimates from the option value model are shown in the first two columns of table 1.2. The income parameter γ (the risk aversion parameter in $U_w[Y_s] = Y_s^\gamma + \omega_s$) is 0.612, suggesting essentially risk neutral preferences. The estimated value of k in $U_r(B_s) = [kB_s(r)]^\gamma + \xi_s$ is 1.477, implying that a dollar without working is worth more than a dollar with work, although the estimate is not significantly different from one. The estimated value of β, 0.895, suggests that future expected or promised income is rather highly discounted relative to income now.

Dynamic Programming Model

The estimated parameters based on the dynamic programming decision rule are shown in the remaining columns of table 1.2. In general, the estimates are similar to those based on the option value rule. The estimated value of γ in the extreme value version (specification 2) is close to, and not significantly different from, one, implying that individuals are risk neutral (that utility is linear in income). The normal version (specification 5) also yields an estimated γ that is not significantly different from one but is substantially larger than the option value estimate (1.19 vs. 0.61). Like the option value results, the dynamic programming results suggest that the value of income together with retirement is substantially greater than the value of income together with work, although the dynamic programming models yield larger estimated values of k. And, like the option value estimates, the dynamic programming estimates indicate that future income is substantially discounted relative to current income in the determination of retirement. The normal specifications yield discount factors close to the option value estimates; the extreme value specification implies larger discount rates.

Estimates of the models including random individual components are reported as specifications 3 and 6. In neither case does inclusion of random individual effects significantly affect other parameter estimates. In the normal version, the variance of the individual effect is not significantly different from zero, implying no variation in taste for retirement versus work among these employees. The extreme value version suggests variation that is significantly different from zero, and the specification fits the data somewhat better than the specifications without the individual component. In neither case does the individual component noticeably improve the prediction of the window plan effects.

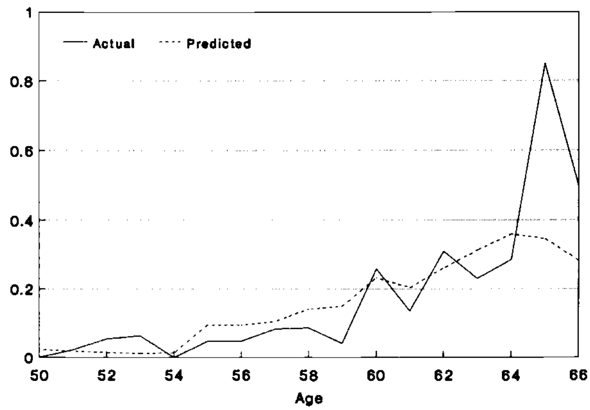
Based on the likelihood values, the more forward-looking models fit the data better than the probit specifications, with the exception of the probit with age. Overall, there is little difference in the likelihood values of the option value and the dynamic programming specifications.

The most informative χ^2 statistics pertain to the prediction of departure rates under the 1982 window plan. In this case, the forward-looking models predict actual departure rates substantially better than the probit specifications.

1.3.2 Graphic Comparisons

The Option Value versus Dynamic Programming Results

The easiest way to compare the models is by graphing their implied departure rates. The option value results (model 2 in table 1.2) are used as a base for comparison, and the relevant results are shown in figures 1.1a and 1.1b. Figure 1.1a shows the within-sample fit. Departure (hazard) rates by age are shown in the top panel. The cumulative departures implied by the departures by age are shown in the bottom panel. For example, according to the observed departure rates, 72.0 percent of persons employed at age 50 would have left the firm by age 62; based on the predicted departure rates, the cumulative percentage is 77.7. In general, the predicted departure rates correspond



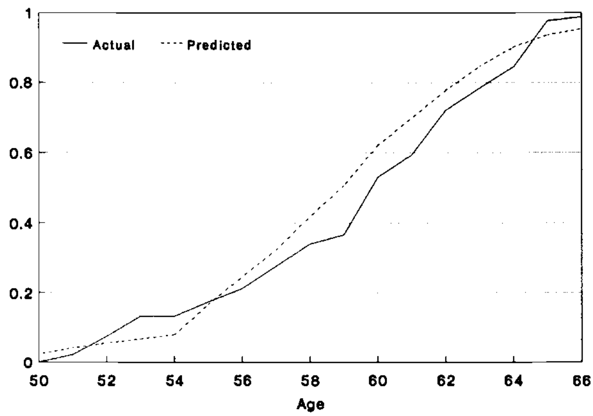


Fig. 1.1a Predicted vs. actual 1980 departure rates and implicit cumulative departures, by age: option value model 2

closely to the actual rates. For example, like the actual rates, the predicted rates show substantial jumps at 55, 60, and 62, all of which correspond to specific pension plan and SS provisions as described in appendix A. A noticeable exception occurs at age 65; among the small proportion of employees still in the firm at that age, a much larger proportion leaves the firm than the model predicts. This finding is common to all employee groups and to all versions of the option value model that we have estimated to date. It is apparently due to an "age-65-retirement effect" that is unrelated to earnings or retirement benefits.

As a test of the predictive validity of the model, the estimates based on 1980 departure rates have been used to predict departure rates under the 1982 window plan. The departure rates of persons offered the window plan bonus were typically about twice as high as they were without this special incentive. Predicted versus actual rates under the window plan are shown in figure 1.1b,

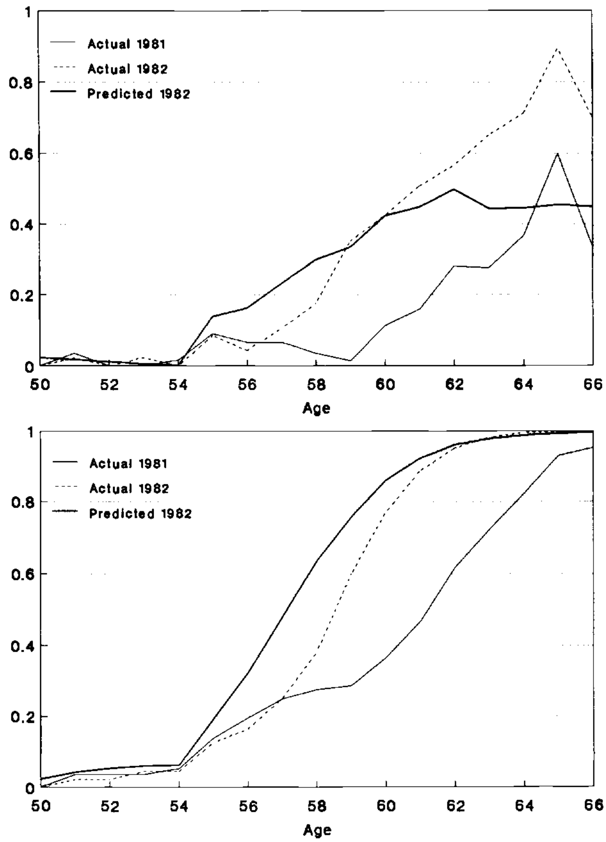


Fig. 1.1b Predicted vs. actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: option value model 2

together with 1981 actual rates. Like the actual rates, the predicted rates under the window plan are much higher than the 1981 rates. Thus, in general, the model predicts an effect that is comparable in order of magnitude to the actual effect. The option value model, however, tends to overpredict departure rates for persons between 55 and 58 and to underpredict rates for those between 63 and 65. Because departures between 55 and 58 are overpredicted, the predicted cumulative departures are higher than the actual cumulative rates through age 62, as shown in the bottom panel of the figure. (The actual and predicted departure rates used in figs. 1.1a and 1.1b are shown in appendix tables 1B.1a and 1B.1b.)

