Volume Title: Studies in the Economics of Aging

Volume Author/Editor: David A. Wise, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-90294-3

Volume URL: http://www.nber.org/books/wise94-1

Conference Date: May 1992

Publication Date: January 1994

Chapter Title: Pension Plan Provisions and Retirement: Men and Women, Medicare, and Models

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

Chapter URL: http://www.nber.org/chapters/c7344

Chapter pages in book: (p. 183 - 222)

6 Pension Plan Provisions and Retirement: Men and Women, Medicare, and Models

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

Our ongoing analysis of the effects of pension plan provisions on retirement is pursued in this paper. The work to date has emphasized the dramatic effect of employer-provided pension plan provisions on age of retirement and the enormous effects of changing the provisions. The work has also highlighted the important limitations of using Social Security provisions to predict retirement behavior, without accounting for the effect of employer pension plan provisions, which, for employees who have such plans, is typically much more powerful than the effect of Social Security provisions.

Two aspects of our work have guided the analysis as the research progressed. The first is that a new method has been used to model retirement decisions. The second is that the empirical analysis has been based on data from individual firms. Thus we have been led to consider whether the model provided accurate predictions of the effects of plan provisions on retirement and whether the behavioral implications of analysis based on data from one firm could be generalized to other firms with different plan provisions.

In two initial papers, Stock and Wise (1990a, 1990b) developed an "option value" model of retirement. The central feature of this model is that in deciding whether to retire employees are assumed to compare the "value" of retiring now to the *maximum of the expected values* of retiring at all future retirement ages. If the maximum of the future values is greater than the value of retirement

Robin L. Lumsdaine is assistant professor of economics at Princeton University and a faculty research fellow of the National Bureau of Economic Research. James H. Stock is professor of political economy at the John F. Kennedy School of Government, Harvard University, and a research associate of the National Bureau of Economic Research. David A. Wise is the 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 from the National Institute on Aging and the Hoover Institution (Wise) is gratefully acknowledged. The authors also thank Jim Smith and John Rust for helpful comments.

now, the employee continues to work. We tested the predictive validity of this model in two ways: first, we considered the "within-sample fit" of the model, by comparing the actual pattern of retirement by age to the pattern predicted by the model, based on the data used for estimation. Second, in papers by Lumsdaine, Stock, and Wise (1990, 1991) we emphasized an external "out-of-sample" check of predictive validity, by considering how well the model predicted the effect on retirement of an unanticipated and temporary change in the pension plan provisions, occasioned by an early retirement window plan. In a subsequent paper, Lumsdaine, Stock, and Wise (1992) compared the predictive validity of the option value model with that of two versions of stochastic dynamic programming models. The stochastic dynamic programming model is close in spirit to the option value model, but the prediction of retirement is based on the comparison of the value of retirement now to the *expected value of the maximum* of the values of future retirement ages. The evidence was that the option value model predicted just as well as the stochastic dynamic programming models, but had the advantage of being much less complex numerically. Ausink (1991) pursued a similar comparison based on retirement from the military and found that the option value version was noticeably better than the stochastic dynamic programming versions.

All of these papers, with the exception of the work by Ausink, are based on data from a single firm. The use of firm data was motivated by the absence of information on pension plan provisions in standard data sources, such as the Retirement History Survey, and by the realization that the incentives inherent in such plans could be very substantial and varied widely among firms, as shown in papers by Bulow (1981), Lazear (1983), and Kotlikoff and Wise (1985, 1987, 1988). In principle, the ideal data source would provide retirement information and pension plan information for a random sample of employees, from a wide range of firms. Such information has not been available. The alternative we followed was to obtain data from several different firms. The hope was that similar results from different firms would tend to confirm the validity of the model, even though the firms themselves could not be considered a random sample of all firms. Thus there is a need to determine whether the results for the single firm are confirmed based on data from other firms.

Therefore, the first goal of this paper is to use data from a new firm to confirm that the age pattern of retirement corresponds to the pension plan provisions. Descriptive analysis confirms that this is the case.

A second emphasis in this paper is the comparison of the retirement behavior of men and women. It is sometimes proposed that women may tend to retire earlier than men because they are typically younger than their husbands and they may tend to retire when their husbands do. From descriptive analysis, it is clear that the retirement patterns of men and women are not appreciably different.

A third goal is to add another observation to the list of comparisons of the predictive validity of the option value versus that of the stochastic dynamic programming model. Predictive validity in this paper is judged by the model predictions of retirement under a special 1983 early retirement incentive plan. The models are estimated based on retirement decisions in 1982. The goal of the comparisons is to accumulate data on the extent to which the different models predict actual retirement choices and thus to determine which specification best approximates the considerations that determine actual retirement decisions. The emphasis is not on which model best approximates the economists' view of the "right" calculation, but rather which best approximates the calculations that the typical person makes. Or, better still, which predicts best the retirement decisions of the typical employee.

A fourth goal is to make limited inferences about the potential effect of Medicare availability on retirement, especially at age 65. We have found in our previous work that model predictions of the age-65 retirement rate are typically much lower than the actual rate. We have attributed the high actual rate to an "age-65 retirement effect." But our work, and we believe the work of others, has ignored the potential effect of Medicare insurance that becomes available at age 65. The approach used here is to consider how retirement rates—especially at age 65—would be affected were Medicare valued according to the average payments to the covered population. The final goal is to compare parameter estimates based on data from two firms with different pension plans.[1]

6.1 Background

6.1.1 The Firm III Plan

Employees are covered by a defined-benefit pension plan with normal retirement at age 65 and early retirement at age 55. Cliff vesting occurs at 10 years of service (YOS) or at age 65, whichever comes first. The normal retirement benefit at 65 depends on earnings, age, and years of service at retirement (that is, at the time of departure from the firm). A person can retire and elect to start receiving benefits before age 65, but the normal benefit will be reduced by 5 percent for each year that receipt of benefits precedes age 65, as shown in figure 6.1. A person who retired at age 55, for example, would receive 50 percent of the normal retirement benefit of a person who left the firm at age 65. (The normal benefit also depends on years of service at the time of retirement.)

However, if a person has 30 years of service at retirement and if the person is age 60 or older, the person is eligible for 100 percent of the normal benefit. Benefits are reduced 5 percent for each year that retirement precedes age 60,

1. In order to maintain numerical consistency throughout our work, we refer to the firm used in this analysis as "firm III." The comparison firm, used in Stock and Wise (1990a, 1990b) and Lumsdaine, et al. (1990, 1991, 1992), is "firm I."

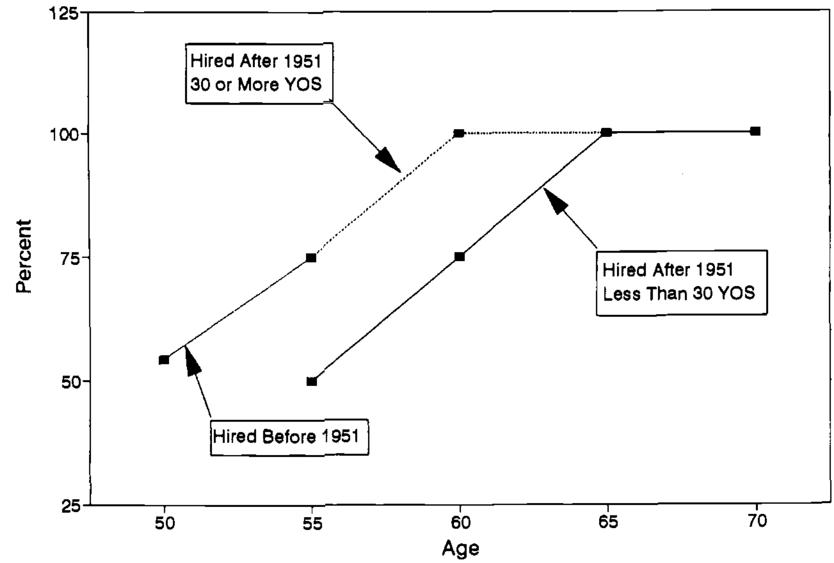


Fig. 6.1 Early retirement benefit (% of normal age-65 benefit)

if the person has 30 years of service. For example, a person who retired at age 55 with 30 years of service would receive 75 percent of the normal benefit.

Even a person who retires before age 55 and is vested can elect to receive benefits from the pension plan as early as age 55, but like the post-55 retiree, benefits are reduced 5 percent for each year that receipt of the benefits precedes age 65. Of course, this person's benefits would be based on earnings, age, and years of service at the time of retirement, unadjusted for earnings inflation, and would thus be lower than the benefits of a person who retired later.

Employees who joined the firm before 1951 can retire as early as age 50 and begin to receive benefits immediately, but at a reduced rate. An employee hired before 1951 had at least 31 years of service in 1982. The reduction for this group is indicated by the extended line that indicates benefits at age 50 to be 54.3 percent of the age-60 benefits for an employee who has 30 or more years of service at retirement.

To understand the effect of the pension plan provisions, figure 6.2*A* shows the expected future compensation of a person from our sample who is 51 years old and has been employed by the firm for 23 years. To compute the data graphed in figures 6.2*A*–6.2*D*, a 5 percent real discount rate and a 6 percent inflation rate are assumed. The discount rate is estimated in the empirical analysis, and the inflation rate is assumed to be 6 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 for which we have no data.) As compensation for working another year, the em-

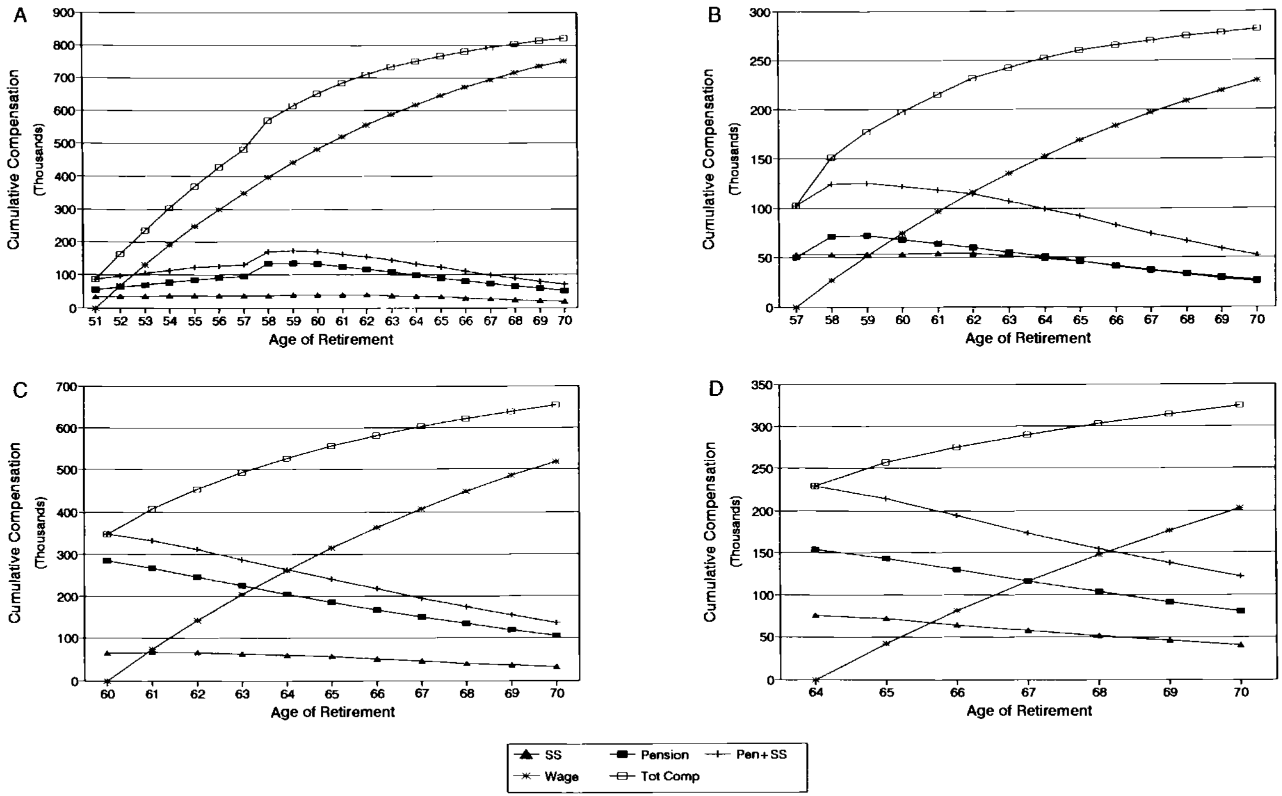


Fig. 6.2 Future compensation: persons (*A*) age 51 with 23 YOS, (*B*) age 57 with 29 YOS, (*C*) age 60 with 38 YOS, and (*D*) age 64 with 45 YOS

ployee receives salary earnings. Compensation is also received 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 be calculated in a similar manner, and is also comparable to wage earnings. Figure 6.2*A* shows the present value at age 51 of expected future compensation in all three forms. Wage earnings represents cumulated earnings, by age of retirement from the firm (more precisely, by age of departure from the firm, since some workers might continue to work in another job). For example, the cumulated wage earnings of this employee between age 51 and age 60, were he to retire at age 60, would be about $482,000, discounted to age-51 dollars. The slope of the wage earnings line represents annual earnings discounted to age-51 dollars.