For comparison, the same graphs are reproduced in figures 1.2a and 1.2b, but with the extreme value dynamic programming (specification 2) predictions added. The two models yield very similar results. Although the likelihood

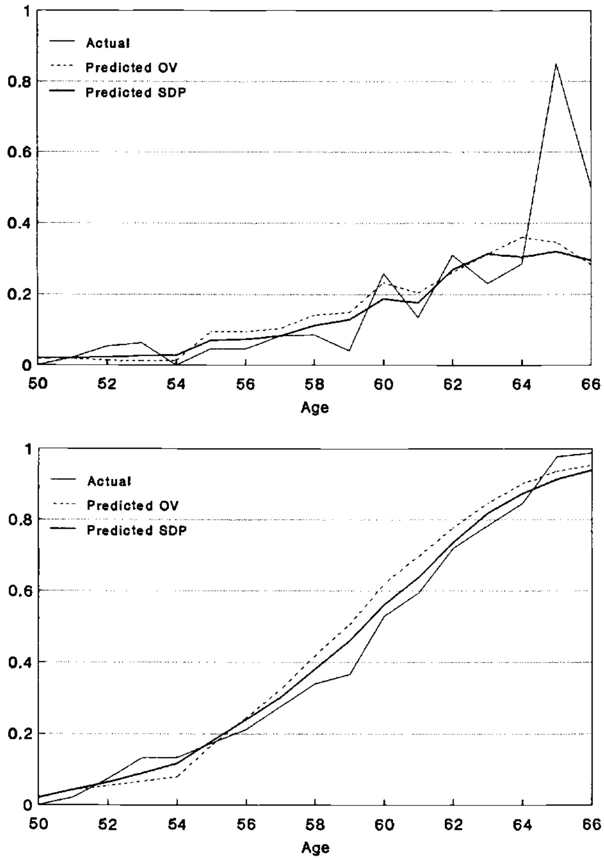
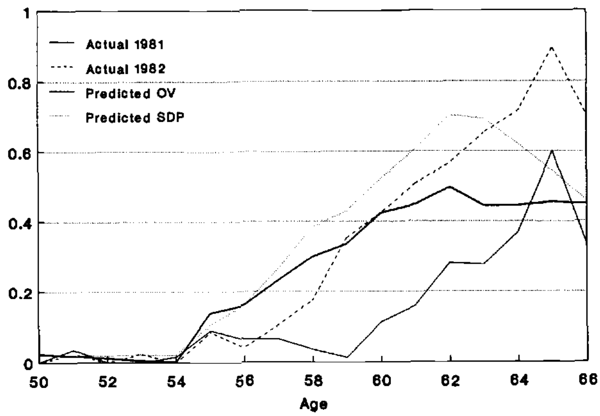


Fig. 1.2a Predicted vs. actual 1980 departure rates and implicit cumulative departures, by age: option value model 2 and stochastic dynamic programming model 2

values from the two models are about the same, the dynamic programming within-sample χ^2 measure of fit is better than the option value measure (as shown in table 1.2), and this is reflected in figure 1.2a. In particular, the dynamic programming model fits departure rates between 55 and 59 somewhat better than the option value model does. Thus, the implied cumulative rates from the dynamic programming model track the actual rates better than the option value model predictions do.

On the other hand, departure rates under the window plan (fig. 1.2b) are predicted better by the option value than by the dynamic programming model, although the differences are not large. The dynamic programming overprediction of departure rates between 55 and 59 is greater than the option value overprediction at these ages. In addition, the dynamic programming model



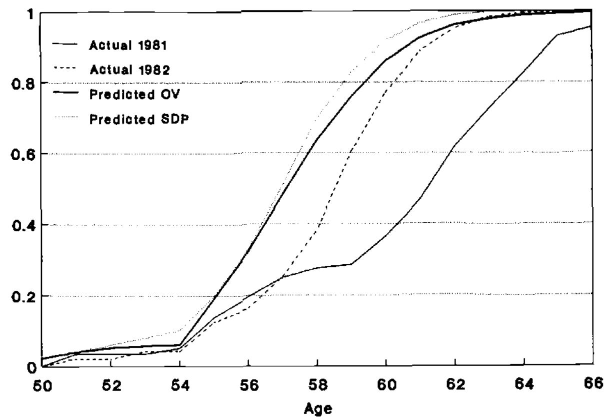


Fig. 1.2b Predicted vs. actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: option value model 2 and stochastic dynamic programming model 2

overpredicts departure rates through age 63 as well, while the option value model underpredicts departure rates beginning at age 61. (The actual and predicted departure rates used in figs. 1.2a and 1.2b are shown in appendix tables 1B.2a and 1B.2b.)

The extreme value and the normal versions of the dynamic programming model are compared in figures 1.3a and 1.3b. As the figures show, there is little difference between the predictions from the two specifications, although the normal version fits actual departure rates under the window plan somewhat better than the extreme value version. The normal model χ^2 sample statistic is slightly larger than the extreme value statistic, but the normal χ^2 window statistic is lower than the corresponding extreme value statistic, as shown in table 1.2.

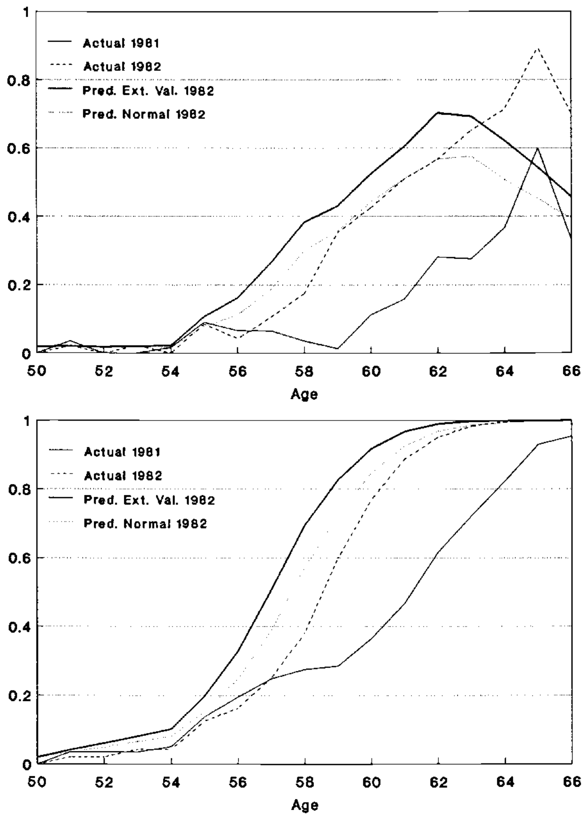


Fig. 1.3a Predicted vs. actual 1980 departure rates and implicit cumulative departures, dynamic programming model, by age: extreme value distribution (model 2) and normal distribution (model 5)

The three models are compared in figure 1.4. The figure shows the difference between the 1982 and the 1980 predicted departure rates based on the three models versus the difference between the actual 1982 and 1980 rates. As the previous figures suggest, the three models yield very similar results, although the option value model tends to underestimate the effects of the window plan whereas the dynamic programming models tend to overestimate the effects.

To summarize, in accordance with the actual effect of the window plan, both the option value and the dynamic programming models predict a large increase in departure rates under the window plan. This comparison does not suggest to us that one model is noticeably better or worse than the other.

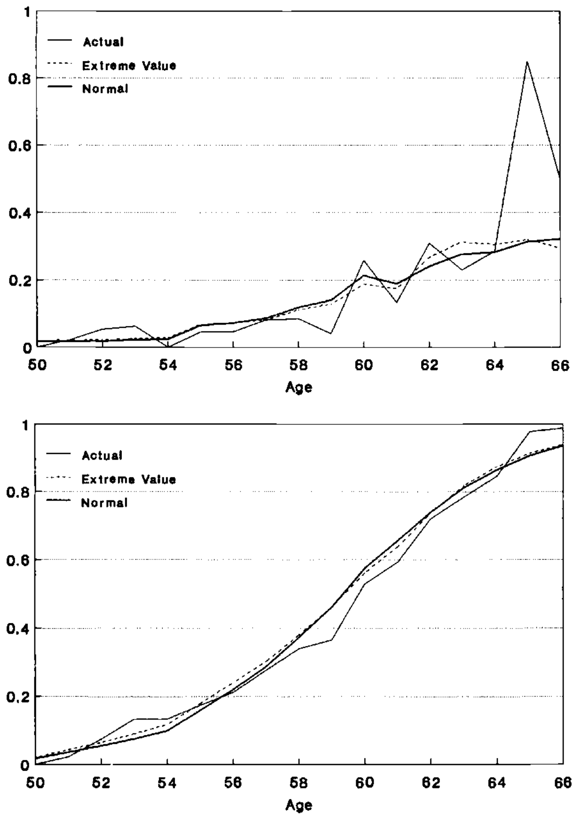


Fig. 1.3b Predicted vs. actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: dynamic programming model 2 (extreme value distribution) and model 5 (normal distribution)

Selected Probit Model Results

The graphs confirm that the probit models are typically inferior to the more behavioral forward-looking models. But probit specifications that include forward-looking variables capture some of the important features of the option value and the dynamic programming rules. The results of the probit model using the calculated option value variable (computed with $\gamma = 1$, $k = 1$, and $\beta = .95$) are graphed in figures 1.5a and 1.5b. This specification shows very little variation in retirement rates with age, as shown in the top panel of figure 1.5a, and the implied cumulative rates yield a poor approximation to the actual rates. The model predicts very little response to the window plan.

By using both the calculated option value variable and age, it is possible to

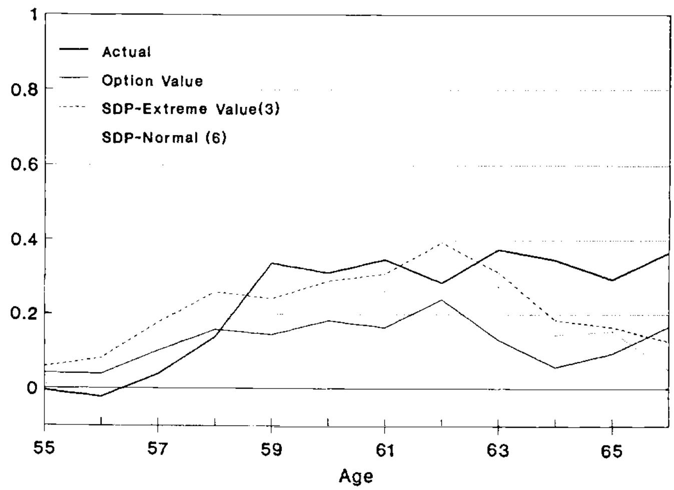


Fig. 1.4 Actual and predicted increases in retirement rates under the 1982 window plan: option value model, stochastic dynamic programming–extreme value model 3, and stochastic dynamic programming–normal model 6

fit the observed departure rates well, as shown in figure 1.6a. But this specification has essentially no behavioral implications: as revealed in figure 1.6b, there is almost no predicted response to the window plan.

The probit specification with the best fit (excluding the specification with age) is based on the current present value of SS and pension benefit entitlements (accumulated SS and pension wealth), the accrual in SS and pension wealth if the person works another year, and expected wage income if the person works another year (specification 8 in table 1.1). This is shown in figure 1.8a. This model fits the sample data about as well as the forward-looking models; indeed, it yields a lower within-sample χ^2 statistic than these more behavioral models. Essentially the same results are obtained when the SS and pension wealth variables are excluded (specification 6 in table 1.1), as shown in figure 1.7a.

But both these probit specifications greatly overpredict retirement rates under the window plan, as shown in figures 1.7b and 1.8b. The window χ^2 statistics also show that the forward-looking models predict the window plan departure rates much better than the probit models do. Aside from the details of functional form, the basic difference between the models is that the probit specification assumes that retirement decisions are based on a rule that involves looking ahead only one period whereas the option value and the dynamic programming rules consider all future potential retirement dates. In this instance at least, a rule that incorporates evaluation of events in the foreseeable future is more consistent with individual behavior than one that limits consideration to events in the next year only.

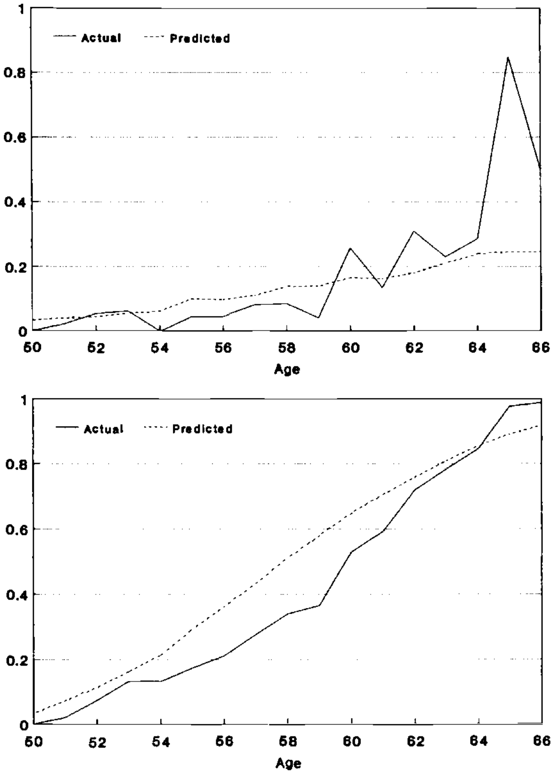


Fig. 1.5a Predicted vs. actual departure rates and implicit cumulative departures, by age: probit model 1