The pension line shows the accrual of firm pension benefits, again discounted to age-51 dollars. The shape of this profile is determined by the pension plan provisions. The present value of accrued pension benefit entitlements at age 51 is about $54,000. The present value of retirement benefits increases between ages 51 and 57 because years of service and nominal 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 nominal dollar earnings at age 53. He could choose to start receiving benefits as early as age 55, the pension early retirement age, but the benefit amount would be reduced 5 percent for each year that the receipt of benefits preceded age 65. Because 5 percent is less than the actuarially fair discount rate, the present value of benefits of a person who leaves the firm before age 55 are always greatest if receipt of benefits begins at 55.

Recall that a person who has accumulated 30 years of service and is 55 or older is entitled to increased retirement benefits that would reach 100 percent of normal retirement benefits at age 60. No early retirement reduction is applied to benefits if they are taken then. So a person at age 60 with 30 years of service who continues to work will no longer gain 5 percent a year from fewer years of early retirement reduction, as occurs before age 60. There is a jump in the benefits of a person younger than 60 who attains 30 years of service. That accounts for the jump in the benefits of the person depicted in the figure 6.2*A*, when he attains 30 years of service at age 58.

The Social Security (SS) accrual profile is determined by the Social Security benefit provisions. The present value of accrued Social Security benefit entitlements at age 51 is about $33,000. Social Security benefits cannot begin until age 62. If real earnings do not change much between ages 51 and 62, then real Social Security benefits at age 62 will not change much either. After age 62, the actuarial adjustment is such that the present value of benefits, evaluated at the age of retirement, does not depend on the retirement age. But the present value of the benefits discounted to the same age (51 in this case) declines.

There is a further drop after age 65, because the actuarial adjustment is reduced from 7 percent to 3 percent.

The top line (Tot comp) shows total compensation. For example, the wage earnings of an employee who left the firm at age 60 would increase $482,000 between ages 51 and 60, shown by the wage earnings line. Thereafter, the employee would receive firm pension plan and Social Security retirement benefits with a present value—at age 51—of about $170,000. The sum of the two is about $652,000, shown by the top line. Compared to total compensation of $575,000 between ages 51 and 60, an average of $63,000 per year, total compensation between ages 60 and 65 would be only $100,000, or $23,000 per year. Thus the monetary reward for continued work declines dramatically with age.

Figures 6.2*B*–6.2*D* show comparable compensation profiles for employees who are ages 57, 60, and 64, respectively, in 1982; they have 29, 38, and 45 years of service, respectively. The person depicted in figure 6.2*B* attains 30 years of service at age 58; thus the jump in pension benefits at that age. The present value of pension plus Social Security compensation (Pen + SS) reaches a maximum at age 59 and declines thereafter. Were this employee to continue to work after age 59, until 65, the present value of total retirement benefits would fall by $33,000, offsetting about 28 percent of the present value of wage earnings over this period ($117,000). A similar prospect faces the employee depicted in figure 6.2*C*, but this employee is already entitled to 100 percent of normal retirement benefits and loses benefits for each year that he continues to work.

The employee who faces the figure 6.2*D* compensation profile is 64 years old and loses both pension and Social Security benefits for each year that retirement is postponed. At age 65, for example, about 54 percent of expected wage earnings would be offset by a reduction in retirement benefits, if retirement were postponed.

6.1.2 The 1983 Window

Under the window plan, which was in effect from January 1 to February 28, 1983, all employees were eligible for a separation bonus, but the most generous payments were available to persons age 55 and older who had at least 21 years of service. Retirement benefits for this group were increased depending on age and years of service. For example, a person age 59 with 28 years of service could receive 100 percent of normal retirement benefits, instead of 70 percent under the regular plan. That is, this person's retirement benefit would be increased by 43 percent. A person who was age 55 with 21 years of service could receive 55 percent of the normal benefits, instead of 50 percent. Persons age 60 or older with 30 years of service were eligible for 100 percent of normal benefits under the regular plan.

In addition, all employees were eligible for a separation bonus equal to one week's pay for every year of service, with a minimum of 2 weeks and a maximum of 26 weeks of pay. Thus even persons who were under 55 and those who were eligible for 100 percent of normal retirement benefits faced an added inducement to retire.

6.1.3 The Data

The data used in the analysis are drawn from the personnel records of all persons employed by the firm at any time between 1979 and 1988. A year-end file is available for each year. Earnings records back to 1979 (or to the date of hire, if after 1979) are available for each employee. In addition, the data contain some demographic information, such as date of birth, gender, marital status, and occupational group. The retirement date of employees who retire is also known. (More generally, the date of any departure is known, and the reason for the departure is recorded.) Thus we are able to determine whether a person who was employed at age a was also employed at age $a + 1$, and if not, the exact age at which the employee left the firm.

The estimation of the retirement model in this paper is based on 1982 data, whether or not an employee left the firm in 1982. (To simplify the determination of age of retirement, only employees born in January and February and who had not retired before March 1, 1982, are used in this analysis.) The primary test of the predictive validity of the model is based on how well the model, estimated on 1982 data, predicts retirement under the 1983 window plan that substantially increased standard retirement benefits.

6.2 Departure Rates for Men versus Women

6.2.1 Life-Cycle Departure Rates

Firm departure rates for employees aged 20–70 are shown in figure 6.3. The graph reflects average departure rates over the years 1979–82. After substantial turnover at younger ages, annual departure rates fall continuously to 1–2 percent at ages 45–54. Employees start to leave the firm in larger numbers at age 55, the early retirement age.

Figures 6.4*A*–6.4*C* compare the departure rates for men and women. Figure 6.4*A* pertains to employees with less than 10 years of service, who are not vested in the firm's pension plan. Figure 6.4*B* pertains to employees with 10–29 years of service, and figure 6.4*C* to those with 30 or more years of service. The striking aspect of the graphs is that there is virtually no difference between the departure rates of men and women, except at the principle child-bearing ages—say, 23 to 37. For example, between ages 37 and 54 the turnover rates of men and women with 10–29 years of service are almost identical. Among employees with less than 10 years of service, there is little difference in the departure rates of men and women between ages 37 and 65. Men and women with 30 or more years of service have almost identical departure rates at all ages.

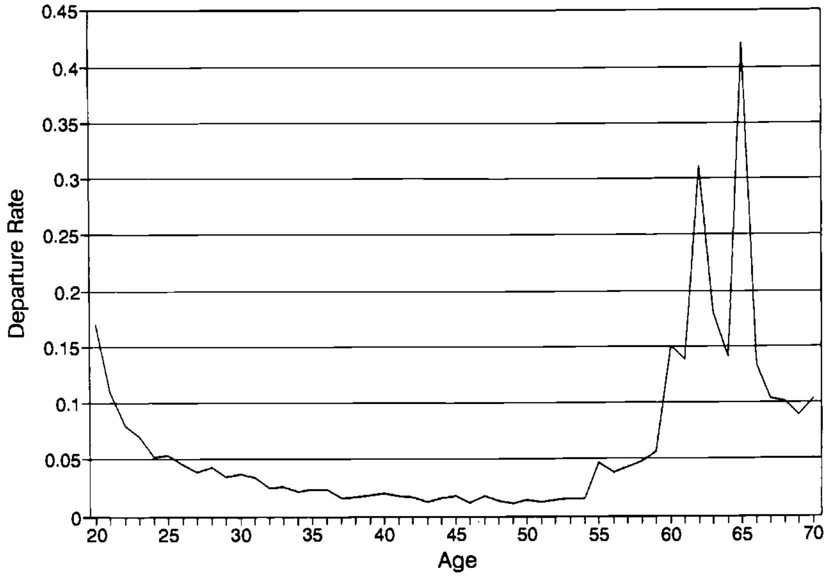


Fig. 6.3 Departure rates by age: men and women together

6.2.2 Retirement-Age Departure Rates

The departure rates for persons aged 50 and above are shown in figure 6.5 for men and women. These rates are based on 1982 data only. There is a noticeable increase in departure rates at age 55, from less than 1 percent for persons 50–54 years old to 3 or 4 percent for employees 55–59 years old. Although the increase in the annual departure may seem small, the cumulative effect of the increase is substantial. For example, with a 4 percent annual departure rate, 19 percent of persons in the firm at age 54 will leave before age 60. At a 1 percent annual rate, only 5 percent will leave.

There is also a sharp increase at age 60, the age at which persons with 30 years of service are entitled to 100 percent of normal (age-65) benefits. The sharp increases at ages 62 and 65 correspond to the Social Security early and normal retirement ages.

The plan provisions suggest that for employees age 55–64, and especially those 55–60 or 61, the departure rate for persons with 30 or more years of service should be higher than the rate for persons with less than 30 years of service. The descriptive data are shown in figure 6.6. The departure rates for men with 30 or more years of service are higher in the 55–61 age range. They are also higher for women at age 60, but the differences at other ages are small. Women with less than 30 years of service appear more likely than men to take early retirement between the ages of 55 and 61.