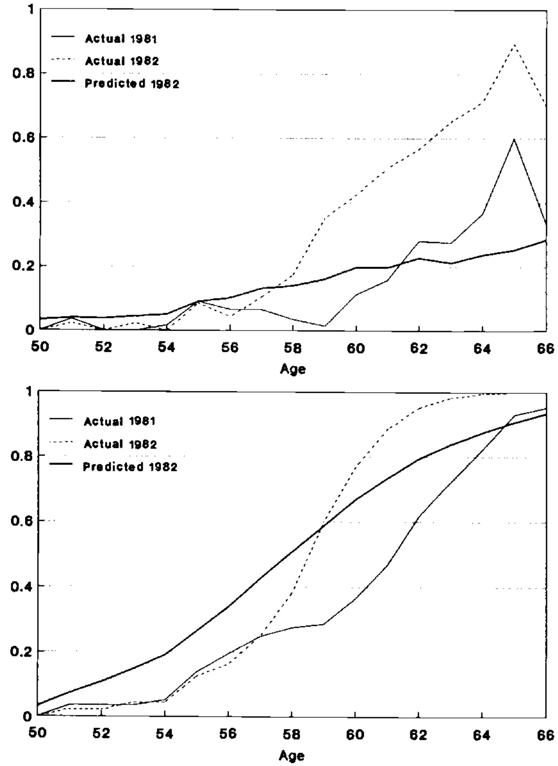


Fig. 1.5b Predicted vs. actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: probit model 1

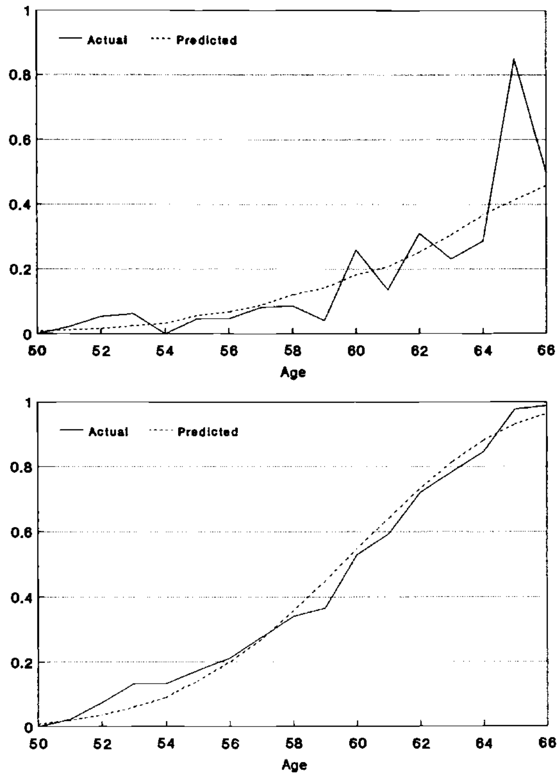


Fig. 1.6a Predicted vs. actual departure rates and implicit cumulative departures, by age: probit model 2

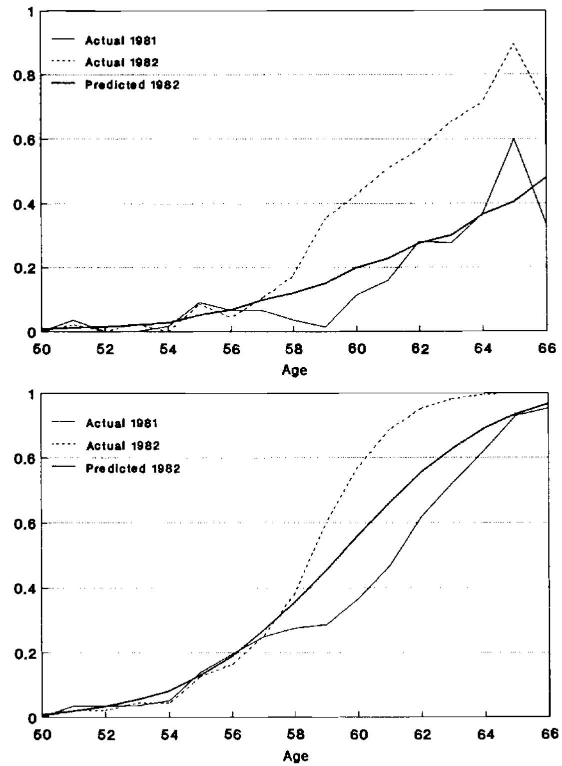


Fig. 1.6b Predicted vs. actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: probit model 2

Fig. 1.7a Predicted vs. actual departure rates and implicit cumulative departures, by age: probit model 6

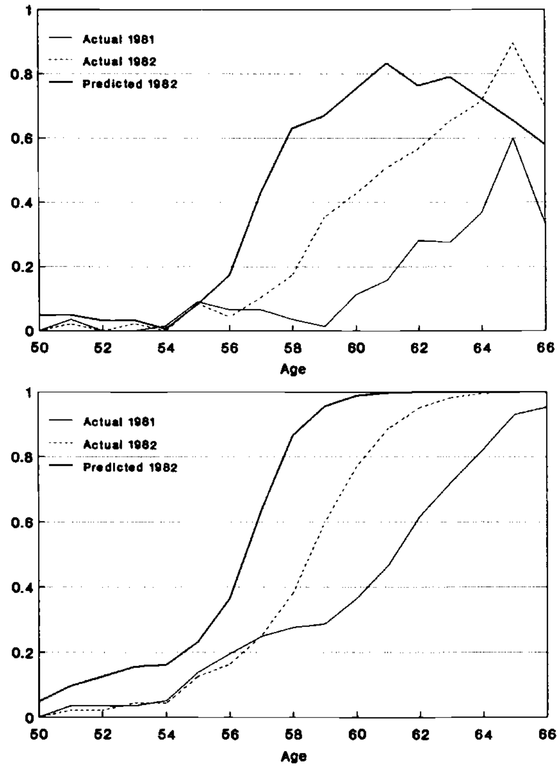


Fig. 1.7b Predicted vs. actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: probit model 6