These data also reveal what may be an individual-specific work effect. Employees with 30 or more years of service who have not retired before age 65 are

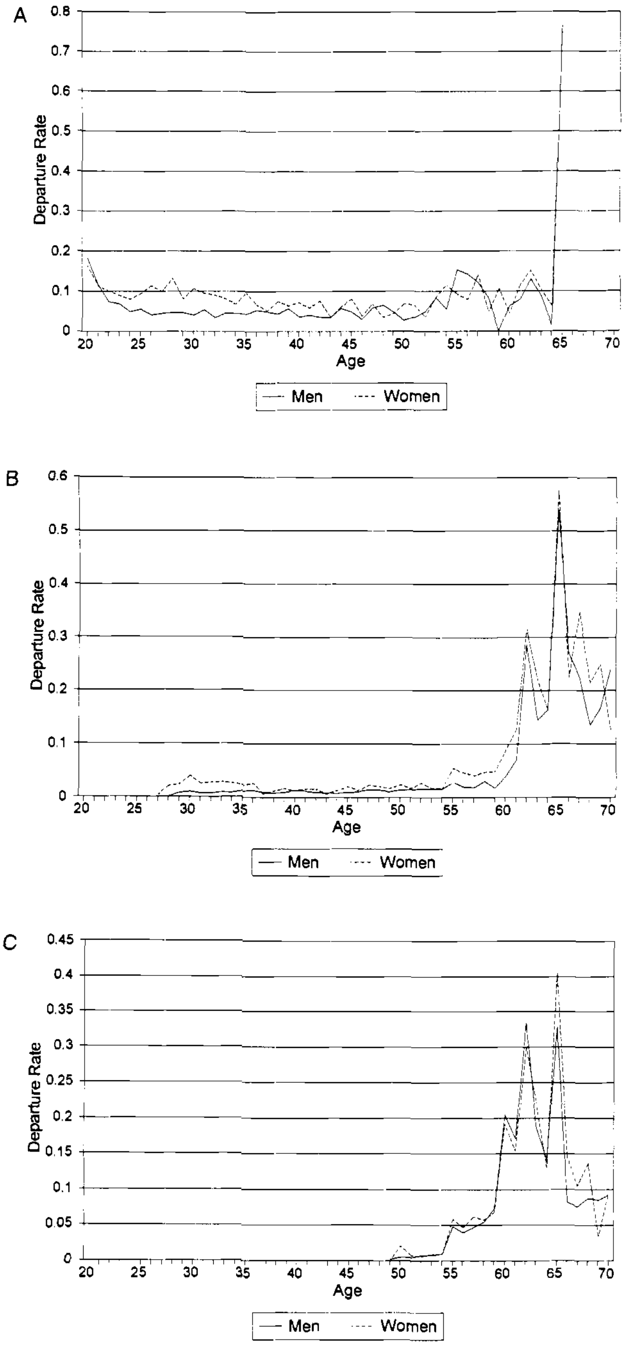


Fig. 6.4 Departure rates by age: men and women (*A*) with less that 10 YOS, (*B*) with 10–29 YOS, and (*C*) with 30 or more YOS

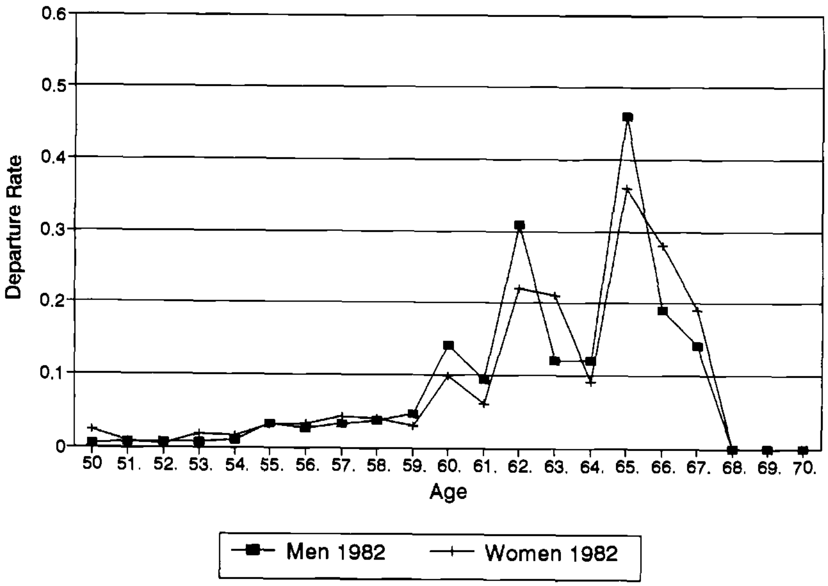


Fig. 6.5 Departure rates: men and women compared, 1982

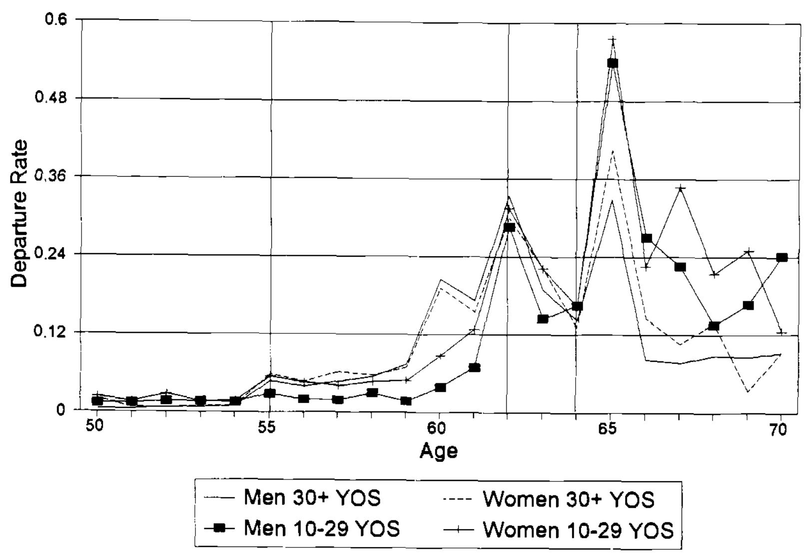


Fig. 6.6 Departure rates by age and YOS: men and women aged 50–70

thereafter less likely to retire than employees with less than 30 years of service.

In summary: even without formal analysis, the graphs make it clear that the pattern of departures reflects the provisions of the pension plan. The pattern is also consistent with Social Security provisions, but the magnitude of the age-65 departure rate seems much more abrupt than the reduction in Social Security benefits at age 65 would suggest. These graphs also make it clear that there is little appreciable difference between the retirement patterns of men and women, with the possible exception of a greater likelihood of early retirement for women with less than 30 years of service. But in general, there is no evidence of a substantial difference in the retirement patterns of men and women.

6.2.3 Window-Plan Retirement Rates

Departure rates under the 1983 window plan are shown for men in figure 6.7*A* and for women in figure 6.7*B*. These rates are contrasted with the average 1982 rates, shown on the same graph. Departure rates under the window plan were typically three to five times as large as the 1982 rates. Like 1982 departures, there was little difference in the departure rates of men and women under the window plan, as shown in figure 6.8. There is, however, some indication that women under age 55 may have been more likely than men to accept the separation bonus.

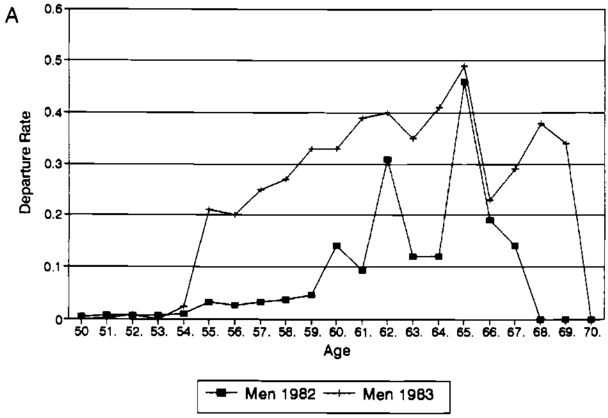
6.3 Formal Models and Prediction of Retirement

6.3.1 Models

Two models are compared during the course of the analysis: the "option value" model and a stochastic dynamic programming model. Both are described in Lumsdaine et al. (1992), and excerpts from that paper are included as an appendix to this chapter. The models are explained only briefly here.

The Option Value Model

At any given age, based on information available at that age, it is assumed 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 74. 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 she retires. With reference to figure 6.1, for example, at age 51 the employee would compare the value of the retirement benefits that she would receive were she to retire then—approximately $87,000—with the value of wage earnings and retirement benefits in



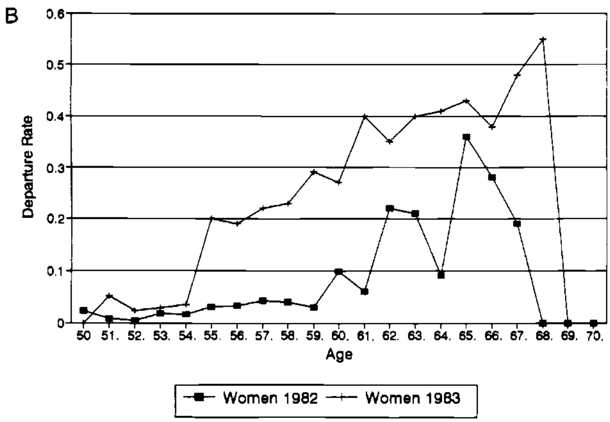


Fig. 6.7 Departure rates: (*A*) men and (*B*) women, 1982 window vs. 1983 window

each future year. The expected present value of retiring at age 60 (discounted to age 51), for example, is about $652,000. Future earnings forecasts are based on the individual's past earnings, as well as on the earnings of other persons in the firm. The precise model specification follows.

A person at age t who continues to work will earn Y_s at subsequent ages 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

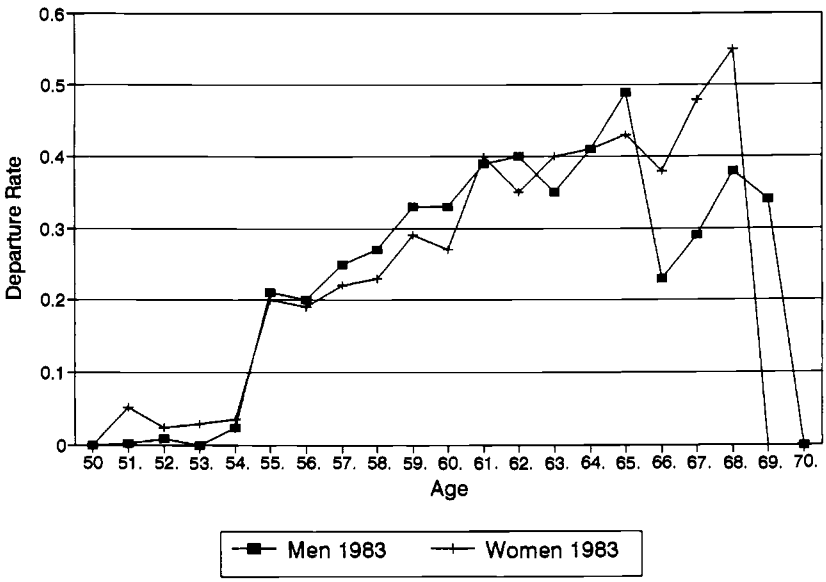


Fig. 6.8 Departure rates: men and women compared, 1983 window

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

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

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

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

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

The utilities of future wage and retirement income are parameterized as

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

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

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

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

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

The parameter k reflects that, in considering whether to retire, the employee's utility *associated* with a dollar of income while retired may be different from her 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 $[k(B_s/Y_s)]^{\gamma}$.

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 at 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.

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 importantly on the error structure that is assumed. Like other users of this type of model, we assume an error structure—and thus a behavioral rule—that simplifies the dynamic programming calculation.[2] In particular, although the option value model allows correlated disturbances, the random disturbances in the dynamic programming model are assumed to be uncorrelated. 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 on its precise implementation. In the version of the dynamic programming model that we implement here, the disturbances are assumed to follow an extreme value distribution.

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.

2. See the Appendix for a more complete description of the error structure.

6.3.2 Results

Parameter estimates are shown in table 6.1. The effect of the special plan provisions for the pre-1951 hires is considered first (cols. [1] and [2]). Estimates for men versus women and stochastic dynamic programming (SDP) versus option value (OV) estimates are then considered (cols. [3]–[8]). Estimates with the "value" of Medicare and firm retiree health insurance benefits counted as equivalent to Social Security benefits, and with firm current employee health insurance benefits counted as equivalent to wage earnings, are also presented (col. [9]). Finally, results imposing firm I parameter estimates (taken from Lumsdaine et al. 1992) on the firm III data are reported (col. [10]).

Pre-1951 Pension Plan Provisions

The estimates in column (1) are based on the assumption that the pre-1951 hires face the same pension plan provisions as later hires. These estimates, as well as those in column (2), are based on a sample of 400 employees. Taken literally, the estimated value of γ (1.045) suggests that with respect to retirement income employees are essentially risk neutral. The estimated value of k is 1.605, implying that a dollar of retirement benefit income—unaccompanied by work—is valued at 60 percent more than a dollar of income accompanied by work. These estimates are very similar to those obtained in our previous work. The estimated value of β, however, is extremely small. If taken literally, it would suggest that in making retirement decisions, future income is given very little weight, compared to income in the current year. Indeed, a value of zero would imply that the decision to retire is based only on the comparison of wage income versus retirement benefits—the replacement ratio—without concern for future possibilities. When the immediate ratio is large enough, the person retires. (Based on our experience elsewhere, we are not inclined to believe this estimate.)

The model fits the data rather well, however, as shown in figures 6.9*A* and 6.9*B*. The principal discrepancy between actual and predicted rates occurs at age 55, where the jump in the predicted rates is noticeably less than the jump in the sample rates. The sample data show a 10 percent departure rate at age 55, which is twice as large as the rates shown in the graphs above, based on larger sample sizes. The predicted and actual cumulative departure rates are very close. Based on the likelihood values, the model fits the data better than a model with dummy variables for each age—that is, better than predictions based on average retirement rates by age. The model does not allow directly for an effect of age on retirement.

The primary test of the predictive validity of the model is how well it predicts retirement rates under the 1983 window plan. The model predictions capture the general pattern of retirement under the window, but substantially overpredict retirement rates between ages 55 and 60, as shown in figure 6.9*C*. The model also predicts some retirements among employees aged 52–54, whereas the actual data show essentially no retirements in this age group.

Table 6.1 **Parameter Estimates by Method and Sample**

Parameter	Pre-1951 Provisions: SDP Model Pre-51 Ignore (1)	Pre-1951 Provisions: SDP Model Pre-51 Correct (2)	Option Value versus SDP Model and Men versus Women: OV Model: Men (3)	OV Model: Women (4)	OV Model: Men and Women (5)	SDP Model: Men (6)	SDP Model: Women (7)	SDP Model: Men and Women (8)	Medicare SDP Model Men and Women (9)	Firm I Comparison SDP Model Men (10)
γ	1.045	0.793	0.599	0.842	0.656	0.898	0.866	0.839	0.789	1.018[a]
	(0.245)	(0.273)	(0.155)	(0.134)	(0.080)	(0.134)	(0.245)	(0.098)	(0.127)	
k	1.605	1.606	2.516	1.628	2.408	1.858	1.446	1.877	1.892	1.881[a]
	(0.147)	(0.052)	(0.722)	(0.360)	(0.272)	(0.250)	(0.362)	(0.035)	(0.073)	
β	0.185	0.185[a]	0.973	0.847	0.963	0.549	0.301	0.564	0.630	0.620[a]
	(0.337)		(0.053)	(0.123)	(0.020)	(0.135)	(0.398)	(0.032)	(0.032)	
σ	0.224	0.152	0.114	0.092	0.120	0.252	0.224	0.229	0.203	0.303
	(0.054)	(0.057)	(0.031)	(0.027)	(0.027)	(0.031)	(0.045)	(0.024)	(0.034)	(0.024)
Log likelihood										
At maximum	122.839	121.768	380.59	100.52	485.64	385.86	101.84	489.69	489.34	387.00
Age	127.051	127.051	362.55	91.26	461.30	362.55	91.26	461.30	461.30	362.55
χ^2										
Fitted data	13.059	12.433	28.24	14.39	30.92	64.30	17.12	59.17	61.78	78.89
Window	78.997	37.782	74.04	44.50	86.52	49.59	15.60	68.69	86.79	89.87
Figure	9A–9C	10A–10C	11A–11C	–	–	12A–12C	–	–	13A–13C	–

Note: Numbers in parentheses are standard errors.

[a]Parameters were fixed at these values.

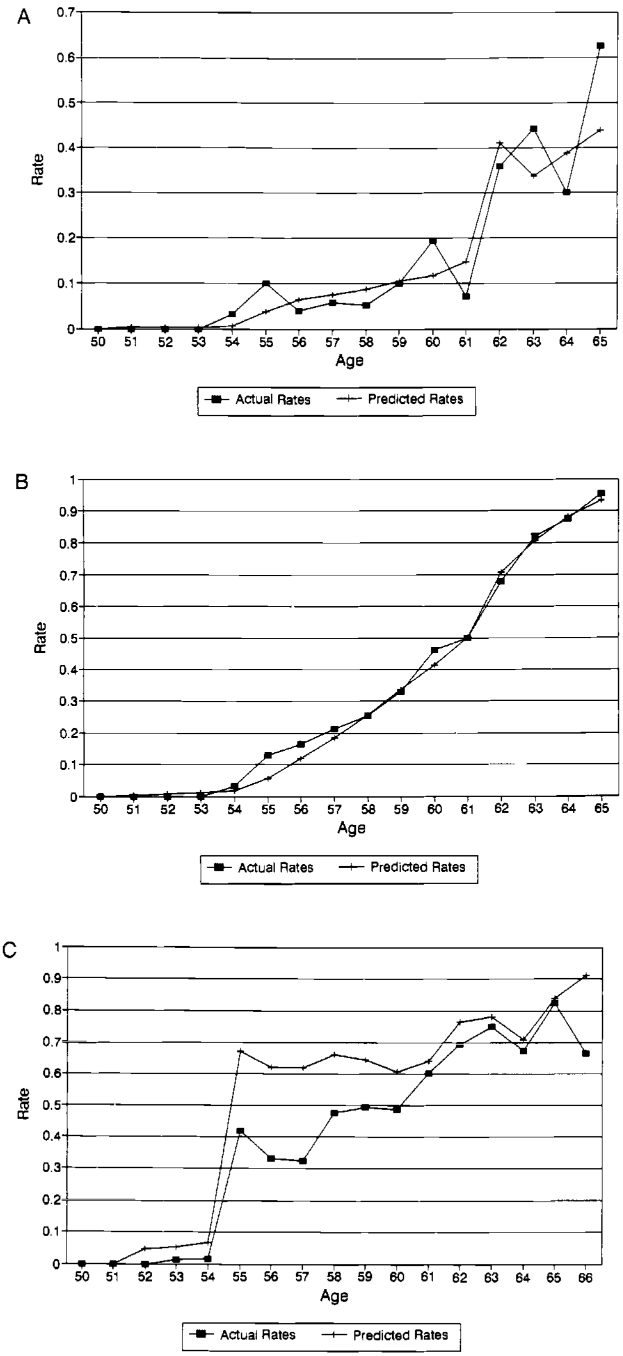


Fig. 6.9 Predicted vs. actual (*A*) annual departure rates, (*B*) cumulative departure rates, and (*C*) window departure rates: men, ignoring pre-1951 provisions

The estimates in column (2) of table 6.1 are obtained if the special pension plan provisions that pertain to employees hired before 1951 are used to determine their options. (Because there are so few retirements among employees younger than 55, we have questioned whether these provisions translate into visible alternatives that are actively considered by older employees, or whether in practice these older employees consider their options to be the same as employees of the same age who are covered by the current plan provisions. Thus we have obtained the two sets of estimates.) The estimated parameter values are very similar to those reported in column (1), although to hasten convergence the discount factor is set in this case. Predicted versus actual rates are shown in figures 6.10*A*–6.10*C*. In general, the fit to actual values is close. The major exception is at age 55, and in this case the actual sample rate is abnormally high; the predicted rate is more in line with typical retirement rates at this age. The difference in the age-55 retirement rate is reflected in the difference between the actual and predicted cumulative rates through age 60. The model predictions of the effects of the 1983 window are very accurate, with the exception of predictions for employees aged 53–54 and 56–57. The actual sample rates for the 56–57 ages are abnormally low; the typical rates are more like the model predictions. Thus for these ages at least, the model predictions give a more accurate indication of actual behavior than the actual sample values.

Stochastic Dynamic Programming versus "Option Value" Estimates

The two sets of parameter values are shown for men (cols. [3] and [6]), for women (cols. [4] and [7]), and for men and women combined (cols. [5] and [8]). In general, the estimated parameters are similar. The most noticeable difference is that the SDP estimated values of β are lower than the OV estimates. For men and women combined, for example, the SDP estimate is 0.564 and the OV estimate is 0.963. The estimated value of k based on the OV model is somewhat larger than the SDP estimate (2.408 vs. 1.877), and the estimated value of γ is somewhat smaller (0.656 vs. 0.839).

The OV model fits the sample data considerably better than the SDP model, based on the likelihood and χ^2 values pertaining to the fitted data. This is revealed graphically for men in figures 6.11*A* and 6.11*B* versus figures 6.12*A* and 6.12*B*. On the other hand, the SDP model predictions of retirement under the 1983 window fit actual retirement rates better than do those of the OV model. This can be seen by comparing figures 6.11*C* and 6.12*C* and in the χ^2 values pertaining to the window. Thus, in general, there is no reason to prefer one model over the other.

Separate Estimates for Men and Women

The estimates for men and women are not statistically different, judged by likelihood ratio tests. The OV model χ^2 statistic is 9.06 (and with four degrees of freedom, the .05 significance level is 9.49), and the SDP model χ^2 statistic

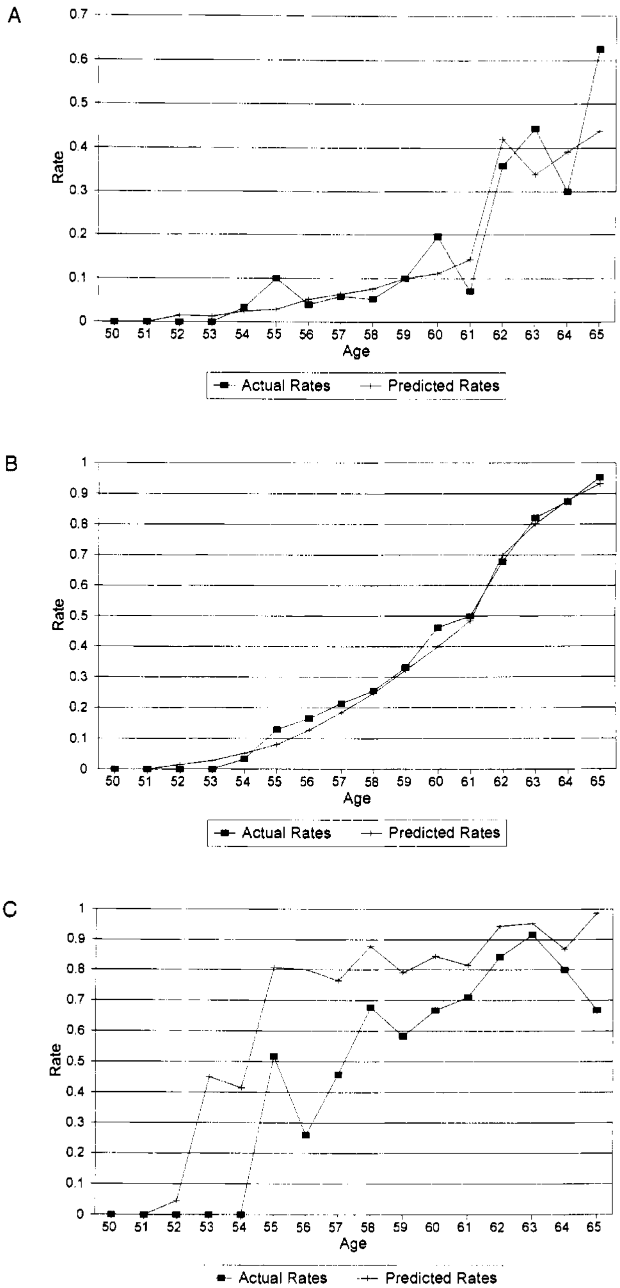


Fig. 6.10 Predicted vs. actual (*A*) annual departure rates, (*B*) cumulative departure rates, and (*C*) window departure rates: men, including pre-1951 provisions