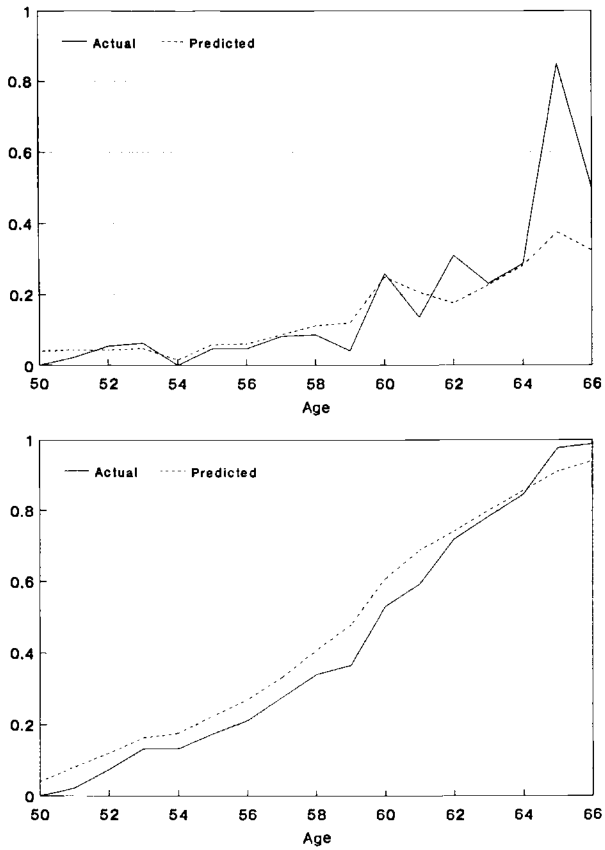


Fig. 1.8a Predicted vs. actual departure rates and implicit cumulative departures, by age: probit model 8

1.4 A Simulation: The Elimination of the Social Security Early Retirement Option

As a further comparison of the models, we have simulated the effect of removing the SS early retirement option so that SS benefits are only available beginning at age 65. A comparison of predicted retirement rates with and without the SS early retirement is shown in table 1.3 by model for ages 60–65.

According to the simulation based on the option value model, eliminating SS early retirement reduces predicted retirement rates among persons 62–64 by about 23 percent. The extreme value dynamic programming specification shows noticeably larger effects, but the effects based on the normal dynamic programming specification are smaller than the option value estimated effects.

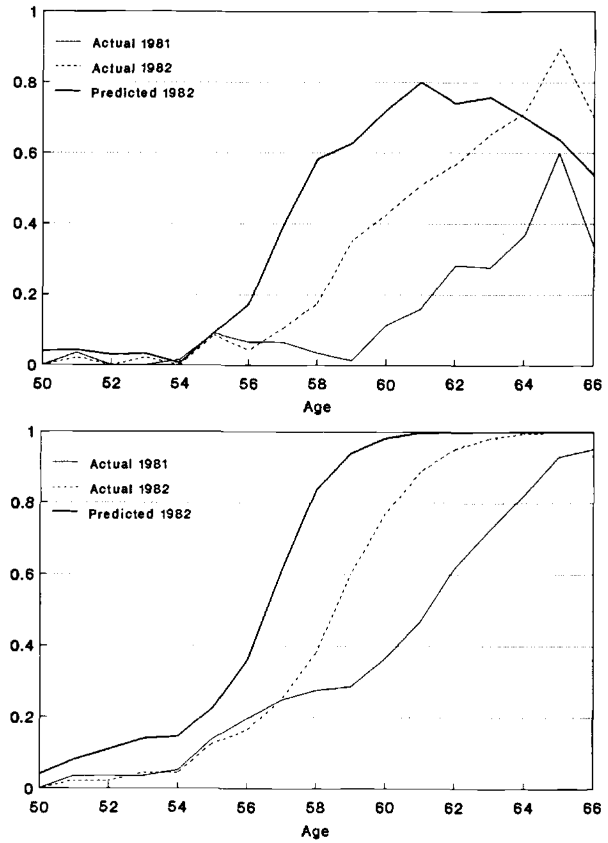


Fig. 1.8b Predicted vs. actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: probit model 8

Because a large proportion of employees in this firm have already left the firm before 62, the reduction applies to only the small proportion of employees who are still working, and thus the effect on the overall retirement is small. To the extent that these reductions generalize to workers not covered by defined benefit plans with incentives for early retirement, these estimates suggest that an increase in the SS early retirement age would have a very substantial effect on labor force participation. A large proportion of retired persons relies almost exclusively on SS benefits for retirement income. According to these estimates, substantially fewer of these employees would leave the labor force if they could not collect SS benefits.

Because of data limitations, it has been common to use parameter estimates from models that exclude firm pension plan data to simulate the effect of changes in SS provisions. To demonstrate the potential effect of the exclusion

Table 1.3 Retirement Rates in 1980 with and without Social Security Early Retirement

			Dynamic Programming					
	Option Value		Extreme Value		Normal		Probit	
Age	With	Without	With	Without	With	Without	With	Without
60	.233	.229	.188	.172	.214	.199	.249	.242
61	.204	.197	.176	.142	.190	.170	.206	.201
62	.262	.218	.269	.177	.241	.205	.175	.136
63	.313	.258	.314	.214	.277	.240	.227	.155
64	.360	.294	.305	.230	.284	.258	.281	.175
65	.346	.346	.320	.320	.314	.314	.375	.375

Note: The entries are the predicted retirement rates from maximum likelihood estimates of option value model 2, dynamic programming model 2, dynamic programming model 5, and probit specification 8. See the notes to tables 1.1 and 1.2. "With" refers to the base (current) specification. "Without" estimates are from a simulation that eliminates the possibility of SS receipt as early as age 62. Under the simulation, SS benefit receipt begins at age 65. Details are provided in the text.

of firm plans, we have estimated the dynamic programming normal model (specification 5) using only SS benefits—instead of SS and the firm pension benefits—and these estimates have been used to simulate the effect of the elimination of SS early retirement. The results are shown in table 1.4, where they are compared to the dynamic programming normal estimates. The estimated effect of the elimination of SS early retirement is much greater when the firm pension is not accounted for. For example, the retirement rate at 62 is reduced from .291 to .081; the base model yields a reduction from .241 to .205.

1.5 Summary

We have compared the in-sample and out-of-sample predictive performance of three models of retirement. The goal was to determine which of the retirement rules most closely matched observed retirement behavior in a large firm. The primary measure of predictive validity was the correspondence between the model predictions of retirement behavior and actual retirement under the firm window plan. Model parameter estimates were obtained on the basis of retirement in 1980. These estimates were then used to predict retirement in 1982 when the window plan was in effect. Retirement rates of persons eligible for the window plan bonus typically doubled in 1982 compared to earlier (and later) years.

The option value and the dynamic programming models fit the sample data equally well, with a slight advantage to the normal dynamic programming model. Both models correctly predicted a very large increase in retirement under the window plan, with some advantage in fit to the option value model.

Table 1.4 **Retirement Rates in 1980 with and without Social Security Early Retirement, Comparison with Estimates Based on Social Security Only, Using Dynamic Programming Normal Specification**

	Dynamic Programming: Normal			
	Base (SS & pension data)		SS Data Only	
Age	With	Without	With	Without
60	.214	.199	.114	.057
61	.190	.170	.167	.067
62	.241	.205	.291	.081
63	.277	.240	.310	.118
64	.284	.258	.334	.191
65	.314	.314	.356	.356

In short, this evidence suggests that the option value and dynamic programming models are considerably more successful than the less complex probit model in approximating the rules individuals use to make retirement decisions but that the more complex dynamic programming rule approximates behavior no better than the simpler option value rule. More definitive conclusions will have to await accumulated evidence based on additional comparisons using different data sets and with respect to different pension plan provisions.