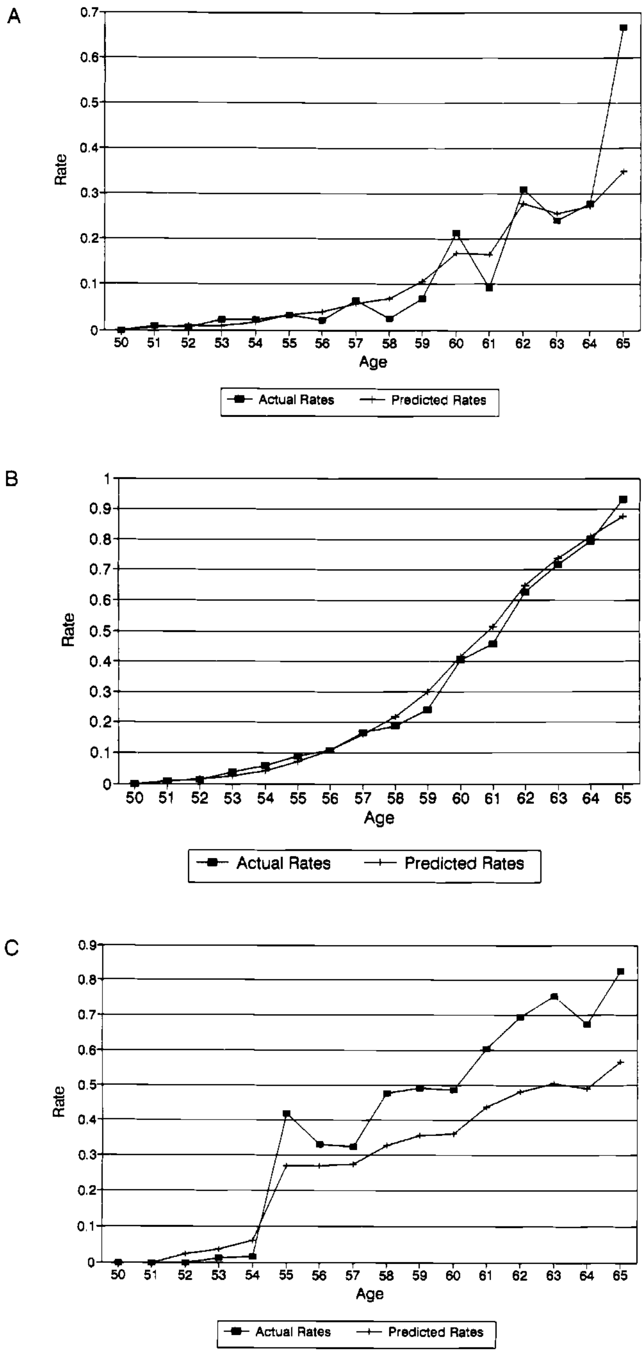


Fig. 6.11 Predicted vs. actual (*A*) annual departure rates, (*B*) cumulative departure rates, and (*C*) window departure rates: men, OV

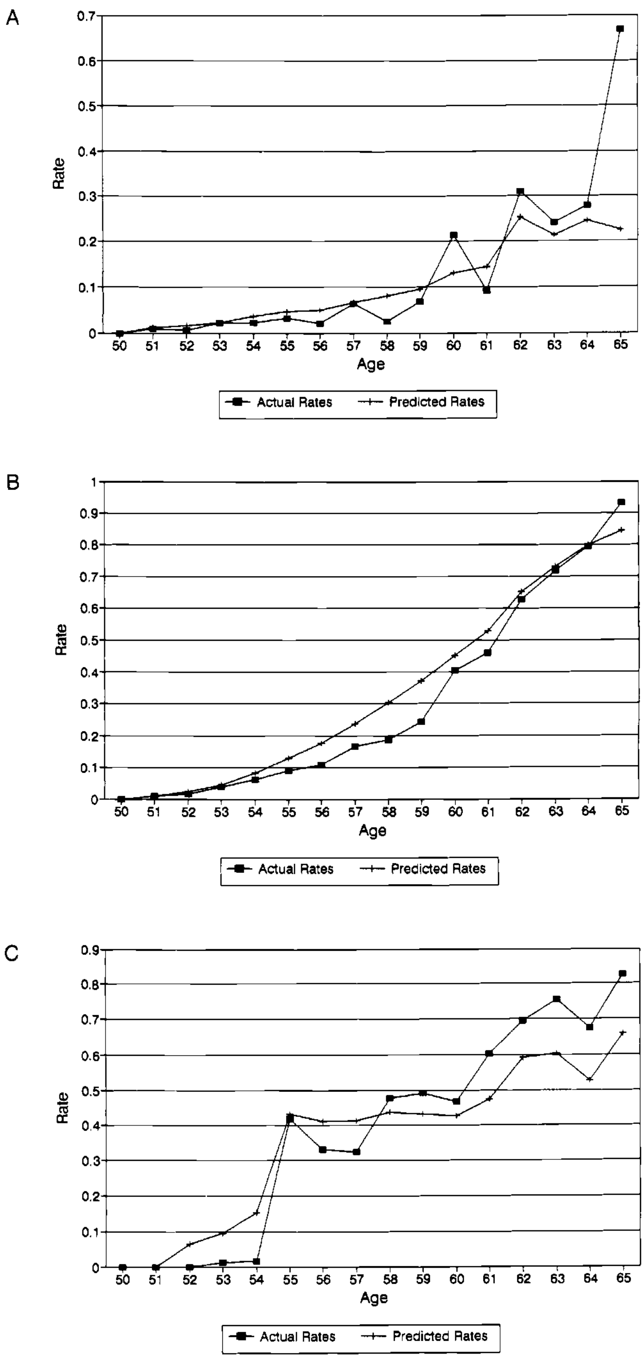


Fig. 6.12 Predicted vs. actual (*A*) annual departure rates, (*B*) cumulative departure rates, and (*C*) window departure rates: men, SDP

is only 3.98, neither of which is statistically significant. The t-statistics for the individual parameters also suggest that the estimates for men and women are not statistically different. Thus the formal estimates appear to be consistent with the graphical evidence in figures 6.4, 6.5, 6.6, and 6.8 showing that departure rates for men and women are virtually indistinguishable after age 40.

Valuing Medicare

The OV and the SDP models underpredict retirement at age 65 for both men and women. A possible reason for the underprediction is that Medicare insurance becomes available at age 65 and provides an inducement to retire similar to the Social Security inducement. But employees at this firm have health insurance while working, and after retirement the same coverage is provided, at no cost to the retiree. For example, a person who retired at age 60 would be covered by retiree health insurance until age 65. After age 65, medical costs up to the Medicare limit would be paid by Medicare, and any additional costs—that are covered by the firm plan—would be paid by the firm retiree insurance. A simple assumption, albeit one that is unlikely to be precisely true, is that medical insurance is valued at its cost, which is treated by employees as comparable to wage or pension benefit compensation. Following this rule, there are three parts to medical coverage: First, while employed at the firm, health benefits are valued at the cost of insurance to the firm.[3] Second, if the person retires before age 65, firm pension benefits are increased by the cost of insurance with coverage comparable to the retiree health insurance.[4] After age 65, Social Security benefits are increased according to the average payment to persons covered by Medicare.[5] Estimates incorporating these assumptions and based on the SDP model for men and women are reported in column (9) of table 6.1.

The parameter estimates are affected very little, relative to comparable estimates without these adjustments, shown in column (8). The likelihood value and the fitted-data χ^2 statistic are almost the same as in the comparable column (8) estimates, that do not account for the value of medical insurance. In particular, the addition of these measures of the value of medical insurance does nothing to explain the departure rate at age 65, as can be seen in figure 6.13*A*. The

3. This cost was estimated by the average cost at large firms for group insurance with coverage like the plan offered by our firm—$105 and $247 for individual and family coverage, respectively, in 1989 dollars. These costs were deflated to 1982 dollars based on a constructed index of Blue-Cross–Blue-Shield premiums per insured person, obtained from the U.S. Health Insurance Institute (1991). We are grateful to Jonathan Gruber for assistance in developing these numbers.

4. The value of this insurance was estimated by increasing the basic group insurance premium according to age, by 5.4 percent per year for each year after age 50. This rate is based on the annual premium costs by age reported by the Congressional Research Service (1988).

5. The costs were estimated based on the average 1986 Medicare payments by age to married and single persons, reported in Shoven, Topper, and Wise (chap. 1 in this volume). The 1986 values were deflated to 1982 dollars based on the Blue-Cross–Blue-Shield index described in n. 3, above. Linear interpolation was used to convert the payments by age interval, reported by Shoven, Topper, and Wise, to payments for each age.

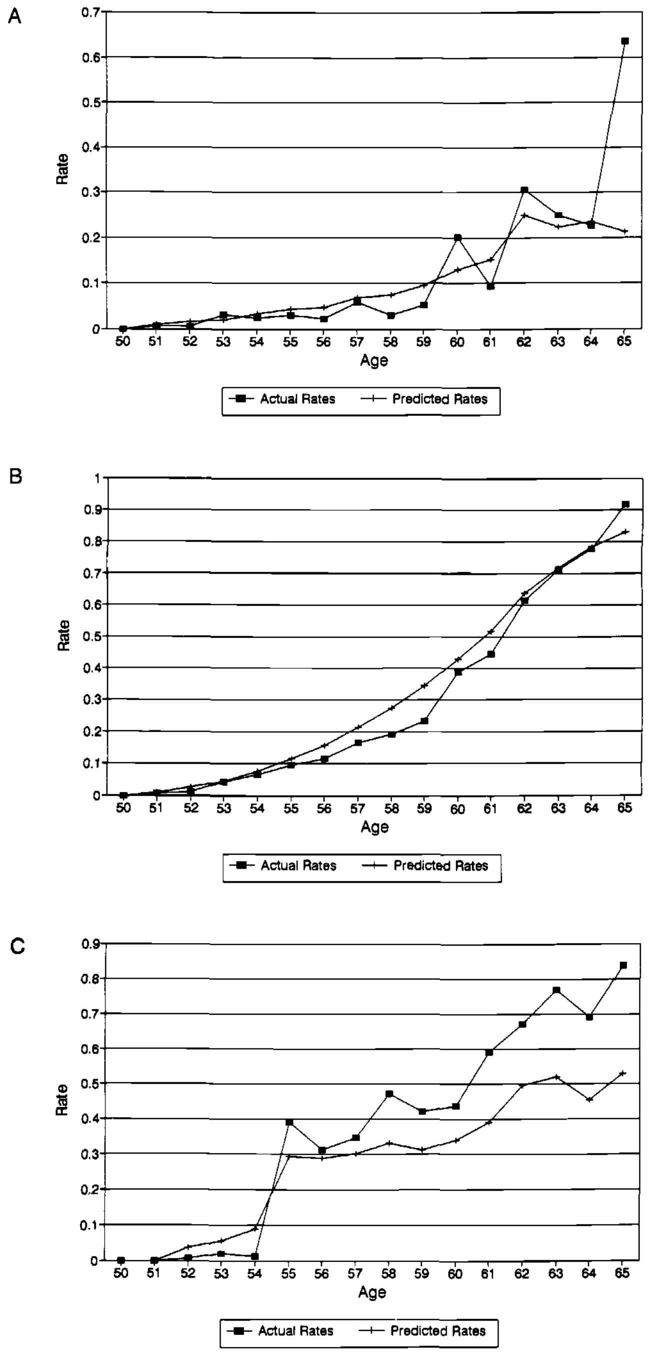


Fig. 6.13 Predicted vs. actual (*A*) annual departure rates, (*B*) cumulative departure rates, and (*C*) window departure rates: men and women, SDP, Medicare

actual rate is .636, but the predicted rate is only .215, somewhat *lower* than the predicted rate without accounting for medical insurance (.277, based on col. [8] estimates). In addition, the model yields worse predictions of retirement under the window plan, judging by the window χ^2 statistic. Thus these results lend no support to the conjecture that retirement at age 65 is strongly affected by the availability of Medicare at that age. However, these exploratory results should not be interpreted to mean that Medicare does not matter. It may well be that the rough specification that we experimented with does not capture the effect of Medicare but that a more careful treatment of the value of medical coverage would show an effect. For example, the assumption that medical insurance is valued at its cost may be incorrect.

Estimates from Firms I and III Compared

The parameter estimates based on firm III data are surprisingly close to those based on firm I data. Results, for firm III, with parameter estimates set to those that we obtained for firm I (Lumsdaine et al. 1992) are shown in column (10) of table 6.1.[6] By comparing the estimates in columns (6) and (10), it can be seen that the firm I estimates for men are very close to the estimates for firm III, based on the SDP specification. The hypothesis that the parameters are the same cannot be rejected, based on a likelihood ratio test. From the χ^2 statistics, however, it is clear that the firm I parameter estimates do not fit the data or predict departure rates under the window plan quite as well as the firm III estimates. Option value model estimates for the two firms (not shown) are also similar but not as close as the SDP estimates, and the hypothesis that the estimates are the same is rejected at the 5 percent level. Again, based on χ^2 statistics, the firm I estimates do not fit the data or predict window departure rates as well as the firm III parameter estimates. On balance, however, the results provide strong confirmation that employees in these two firms react similarly to the incentives inherent in pension plan provisions.

6.4 Summary and Conclusions

The data for firm III confirm a principal conclusion based on firm I data. It is clear that the changes in retirement rates by age correspond closely to provisions of the pension plan. And, as for the results based on firm I, we find that the OV and the SDP models yield similar results. There is no apparent reason to choose one over the other, except based on numerical simplicity. In this case, the OV model fits the sample data better than the SDP model, but the SDP model predicts the window plan retirement rates better than the OV model. We also find that there is essentially no difference between the retirement behavior of men and women. There is some indication that women may be slightly more

6. The scale parameter σ was estimated; this normalization enables comparison of the results from the two firms.

likely than men to take early retirement between ages 55 and 60. But at most ages, the annual retirement rates of men and women are very close. In addition, we explored the possibility that retirement at age 65 is induced by Medicare benefits that become available at that age. Our method of incorporating medical insurance, however, did little to explain the large retirement rates at age 65. Thus we are still left with an "age-65 retirement effect" that is not explained by monetary gain.

Appendix

The "option value" and stochastic dynamic programming models used in the analysis are described.

The Option Value Model

Given the specification as described via equations (1)–(5) in the text, the function $G_t(r)$ can be decomposed into two components,

$$G_t(r) = g_t(r) + \phi_t(r)\,, \tag{A1}$$

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 of the individual effects ω_s and ξ_s, $g_t(r)$ and $\phi_t(r)$ are given by

$$g_t(r) = \textstyle\sum_{s=t}^{r-1}\beta^{s-t}\pi(s|t)E_t(Y_s^\gamma) + \sum_{s=r}^{S}\beta^{s-t}\pi(s|t)[E_t(kB_s(r))^\gamma] - \sum_{s=t}^{S}\beta^{s-t}\pi(s|t)[E_t(kB_s(t))^\gamma], \tag{A2a}$$

$$\phi_t(r) = \textstyle\sum_{s=t}^{r-1}\beta^{s-t}\pi(s|t)E_t(\omega_s - \xi_s), \tag{A2b}$$

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

$$\phi_t(r) = \textstyle\sum_{s=t}^{r-1}\beta^{s-t}\pi(s|t)\rho^{s-t}(\omega_t - \xi_t) = K_t(r)\upsilon_t\,, \tag{A3}$$

where $K_t(r) = \sum_{s=t}^{r-1}(\beta\rho)^{s-t}\pi(s|t)$ and $\upsilon_t = \omega_t - \xi_t$. The simplification results from the fact that at time t the expected value of $\upsilon_s = \omega_s - \xi_s$ is $\rho^{s-t}\upsilon_t$, 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.) Thus, $G_t(r)$ may be written simply as

(A4) $$G_t(r) = g_t(r) + K_t(r)\upsilon_t \,.$$

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

(A5) $$Pr[\text{Retire in year } t] = Pr[g_t(r_t^\dagger)/K_t(r_t^\dagger) < -\upsilon_t] \,.$$

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

(A6) $$Pr[R = \tau] = Pr[g_t(r_t^\dagger)/K_t(r_t^\dagger) > -\upsilon_t, \ldots, \\ g_{\tau-1}(r_{\tau-1}^\dagger)/K_{\tau-1}(r_{\tau-1}^\dagger) > -\upsilon_{\tau-1}, g_\tau(r_\tau^\dagger)/K_\tau(r_\tau^\dagger) < -\upsilon_\tau] \,.$$

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

(A7) $$Pr[R > T] = Pr[g_t(r_t^\dagger)/K_t(r_t^\dagger) > -\upsilon_t, \ldots, \\ g_{T-1}(r_{T-1}^\dagger)/K_{T-1}(r_{T-1}^\dagger) > -\upsilon_{T-1}, g_T(r_T^\dagger)/K_T(r_T^\dagger) > -\upsilon_T] \,.$$

This is a multinomial discrete choice probability with dependent error terms υ_s.

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

(A8) $$\upsilon_s = \rho\upsilon_{s-1} + \varepsilon_s, \quad \text{with } \varepsilon_s \text{ i.i.d. } N(0, \sigma_\varepsilon^2),$$

where the initial value, υ_t, is i.i.d. $N(0,\sigma^2)$ and is independent of ε_s. The covariance between υ_τ and $\upsilon_{\tau+1}$ is $\rho\mathrm{Var}(\upsilon_\tau)$, and the variance of υ_τ, for $\tau > t$, is $(\rho^{2(\tau-t)})\sigma^2 + (\sum_{j=0}^{\tau-t-1} \rho^{2j})\sigma_\varepsilon^2$.

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 υ_τ and $\upsilon_{\tau+1}$ is $\mathrm{Var}(\upsilon_\tau)$, and the variance of υ_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. (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.)

The Stochastic Dynamic Programming 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

$$W_t = \max\{E_t[U_w(Y_t) + \varepsilon_{1t} + \beta W_{t+1}], E_t[\textstyle\sum_{\tau=t}^{S} \beta^{\tau-t}(U_r(B_\tau(t)) + \varepsilon_{2\tau})]\},$$
$$\text{(A9)} \qquad \text{with } W_{t+1} = \max\{E_{t+1}[U_w(Y_{t+1}) + \varepsilon_{1t+1} + \beta W_{t+2}],$$
$$E_{t+1}[\textstyle\sum_{\tau=t+1}^{S} \beta^{\tau-t-1}(U_r(B_\tau(t+1)) + \varepsilon_{2\tau})]\},$$

etc., 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, i.e., 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 (A9) can be written as

$$W_t = \max\{\bar{W}_{1t} + \varepsilon_{1t}, \bar{W}_{2t} + \varepsilon_{2t}\}, \quad \text{where}$$
$$\text{(A10)} \qquad \bar{W}_{1t} = U_w(Y_t) + \beta\pi(t+1 \mid t)E_t W_{t+1} \quad \text{and}$$
$$\bar{W}_{2t} = \textstyle\sum_{\tau=t}^{S} \beta^{\tau-t}\pi(\tau \mid t)U_r(B_\tau(t)).$$

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 (74), he retires and receives expected utility $\bar{W}_{2t_{74}}$.

Recursions and Computation

With a suitable assumption on the distribution of the errors ε_{it}, the expression (A10) 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 (1988), 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

$$E_t W_{t+1}/\sigma \equiv \mu_{t+1}$$
$$\text{(A11)} \qquad = \gamma_e + \ln[\exp(\bar{W}_{1t+1}/\sigma) + \exp(\bar{W}_{2t+1}/\sigma)]$$
$$= \gamma_e + \ln[\exp(U_w(Y_{t+1})/\sigma)\exp(\beta\pi(t+2 \mid t+1)\mu_{t+2}) + \exp(\bar{W}_{2t+1}/\sigma)],$$

where γ_e is Euler's constant. Thus equation (A11) can be solved by backward recursion, with the terminal value coming from the terminal condition that $\mu_{t_{74}} = \bar{W}_{2t_{74}}$.

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

$$\text{(A12)} \quad Pr[\text{Retire in year } t] = Pr[\bar{W}_{1t} + \varepsilon_{1t} < \bar{W}_{2t} + \varepsilon_{2t}]$$
$$= \exp(\bar{W}_{2t}/\sigma)/[\exp(\bar{W}_{1t}/\sigma) + \exp(\bar{W}_{2t}/\sigma)].$$

Gaussian errors. Following Daula and Moffitt (1989), 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:

$$\text{(A13)} \qquad Pr[\text{Retire in year } t] = Pr[(\varepsilon_{1t} - \varepsilon_{2t})/\sqrt{2}\sigma < (\bar{W}_{2t} - \bar{W}_{1t})/\sqrt{2}\sigma] = \Phi(a_t),$$

where $a_t = (\bar{W}_{2t} - \bar{W}_{1t})/\sqrt{2}\sigma$. Then the recursion (A10) becomes

$$\text{(A14)} \qquad E_t W_{t+1}/\sigma \equiv \mu_{t+1} = (\bar{W}_{1t+1}/\sigma)[1 - \Phi(a_{t+1})] + (\bar{W}_{2t+1}/\sigma)\Phi(a_{t+1}) + \sqrt{2}\phi(a_{t+1})$$

where $\phi(\bullet)$ denotes the standard normal density and $\Phi(\bullet)$ denotes the cumulative normal distribution function. As in equation (A13), $\Phi(a_t)$ 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 (A14).

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

$$\text{(A15a)} \qquad U_w(Y_t) = Y_t^\gamma ,$$

$$\text{(A15b)} \qquad U_r(B_t(s)) = [\eta k B_t(s)]^\gamma ,$$

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: $\eta = \exp(\lambda z + 1/2\, \lambda^2)$, 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

$$\text{(A16a)} \qquad U_w(Y_t) = Y_t^\gamma + \zeta,$$

$$\text{(A16b)} \qquad U_r(B_t(s)) = [kB_t(s)]^\gamma ,$$

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

In summary, the dynamic programming models are given by the general recursion equation (A9). It is implemented as shown in equation (A11) under the assumption that the ε_{it} are i.i.d. extreme value, and as shown in equation (A14) under the assumption that ε_{it} are i.i.d. normal. The retirement probabilities are computed according to equations (A12) and (A13) respectively. The fixed effects specifications are given by equations (A15) and (A16). The unknown parameters to be estimated are γ, k, β, σ, and 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.*

References

Ausink, John A. 1991. The effect of changes in compensation on a pilot's decision to leave the air force. Ph.D. dissertation, Harvard University.

Berkovec, James, and Steven Stern. 1991. Job exit behavior of older men. *Econometrica* 59(1): 189–210.

Bulow, J. 1981. Early retirement pension benefits. NBER Working Paper no. 654. Cambridge, Mass.: National Bureau of Economic Research.

Congressional Research Service. 1988. Costs and effects of extending health insurance coverage. Washington, D.C.: Library of Congress.

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. Gilray, D. K. Horne, and D. Alton Smith. Alexandria, Va.: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kotlikoff, Laurence J., and David A. Wise. 1985. Labor compensation and the structure of private pension plans: Evidence for contractual versus spot labor markets. In *Pensions, labor, and individual choice,* ed. D. Wise, 55–85. Chicago: University of Chicago Press.

———. 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.

———. 1988. Pension backloading, wage taxes, and work disincentives. In *Tax policy and the economy,* vol. 2, ed. L. Summers, 161–96. Cambridge: MIT Press.

Lazear, Edward P. 1983. Pensions as severance pay. In *Financial aspects of the United States pension system,* ed. Z. Bodie and J. Shoven, 57–85. 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. Windows and retirement. *Annales d'Economie et de Statistique* 20/21:219–42.

———. 1992. Three models of retirement: Computational complexity versus predictive validity. In *Topics in the economics of aging,* ed. D. Wise, 19–57. Chicago: University of Chicago Press.

Stock, James H., and David A. Wise. 1990a. Pensions, the option value of work, and retirement. *Econometrica* 58, no. 5 (September): 1151–80.

———. 1990b. The pension inducement to retire: An option value analysis. In *Issues in the economics of aging,* ed. D. Wise, 205–24. Chicago: University of Chicago Press.

U.S. Health Insurance Institute. 1991. *Sourcebook of health insurance data, 1990.* New York: U.S. Health Insurance Institute.