Appendix A
The Firm Retirement Plan

To understand the effect of the pension plan provisions, figure 1A.1 shows the expected future compensation of a person from our sample who is 50 years old and has been employed by the firm for twenty years. For convenience, figure 1A.1 assumes a 5 percent real discount rate and zero inflation. In the estimated model reported in section 1.3, the discount rate is estimated, and the inflation rate is assumed to be 5 percent. Total compensation from the firm can be viewed as the sum of wage earnings, the accrual of pension benefits, and the accrual of Social Security benefits. (This omits medical and other unobserved benefits that should be included as compensation but on which we do not have data.) As compensation for working another year, the employee receives salary earnings. He also receives compensation in the form of future pension benefits. The annual compensation in this form is the change in the present value of the future pension benefits entitlement due to working an additional year. This accrual is comparable to wage earnings. The accrual of

Social Security benefits may also be calculated in a similar manner and is also comparable to wage earnings. Figure 1A.1 shows the present value at age 50 of expected future compensation in all three forms. The line labeled "wage earnings" represents cumulated earnings by age of retirement (more precisely, by age of departure from the firm, since some workers might well continue to work in another job). For example, if the person were to retire at age 62, his cumulated earnings between age 50 and age 62, discounted to age-50 dollars, would be about $144,000. The slope of the earnings line represents annual earnings discounted to age-50 dollars.

The solid line shows the accrual of firm pension plus Social Security benefits, again discounted to age-50 dollars. The shape of this profile is determined primarily by the pension plan provisions. The plan's normal retirement age is 65, and the early retirement age is 55. Cliff vesting occurs at ten years of service. Normal retirement benefits at age 65 are determined by age times years of service, multiplied by some constant factor. The most important additional provisions—those that determine the shape of the profile in figure 1A.1—are described here; full details of the plan provisions are presented in Kotlikoff and Wise (1987). The present value of retirement benefits increases between 50 and 54 because years of service, and possibly earnings, increase. An employee could leave the firm at age 53, for example. If he were to do that, and if he were vested in the firm's pension plan, he would be entitled to normal retirement pension benefits at age 65, based on his years of service and *current* dollar earnings at age 53. He could start to receive benefits as early as age 55, the pension early retirement age, but the benefit amount would be reduced actuarially. Thus, in present value terms, the stream of benefits received beginning at 55 would be equal to the stream of benefits beginning at 65; the annual benefit amount would be reduced just enough to offset the receipt of benefits for ten more years. If he started to receive benefits at age 55, they would be only 36 percent of the dollar amount he would receive at age 65. If, however, he were to remain in the firm until the early retirement age, the situation would be quite different. He would be entitled to normal retirement benefits based on his years of service and salary at age 55. But if he were to start to receive them at age 55, the benefits would be reduced less than actuarially, about 3 percent for each year that retirement precedes age 65, instead of 6 or 7 percent.

In addition, the plan has a Social Security offset provision. Pension benefits are offset by a specified amount, depending on the firm estimate of Social Security benefits. But if the person takes early retirement, between 55 and 65, the Social Security offset is not applied to benefits received before age 65. These two provisions create the large discontinuous jump in retirement benefits at age 55—from about $33,000 to $56,000. This increase is equivalent to more than 130 percent of his annual wage earnings at 55. Thus, there is an enormous bonus for remaining with the firm until that age. After age 55, however, the person who does not retire forgoes the opportunity of taking pension

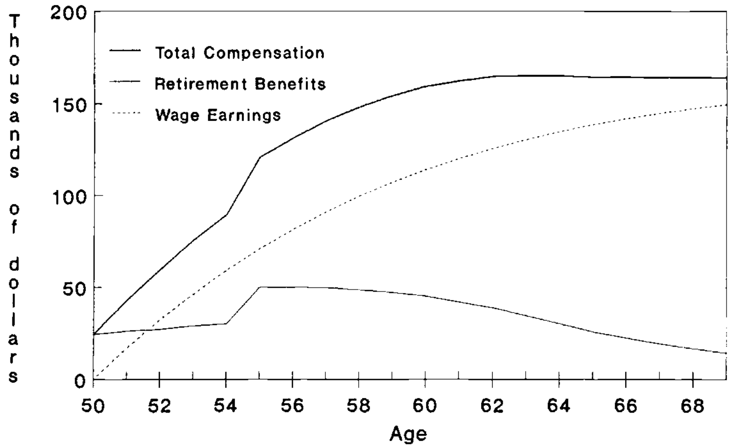


Fig. 1A.1 Present discounted values of future earnings and retirement benefits, as a function of date of retirement

benefits on very advantageous terms—thus the minimal change in the discounted value of benefits between 55 and 60.

If a person has thirty years of service at age 60, he is entitled to full normal retirement benefits. No early retirement reduction is applied to benefits if they are taken then. That is, by continuing to work, he will no longer gain from fewer years of early retirement reduction, as he did before age 60—thus the kink in the profile and the decline thereafter.

The top line shows total compensation. For example, if the employee were to leave the firm at age 60, his wage earnings between 50 and 60 would be $126,000, shown by the wage earnings line. Thereafter, he would receive firm pension plan and Social Security retirement benefits with a present value—at age 50—of about $58,000. The sum of the two is about $184,000, shown by the top line. The large jump at 55 reflects the early retirement provisions of the pension plan. Total compensation declines modestly each year through age 60 and very rapidly thereafter. After age 62 or 63, annual total compensation is close to zero.

Appendix B
Tabulations of Predicted and Actual Retirement Rates

This appendix presents tabulations of the values presented graphically in figures 1.1–1.2. These figures are the predicted and actual retirement rates, or hazard rates, for the employees in the data set and the associated cumulative retirement rates.

The actual retirement rates for each age group are the fraction of workers of that age who retire during the indicated year. The predicted retirement rates are the aggregate rates predicted by the indicated model; that is, the predicted retirement rate is the average predicted probability of retiring for all workers of the indicated age.

The cumulative retirement rates are computed from the single-year retirement rates by following a cohort of one hundred 50-year-olds at the firm for the next twenty years, assuming that the annual retirement rates for this cohort are the same as the annual retirement rates for the indicated year, predicted or actual, as the case may be. For example, in 1980, the actual retirement rates (in our sample of 993 workers) of 50-, 51-, and 52-year-olds were, respectively, .00, .022, and .054. Thus, the cumulative retirement rate for 52-year-olds is $1 - (1 - .00)(1 - .022)(1 - .054) = .075$.

The numbering of the tables in this appendix corresponds to the numbering of the figures in the text: the values plotted in figure 1.1a appear in table 1B.1a, etc.

Table 1B.1a **Data for Figure 1.1a**

		Cumulative Retirement Rates		Annual Retirement Rates	
Age	No. of Observations	Actual	Predicted	Actual	Predicted
50	83	.000	.023	.000	.023
51	89	.022	.042	.022	.019
52	74	.075	.055	.054	.014
53	64	.133	.067	.063	.012
54	77	.133	.079	.000	.013
55	64	.174	.167	.047	.095
56	64	.212	.246	.047	.095
57	61	.277	.324	.082	.104
58	81	.340	.419	.086	.141
59	74	.366	.506	.041	.149
60	85	.530	.621	.259	.233
61	37	.594	.698	.135	.204
62	42	.720	.777	.310	.262
63	39	.784	.847	.231	.313
64	35	.846	.902	.286	.360
65	20	.977	.936	.850	.346
66	4	.988	.954	.500	.283

Note: The actual retirement rates were computed for the 1,000 persons in the sample. The predicted retirement rates are based on option value model 2.