Comment John Rust

This paper continues and extends earlier research by the authors (Lumsdaine, Stock, and Wise 1991) that compares two different dynamic structural models—the Stock-Wise (1990) "option-value" (OV) model and a dynamic programming (DP) model—on the basis of their ability to make accurate out-of-sample predictions of the effect of "window plans," which provide special pension incentives to retire from a firm. I like this research very much, especially for its clever use of the window plans as "natural experiments" to evaluate predictive validity of structural econometric models. In my survey of the literature (Rust, in press), I cite the authors' results as one of the clearest demonstrations of the potential payoffs to dynamic structural modeling: both the OV and DP models do a much better job of predicting the large increase in firm departure rates than any of the traditional "reduced-form" econometric specifications. These are the kind of results that Marschak (1953) and Lucas (1976) must have had in mind when they wrote their critiques of the traditional reduced-form econometric approaches to policy evaluation. However, for reasons I discuss below, the structural forecasts from these models are still a long way from being definitive. I think the most important contribution of this paper is that it provides an excellent example of the kinds of discriminating out-of-sample predictive tests we ought to be subjecting our econometric models to. Unfortunately, too many structural estimation exercises are little more than displays of technique that do not seriously attempt to evaluate model performance. I believe that this research will set a new standard for demonstrating the credibility of structural econometric models.

Although there are some differences in the predictions of the two models (with the OV model fitting better in-sample and the DP model fitting better out-of-sample), as well as significant differences in the parameter estimates (with the OV model yielding a significantly higher estimate of the postretirement "leisure value" of income K, and the DP model yielding a significantly lower estimate of the discount factor β), I am more struck by the overall similarity of the predictions of the two models rather than their differences. The similarity is striking in view of the large conceptual differences in the two approaches and is suggestive of an identification problem I have discussed elsewhere (Rust, in press). Even within the class of DP models, there are different parameter combinations that generate similar retirement behavior: for example, the predictions of DP models with a low value of β, a high value of K, and a high value of γ (the coefficient of relative risk aversion, with $\gamma = 1$ corresponding to risk neutrality) look very similar to models with a high value for β and lower values for K and γ. We can also see this effect in comparing the estimated parameters of the OV and DP models in columns (3) and (6) of

John Rust is professor of economics at the University of Wisconsin and a research associate of the National Bureau of Economic Research.

table 6.1. The fact that very different parameter estimates yield similar behavioral predictions is a bit disturbing. The problem is even more pronounced in the authors' earlier paper, in which a DP model with normally distributed unobservables has $\hat{\beta} = .916$ and $\hat{K} = 2.975$, whereas a comparably fitting extreme value specification has $\hat{\beta} = .62$ and $\hat{K} = 1.88$. Despite the fact that the normal specification seems to put much more weight on the future than the extreme value specification, the fit and predictions of the two models is very similar. Undoubtedly, the reason for this similarity can be traced to important differences in the variance of the error terms and the specification of unobserved heterogeneity. However, for prediction purposes, this raises some difficult issues: which is the "right" behavioral model, OV or DP? Even if we limit ourselves, a priori, to the narrower class of DP models, it is not difficult to show that there are equivalence classes of distinct error distributions and parameter values that yield similar behavioral predictions. We might be tempted to believe that as long as the predictions of these different models are similar, the question of deciding which is the "right" model is moot. For reasons I discuss below, I think it is likely that there are many hypothetical pension policies for which the predictions of these models will be significantly different, such as pension plans where there is nonnegligible uncertainty regarding future payoffs (as in defined contribution plans). I would like to see more work in delineating the conditions under which the predictions of their various models diverge—this might suggest new experiments that might help us discriminate between the various models.

Another problem is that the structural coefficient estimates vary substantially from data set to data set. For example the OV model estimated in the previous paper had $\hat{\beta} = .895$, which is significantly lower than the .962 value reported here, and the estimated value of K is 2.42 in this data set, versus 1.47 in the earlier paper. How are we to interpret this parameter variability? If people really are this heterogeneous, will we have to gather separate data to forecast the retirement behavior at each different firm? One way to get a handle on this issue is to see how well out-of-sample forecasts using the parameters from their previous data set perform on the current data, and vice versa.

The remainder of my comments focus on identifying the situations where the predictions of the OV and DP models will be similar and the situations where their predictions will diverge. As the authors have noted in their paper, the OV and DP models differ primarily in the interchange of the orders in which the maximization and expectation operators are taken. In the OV model, at each time t the worker calculates the maximum of the expected discounted utility of retiring at different future dates, whereas in the DP model the worker compares the value of retiring now with the value of continuing to work at least one more period, where the value function (or "shadow price") is calculated as the expectation of the maximum of the discounted utility retiring next period versus continuing to work in period $t + 1$. In the OV model, a worker retires in the first period t in which the expected utility of immediate retirement exceeds the expected utility of retiring at any future date $t + r$, whereas in the DP model, the worker retires in the first period t where the value of retiring at time t exceeds the value of continuing to work until date $t + 1$. In a certain sense, the OV model seems to take a longer view of the future, since it calculates the value of retiring at all future dates $t + r$, whereas the DP model appears to look only one period ahead at the value function V_{t+1}. However, since V_{t+1} is calculated by backward induction, it provides the correct valuation of discounted expected utility, recursively generating the optimal retirement policy. The OV calculations are not based on a correct evaluation of expected discounted utility and, hence, will generally not yield an optimal decision rule. Indeed, the very terminology "option value model" is actually a misnomer since the Stock-Wise model ignores the option value of staying on the job to take advantage of a possible increase in future salary, an option that is lost if one retires (since retirement is treated as an absorbing state in this model). By evaluating future utility as the expected value of the maximum of the options of staying with the firm at least one more period versus retiring now, the DP model assigns the correct value to this option. Given that an OV decisionmaker is in a sense myopic, or "time inconsistent," why is it that the predictions of the DP and OV models are so similar?

The answer is that the predictions of the two models will coincide provided that the level of future uncertainty is sufficiently small: in that case the maximum of the expected values is close to the expected value of maximum. In particular, if there is little uncertainty about future wages, then the option value of possible future wage increases is negligible, so the predictions of the OV model, which ignores this option, will be very similar to the DP model, which explicitly accounts for it. However in models where there are significant future uncertainties, we can expect that the predictions of the models will be very different. Steven Stern (1994) was one of the first to make this point, and his paper provides numerical examples showing conditions under which the OV and DP decision rules diverge. His Monte Carlo experiments show that if agents are really behaving according to a DP model and one tries to approximate this behavior using an OV model, the parameter estimates and predictions of the OV model can be significantly biased "because . . . the option value of working is significant relative to the option value of retiring" (p. 6). On the other hand, Stern found that biases resulting from incorrect specification of the particular distribution of the error term were insignificant, provided one stays within the basic DP framework.

It is easiest to illustrate the conditions under which the DP and OV solutions differ in a simple two-period model, but the logic extends to any number of periods. Consider the model with no uncertainty, first. Suppose the worker is risk neutral and earns wage $W_1 = W_2$ if working in periods 1 and 2, and retirement benefit rW_1 if retired, where r is the replacement rate (if the worker retires at the beginning of period 2, he also receives retirement benefits rW_1). Thus, at time t the utility of working is W_1 and the utility of retiring is KrW_1, where

K is the parameter reflecting the additional value of leisure in retirement. For concreteness assume that $r = 3/4$ and $K = 8/5$, so the utility of retiring equals $(6/5)W_1$. Now consider the worker's decision problem at time 1 under the OV model: the worker compares the utility of retiring immediately, $(6/5)W_1 + (6/5)W_1 = (12/5)W_1$, with the utility of retiring at time 2, $W_1 + (6/5)W_1 = (11/5)W_1$, or of not retiring at all, $W_1 + W_2 = (10/5)W_1$. The optimal decision is clearly to retire at time 1. Now consider solving the problem by DP. We start in the last period and note that if the worker is already retired his value function is $V_2 = (6/5)W_2 = (6/5)W_1$, whereas if the worker has not yet retired his value function $V_2 = \max[W_2, (6/5)W_1]$, which also equals $(6/5)W_1$. Then in period 1 an unretired worker compares the value $W_1 + V_2 = (11/5)W_1$ of continuing to work with the value $(6/5)W_1 + V_2 = (12/5)W_1$ of retiring immediately, and we see that the DP and OV decision rules coincide with the recommendation to retire in period 1.

Now suppose that period 2 wages are uncertain but are not expected to increase: $E\{\tilde{W}_2|W_1\} = W_1$. It is easy to see that the OV model predicts that retirement in period 1 is still optimal: since $\tilde{W}_2$ has conditional expectation W_1, the earlier calculations are unchanged. The presence of uncertainty does change the DP decision rule, which differs from the OV decision rule when the variance in $\tilde{W}_2$ is sufficiently large. In period 2, we have $V_2 = (6/5)W_1$ if the worker is already retired, and $V_2 = \max[\tilde{W}_2,(6/5)W_1]$ if the worker is not yet retired. The optimal decision rule is to continue working in period 2 provided $\tilde{W}_2 > (6/5)W_1$. In period 1, an unretired worker compares the expected value of retiring immediately, $(12/5)W_1$, with the expected value of continuing to work one more period, $W_1 + E\{V_2|W_1\}$. If $E\{V_2|W_1\} > (7/5)W_1$, then the worker will decide to continue to work rather than retire in period 1, and the difference in these terms gives the loss in utility of following the OV decision rule. This difference is positive provided that the variance in $\tilde{W}_2$ is sufficiently large. In such a situation a worker would want to use DP to calculate the optimal retirement policy, since it correctly accounts for the option value of continued employment, yielding higher utility than the OV decision rule.

In the specifications estimated in this paper, the major computational simplification is the assumption that there is no future uncertainty about wages: workers have perfectly certain point predictions about their future wage profiles at the firm (computed from regressions on individuals' wage histories). The only uncertainty in the OV model comes via two error terms corresponding to unobservable factors affecting the utility of work and retirement. Although these error terms are treated as random walk processes, the estimated standard deviation of the innovation of the random walk is relatively small (.11), so the error components are not a major source of future uncertainty. In the DP model, the error terms entering the period utility functions are assumed to be i.i.d., and the estimated standard deviation of these error terms is also small (.22). Thus, there are two factors responsible for the similarity in the predictions of the OV and DP models: (1) the authors do not allow for uncertainty in future wage profiles, and (2) the estimated variances of random utility components are small.

My guess is that we would see much more significant differences between the DP and OV models once we allow for future wage uncertainty. It is not clear how uncertain workers are about their future wages, but this is clearly an empirical matter that needs to be addressed. At a minimum, it would be interesting to see the R^2 values from the wage regressions (which are not reported). I have estimated DP models that allow for wage uncertainty (Phelan and Rust 1991; Rust and Phelan 1993), predicting the entire distribution of future wages, not just the mean. I find substantial variation in future wage earnings even conditional on the worker's past wage history, in real terms. Part of this uncertainty is undoubtedly an artifact of aggregation of planned hours of work, since my model forecasts future earnings over an entire year, conditional on next year's employment being in one of two categories, full- or part-time. It is plausible that many workers have a fairly good idea of their future wages on a week-by-week or month-by-month basis. Thus, to some extent, the appropriate level of variance in future wages depends on the fineness of the time grain of the decision problem. However, many workers do face significant uncertainties regarding the level of future bonuses (including the window plan itself). These uncertainties may induce a significant option value for remaining with the firm. I would have liked to have seen a fuller discussion of this issue, since it seems to be the most important factor underlying the similarities in the predictions of these two very different models.

Beyond wage earnings, the major source of uncertainty neglected by the models in this paper is health-care costs. All of the models estimated in this paper and in the authors' earlier work miss the big peak in retirements at age 65. Although one might argue that the age-65 peak reflects a sociological effect, I think a much more compelling economic story is that the peak reflects an interaction effect between private health insurance and the "Medicare option": unhealthy workers choose to remain under the firm's more generous group health plan until they are able to supplement their pension health-care coverage with Medicare insurance at age 65, whereas relatively healthy workers take advantage of the window plan, since they are able to purchase supplemental private medical insurance at attractive rates. The authors include Medicare by adding the average value of Medicare payments to the retirement benefit calculated without Medicare and find "little support for the value of Medicare that we have used here." I do not think that this is the proper way to treat Medicare: risk-averse workers are concerned about the small chance of uninsured catastrophic health-care costs, so the certainty equivalent value of Medicare is much higher than the expected value of benefits paid. It seems likely that many of the workers in the age-65 peak could have health problems that would make it prohibitively costly for them to purchase supplemental private health insurance if they retire before age 65. Many of these workers will prefer to remain with the firm to take advantage of the group health-care benefits until Medicare kicks in at age 65, removing a substantial share of the risks of catastrophic health-care costs. Indeed, this is what I have found after explicitly incorporating the uncertainties surrounding health-care costs into my DP model (with health-care expenditures turning out to be almost perfectly approximated by the Pareto distribution). Our recent paper (Rust and Phelan 1993) shows that by explicitly modeling the distribution of health-care costs (as opposed to treating it as an expected value), we obtain DP models that are able to capture the peak in retirement at age 65 that the Lumsdaine, Stock, and Wise approach misses.

There is a final option that the models in this paper ignore: the option of returning to work after retirement. The authors' models rule out this option by assumption: retirement is treated as an absorbing state. One might argue that future re-employment opportunities are implicitly captured in the error terms. However, the fact that the error terms are i.i.d. in the DP model makes this interpretation less plausible: the calculation of the value of retiring does not properly value the option of finding an attractive new job after retiring from the current job. In reality, some workers know that their employment record is sufficiently good that they will always have the option of returning to their career job or a similar job if they find that retirement does not suit them, experience unexpected financial problems, or encounter an unexpected, attractive job offer. The possibility of the re-employment option may be a crucial element of a risk-averse worker's decision to take advantage of an early retirement plan. In Rust (1990) and Rust and Phelan (1993), I show that in the Retirement History Survey, less than half of all work histories involve the discontinuous employed/retired trajectory implicit in the authors' model. Indeed, at least one-third of all work histories involve multiple transitions in and out of the labor force, suggesting that postretirement work in a noncareer job is a fairly common phenomenon (similar results can be found in Berkovec and Stern 1991, using the National Longitudinal Survey). It is not clear what kinds of specification errors are generated by ruling out the possibility of postretirement work. For the fraction of the sample who follow the "traditional" retirement profile, preferring to remain permanently out of work, the error may not be substantial. However for risk-averse workers who face substantial financial uncertainties at the brink of retirement, ruling out the re-employment option substantially increases the riskiness of the retirement decision. The model may be able to generate a good overall fit to the data by underestimating β and overestimating the value of leisure, K, but it is not clear whether the resulting model will do a good job of tracking the behavior of various subgroups of employees with differing views of their re-employment prospects.

From a larger perspective, we are not only concerned about departure rates from particular firms, but the overall dynamics of employment at the end of the life cycle. The simple models estimated here are unable to address this larger question. Although the DP model is sufficiently flexible that it can be generalized to address a broad variety of issues, including uncertain future health, earnings, and re-employment prospects (see Rust 1989; Berkovec and

Stern 1991; Rust and Phelan 1993), it is unclear whether the OV model can be generalized to encompass these features. In their earlier paper, the authors stressed the point of computational complexity, arguing that the OV calculations are simpler to implement than the DP calculations. However I have seen no clear-cut evidence that this is so: both the OV and DP models can be solved very quickly on a 386 computer, and if there is any speed advantage to the OV model, it does not seem to be substantial. Viewing the brain as a massive parallel processor, I think few would disagree that people have far greater reasoning abilities than a 386 computer. Therefore I find it hard to buy the story that humans adopt the OV decision rule because it is simpler to implement than a DP decision rule.

Once we allow for the possibility of re-employment after retirement, there is no obvious way to calculate the value of retirement other than the backward induction process of DP. While I view the OV model as a very useful point of departure for addressing the specific question of modeling the effect of firm pension plan provisions on exit rates, I think it has real limitations in its ability to properly account for the options arising from uncertainty about future earnings and employment possibilities. However at the same time, I must stress that there is nothing sacred about DP models: which type of model best describes people's behavior is ultimately an empirical issue. Even though all of these models will probably ultimately be regarded as stepping stones toward more realistic behavioral models, it is clear that the authors' careful attempts to discriminate between their simple OV and DP specifications have contributed a great deal to our understanding of retirement behavior.

References

Berkovec, J., and S. Stern. 1991. Job exit behavior of older men. *Econometrica* 59(1): 189–210.

Lucas, R. E., Jr. 1976. Econometric policy evaluation: A critique. In *The Phillips curve and labour markets,* ed. K. Brunner and A. K. Meltzer. Amsterdam: North-Holland.

Lumsdaine, R., J. Stock, and D. Wise. 1991. Three models of retirement: Computational complexity vs. predictive validity. In *Topics in the economics of aging,* ed. D. A. Wise, 19–57. Chicago: University of Chicago Press.

Marschak, T. 1953. Economic measurements for policy and prediction. In *Studies in econometric method,* ed. W. C. Hood and T. J. Koopmans. New York: Wiley.

Phelan, C., and J. Rust. 1991. U.S. Social security policy: A dynamic analysis of incentives and self-selection. Department of Economics, University of Wisconsin—Madison. Manuscript.

Rust, J. 1989. A dynamic programming model of retirement behavior. In *The economics of aging,* ed. D. A. Wise, 359–98. Chicago: University of Chicago Press.

———. 1990. Behavior of male workers at the end of the life-cycle: An empirical analysis of states and controls. In *Issues in the economics of aging,* ed. D. A. Wise. Chicago: University of Chicago Press.

———. 1992a. Do people behave according to Bellman's principle of optimality? Hoover Institution Working Paper no. E-92-10. Stanford, Calif.: Hoover Institution.

———. 1992b. A dynamic programming model of retirement behavior. In *Computer* *assisted analysis and modelling on the IBM 3090,* ed. K. R. Billingsley, H. U. Brown III, and E. Derohanes, vol. 2, 887–912. Athens: Baldwin Press, University of Georgia.

———. In press. Structural estimation of Markov decision processes. In *Handbook of econometrics,* ed. R. Engle and D. McFadden, vol. 4.

Rust, J., and C. Phelan. 1993. How Social Security and Medicare affect retirement behavior in a world of incomplete markets. Department of Economics, University of Wisconsin—Madison. Manuscript.

Stern, S. 1994. Approximate solutions to stochastic dynamic programs. *Econometric Theory.* Forthcoming.

Stock, J., and D. Wise. 1990. Pensions, the option value of work and retirement. *Econometrica* 58(5): 1151–80.

Comment James P. Smith

In the first-generation economic research on retirement, the emphasis was on learning about the role of Social Security. In that work, private pensions were assigned a largely secondary role, partly because the data on private pensions were so limited. Largely due to the impetus of the NBER's program project on the economics of aging, the best of the second-generation studies demonstrated how critical the provisions of private pension plans were for retirement. Lumsdaine, Stock, and Wise have been the major contributors to that literature, and in this paper, they offer us another impressive addition to their body of work. In my view, this trio of authors is largely responsible for convincing the economics and policy community about the importance of private pensions. Their earlier work was based on data from a single firm. This paper generalizes that work in an important and necessary direction by using data from another large private-sector firm.

This paper sets out four principal goals: (1) to compare the predictive value of options value and dynamic programming models, (2) to evaluate how well pension plan provisions predict retirement behavior, (3) to compare how well these models predict retirement behavior of men and women, and (4) to evaluate the effect of Medicare on retirement. Much of the emphasis is devoted to the first two aims, with considerably less attention paid to the last two. My main suggestion, in fact, is that the value added of their work could be higher if that emphasis were reversed.

I found the battle between the two dominant theoretical models the least interesting part of the paper. This battle is part of a war no one can win. Indeed, winning or losing the battle as described here has little to do with the war. As the authors recognized, their dynamic programming model is a very specialized and quite restrictive case of an infinite variety of equally plausible alterna-

James P. Smith is RAND Chair Economist in Labor Markets and Demographic Studies and director of Rand's Labor and Population Program.

tives. They are able to test particular variants of two theoretical frameworks, but this is certainly not a test of the relative usefulness of the option value and the dynamic programming approaches. Their work here, and similar tests contained in their earlier papers, demonstrate that both theoretical approaches are useful and give reasonable results. Perhaps we should leave it at that and, within either model, get on to the more substantive issues that occupy the rest of their paper.

To obtain out-of-sample tests of their model, Lumsdaine et al. predict departure rates during years when the firm instituted a window plan. This use of window plans to test retirement models is growing in popularity, but I want to sound some cautionary notes because it makes me uneasy. In a nutshell, my problem is that, when firms introduce these windows, many other fundamental changes are taking place in the work environment. These other, coincident changes may also affect departure rates so that the experiment may be contaminated.

Windows are typically used, as appears to be the case here, to achieve a large-scale downsizing of the work force. My concern stems from my work with another large Fortune 500 firm which also introduced pension plans. I am familiar with this firm because they were involved in litigation about their pension plan, including its window provisions.

In my case, when the window plan was introduced, the very viability of the firm was at stake. In fact, the smart bet, as confirmed by its low single-digit stock price, was that the firm would not make it. The prospect for the employees of that firm were dim indeed. They were understandably worried about whether they would even have their current salaries or jobs in a few years. There were allegations that employees were told that they would be fired if they did not participate in the pension plan. Most instances when firms use window plans are not so severe as this one, but the more general principle applies. When window plans are used, firms are trying to seriously restructure the size and composition of their work force. They may be dropping entire lines in which older workers' skills are specialized, and they have little incentive to invest in new skills for these older employees. At best, prospects for salary increases and promotions for these workers become much worse than they were. If this is a reasonably accurate characterization of the time period when window plans are used, it may be problematic to use the window period as a test of the parameters of the pension plan.

An alternative way of providing out-of-sample tests is available, which at the very least is complementary to the use of windows. Lumsdaine et al. have previously estimated the parameters of this model with data from another private-sector firm. Using the parameters estimated in the previous firm, the model can be used to predict retirement behavior in this new firm. One objection is that the worker or the nature of the firm are too different to provide a meaningful test. But at some level that objection has to fail. If firms are so heterogeneous that they have little in common, we may not be learning much from data on individual firms.

In future work, I hope that the emphasis of the authors' research shifts to looking at the second two goals of the paper—predicting women's retirement rates and incorporating other components of the fringe package into their model, particularly those involving health benefits. Almost all existing research examines male retirement decisions. The eye-catching result here is that women's hazard rates of leaving the firm are little different from those of men. The reason this seems surprising at first blush is that it runs against the grain of most of the labor supply literature. Female labor-force participation begins to decline at a younger age and more rapidly than men's. In addition, female labor supply appears to be sensitive to male characteristics such as husband's wages. In contrast, Lumsdaine et al. are able to predict similar female exit rates without knowing anything about husbands, including whether they have retired.

The patterns of these female departure rates are fascinating, and Lumsdaine et al. are on the brink of an important contribution. The differences with the age pattern of female labor-force participation rates may be reconciled, in part, because their data actually measure departure rates from the firm. Any departures from the firm for women are more likely to be out of the labor force than to another job. In a similar vein, the departure rates in their paper are conditional on years of service. Most of the male-female difference may come in the distribution of years of service rather than in these conditional departure rates.

But even this reconciliation fails to address the question of why they are so successful in predicting female departure rates without any knowledge about the current situation of the husbands of these women. Since the husbands are typically older, does this mean that the family waits until both members retire optimally (based on their own financial incentives) before entering retirement life (including migrating to more amenable areas)? Current sex discrimination laws have had the bizarre interpretation of not permitting the longer life expectancies of women to affect the yearly flows of benefits. What is the impact of that sex bias in favor of women in their retirement behavior? The Lumsdaine et al. paper is breaking new ground in which we will be able for the first time to explore these fascinating questions.