Table 1B.1b **Data for Figure 1.1b**

	Cumulative Retirement Rates			Annual Retirement Rates		
Age	Actual 1981	Actual 1982	Predicted 1982	Actual 1981	Actual 1982	Predicted 1982
50	.000	.000	.023	.000	.000	.023
51	.036	.022	.042	.036	.022	.019
52	.036	.022	.053	.000	.000	.012
53	.036	.044	.059	.000	.023	.006
54	.052	.044	.062	.017	.000	.003
55	.139	.126	.192	.091	.085	.139
56	.195	.163	.323	.066	.043	.162
57	.249	.251	.480	.066	.105	.232
58	.276	.382	.635	.036	.175	.299
59	.286	.600	.758	.014	.352	.335
60	.366	.770	.860	.113	.425	.424
61	.467	.887	.923	.159	.508	.448
62	.617	.951	.961	.281	.566	.498
63	.723	.983	.978	.276	.652	.444
64	.824	.995	.988	.367	.714	.445
65	.930	.999	.993	.600	.895	.454
66	.953	1.000	.996	.333	.700	.449

Note: Based on 1980 option value model 2 parameter estimates, reported in table 1.2. The simulation is described in the text.

Table 1B.2a **Data for Figure 1.2a**

		Cumulative Retirement Rates			Annual Retirement Rates		
Age	No. of Observations	Actual	Option Value	Dynamic Programming	Actual	Option Value	Dynamic Programming
50	83	.000	.023	.021	.000	.023	.021
51	89	.022	.042	.043	.022	.019	.022
52	74	.075	.055	.065	.054	.014	.023
53	64	.133	.067	.090	.063	.012	.027
54	77	.133	.079	.117	.000	.013	.029
55	64	.174	.167	.179	.047	.095	.070
56	64	.212	.246	.240	.047	.095	.074
57	61	.277	.324	.303	.082	.104	.082
58	81	.340	.419	.381	.086	.141	.112
59	74	.366	.506	.461	.041	.149	.129
60	85	.530	.621	.562	.259	.233	.188
61	37	.594	.698	.639	.135	.204	.176
62	42	.720	.777	.736	.310	.262	.269
63	39	.784	.847	.819	.231	.313	.314
64	35	.846	.902	.874	.286	.360	.305
65	20	.977	.936	.914	.850	.346	.320
66	4	.988	.954	.940	.500	.283	.295

Table 1B.2b **Data for Figure 1.2b**

	Cumulative Retirement Rates				Annual Retirement Rates			
Age	Actual 1981	Actual 1982	Predicted Option Value 1982	Predicted Dynamic Programming 1982	Actual 1981	Actual 1982	Predicted Option Value 1982	Predicted Dynamic Programming 1982
50	.000	.000	.023	.021	.000	.000	.023	.021
51	.036	.022	.042	.043	.036	.022	.019	.022
52	.036	.022	.053	.062	.000	.000	.012	.020
53	.036	.044	.059	.082	.000	.023	.006	.022
54	.052	.044	.062	.103	.017	.000	.003	.023
55	.139	.126	.192	.199	.091	.085	.139	.107
56	.195	.163	.323	.329	.066	.043	.162	.162
57	.249	.251	.480	.506	.066	.105	.232	.264
58	.276	.382	.635	.696	.036	.175	.299	.384
59	.286	.600	.758	.827	.014	.352	.335	.430
60	.366	.770	.860	.917	.113	.425	.424	.524
61	.467	.887	.923	.967	.159	.508	.448	.604
62	.617	.951	.961	.990	.281	.566	.498	.703
63	.723	.983	.978	.997	.276	.652	.444	.693
64	.824	.995	.988	.999	.367	.714	.445	.622
65	.930	.999	.993	.999	.600	.895	.454	.543
66	.953	1.000	.996	1.000	.333	.700	.449	.457

References

Berkovec, James, and Steven Stern. 1991. Job Exit Behavior of Older Men. *Econometrica* 59 (1):189–210.

Burtless, Gary. 1986. Social Security, Unanticipated Benefit Increases, and the Timing of Retirement. *Review of Economic Studies* 53 (October):781–805.

Daula, Thomas V., and Robert A. Moffitt. 1991. Estimating a Dynamic Programming Model of Army Reenlistment Behavior. In *Military Compensation and Personnel Retention: Models and Evidence,* ed. C. L. Gilroy, D. K. Horne, and D. Alton Smith. Alexandria, Va.: U.S. Army Research Institute for the Behavioral and Social Sciences.

Gustman, Alan, and Thomas Steinmeier. 1986. A Structural Retirement Model. *Econometrica* 54:555–84.

Hausman, Jerry A., and David A. Wise. 1985. Social Security, Health Status, and Retirement. In *Pensions, Labor, and Individual Choice,* ed. David A. Wise, 159–91. Chicago: University of Chicago Press.

Hurd, Michael, and Michael Boskin. 1981. The Effect of Social Security on Retirement in the Early 1970s. *Quarterly Journal of Economics* 46 (November):767–90.

Kotlikoff, Laurence J., and David A. Wise. 1987. The Incentive Effects of Private Pension Plans. In *Issues in Pension Economics,* ed. Z. Bodie, J. Shoven, and D. Wise, 283–339. Chicago: University of Chicago Press.

Lumsdaine, Robin L., James H. Stock, and David A. Wise. 1990. Efficient Windows and Labor Force Reduction. *Journal of Public Economics* 43:131–59.

———. 1991. Fenêtres et retraites (Windows and retirement). *Annales d'Économie et de Statistique* 20/21:220–42.
Rust, John P. 1989. A Dynamic Programming Model of Retirement Behavior. In *The Economics of Aging,* ed. David A. Wise, 359–98. Chicago: University of Chicago Press.
Stock, James H., and David A. Wise. 1990a. The Pension Inducement to Retire: An Option Value Analysis. In *Issues in the Economics of Aging,* ed. David A. Wise, 205–24. Chicago: University of Chicago Press.
———. 1990b. Pensions, the Option Value of Work, and Retirement. *Econometrica* 58 (5):1151–80.

Comment Sylvester J. Schieber

In their paper, Robin L. Lumsdaine, James H. Stock, and David A. Wise evaluate the complexity of models versus their validity in predicting the retirement rates of older male workers in a large firm. The firm, coincidentally, offered an early retirement window that allowed them to test the alternative models being considered by comparing predicted retirement rates at various ages with actual rates.

The paper is extremely interesting and is headed down an important track. Understanding the responses to retirement incentives that employers offer their employees is of interest to both the public policy and the employer communities. It is important to policymakers in the development of macro policy as shown by the authors' analysis of the effect of the early retirement option under Social Security. It is important to employers in the development of micro policies that are aimed at controlling their work forces through the structuring of incentives encouraging the continued work or retirement of older workers. Employers are particularly interested in anticipating the responses to special incentive programs that they introduce to encourage some of their workers to leave their current jobs. The introduction of an early retirement window by the firm on which the authors had data offered an ideal opportunity to test these models' relative predictive capabilities. While I found the paper interesting and to be moving in the right direction, I will voice a number of criticisms toward the end of my comments.

The authors test three different models in the course of their analysis. Their basic model, or at least the first one evaluated in the paper, is an "option value" model, described in Stock and Wise's earlier work.[1] In the option value model, an employee compares the value of retiring today with the maximum

Sylvester J. Schieber is a vice president of the Wyatt Company and director of its Consulting Support Services in Washington, D.C.

1. James H. Stock and David A. Wise, "Pensions, the Option Value of Work, and Retirement," *Econometrica* 58, no. 5 (September 1990): 1151–80.

of the value of retiring at discrete times in the future. If the difference is greater than zero, then the individual continues to work.

I find the option value model appealing because it corresponds with what I believe employers implicitly assume in designing their retirement programs. Employers today are in a position that they cannot systematically terminate older workers because of the Age Discrimination in Employment Act (ADEA). Yet they have concluded that the relative productivity of workers begins to decline at some age between 55 and 65.[2] Given the ADEA restrictions, the only way employers can get rid of workers with declining productivity is to bid them out of the firm. Using the option value model, a person will continue to work if the option value of postponing retirement is greater than zero. The employer typically structures the retirement plan in such a way that the option value of continuing to work is less than zero for most workers attaining some age. They do this by providing early retirement incentives in their pension plans. They can selectively override these early retirement incentives (i.e., manipulate the option value of continuing to work or retiring) by increasing current pay for those older workers they want to keep.

The parameters in the option value model are developed by estimation. The second model the authors consider, their probit model, is a variant of the option value model, except that the gains from added work life are based on assumed values and discount rates. The third model, a stochastic dynamic programming model, considers the expected value of maximum current options versus future options. Under this model, at the beginning of each year the worker decides to take up the utility of the retirement benefits or to work another year, deriving the utility of the related earnings and retaining the option to make another choice next year.

In terms of computational complexity, the models would be ranked from the easiest (the probit model) to the most complex (the stochastic dynamic programming model). The probit model does not generate nearly as good predictions of retirement under the window plan as the other two models, which are roughly equivalent.

One criticism of the paper is that the authors are looking at quit rates rather than retirement. There is likely to be a considerable amount of second career activity going on for individuals who are eligible to partake of many early retirement incentive programs, especially early retirement window plans. Certainly, workers looking at other job opportunities would be facing significantly different option values for quitting a firm under the circumstances described here than the option values of retirement they would face.

Another aspect of the value of continued work as it relates to quit or retirement decisions overlooked in the analysis is the availability of pre-65 retiree

2. There is no explicit information available documenting employers' conclusion that productivity begins to decline at a specific age. But the incentive effects in pension plans encouraging retirement at specific ages are strong implicit evidence that employers have concluded that they want certain workers to leave the firm.

health benefits and the health care needs of workers and their dependents. In some cases, the value of early retirement health benefits may actually exceed the value of lifetime pension benefits. In a similar vein, the availability of a defined contribution plan that would supplement the basic defined benefit plan should be considered in modeling retirement behavior.

The application of the three models to only one firm might render the consideration of early retirement benefits and defined contribution plan accumulations irrelevant if the firm did not have either type of plan. But that in itself is a problem. The models should be tested further over a range of firms with varying incentives in their basic pension benefits, but also with varying availability and generosity of these other supplemental plans.

In evaluating the relevance of being able to predict retirement under situations where explicit incentives are being offered, it is important to be able to predict how many workers will retire. The option value and stochastic dynamic programming models do a relatively good job in that regard. It is equally important, however, for employers to be able to predict which workers will quit under the plan. So the models should be evaluated on their ability to predict the quitters quitting and the stayers staying.

When employers offer early retirement windows, they never expect all eligible workers to take advantage of the incentives. They are typically looking to get rid of the less productive workers among the whole group to whom they offer the incentive. The window is considered to be only partly successful if the wrong people retire. The models were not tested in this regard. One problem with applying such a test would be in identifying which people the company wanted to keep and which they did not. One possible way of discriminating between the two groups is to look at wage increases over the prior two or three years. Presumably, a substantial wage increase above the norm for other workers in the company would be the company's way of indicating the employee's relative productivity. A worker who received a relatively low wage increase would presumably be perceived as relatively marginal.

The conventional wisdom on window plans is that they often, if not almost always, encourage the wrong people to quit the firms offering them. This suggests, at the applied level, that the options value model is not working, or at least that it does not appear to be working within the context that Lumsdaine, Stock, and Wise are testing it. If the firm is acting rationally and giving the workers they desire to keep good raises, it would suggest that the workers' option value for continuing to work should be positive. The converse is true for the workers the firm desires to have quit. Yet the keepers seem to leave, and the dregs seem to stay under window plans. If this is the case, it would be an interesting phenomenon to evaluate with the models being tested here. To do so would require added information on the quitters' subsequent work behavior. I suspect that there is a rational explanation for what generally seems to happen under these window plans that could be tested.

Through the process of annual reviews, raises, bonuses, etc., most employers tell their good workers that they are good. When they communicate a window plan, they tell their workers that things are not going well at the firm. The workers who are good and know that they are good, because they have been told, find the option value of quitting the firm greater than zero but the option value of retiring less than zero. They conclude, "I'm good, but my situation is bad, and my employer is offering me a bonus to tide me over during a transition." They take the money and run.

The poor performers, on the other hand, have been getting the message that they are not very productive people. The window communicates to them that the economic situation in the marketplace is not good. They perceive that their option value of continuing to work is greater than zero and that the value of continuing in their current job exceeds that of the alternatives they face. They conclude, "I may not get any more pay raises if I stay here, but at least I will get paid. Also, I do not have a good record to take to a bad market." Therefore, they stay put.

Finally, in closing, I would like to propose an alternative model to the ones tested by the authors. It is a simplified options value model, which I would characterize as a goals attainment model. Most workers do not calculate the present value of alternative income streams from working or retiring to decide when the combined value of a retirement income stream plus their added leisure more than offsets the value of their income stream if they continue to work.

Most workers want to maintain some preconceived standard of living in retirement. Employers gear much of the employee communications material to explaining the level of benefits their retirement plans provide and the target levels workers have to attain to reach their retirement income goals. Employees use this material to help define their goals, but they calculate their options values by looking around. They know people they have worked with over the years who have retired under their employer's plans or similar plans provided by other employers. They can tell which of those former workers are able to maintain a standard of living to which they themselves aspire and which are not. Through a process of elimination, they determine which of those former employees have the characteristics that they have to roughly match in order for their retirement goals to be met. Once they match the characteristics that they believe correlate with their retirement income goals, they retire.