JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **19**: 795–807 (2004)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jae.798

A DYNAMIC PROGRAMMING APPROACH TO MODEL THE RETIREMENT BEHAVIOUR OF BLUE-COLLAR WORKERS IN SWEDEN

ANDERS KARLSTROM,[a] MARTEN PALME[b]* AND INGEMAR SVENSSON[c]

[a] *Systems Analysis and Economics, Royal Institute of Technology, Stockholm, Sweden*
[b] *Department of Economics, Stockholm University, Sweden*
[c] *National Social Insurance Board, Stockholm, Sweden*

SUMMARY

This paper presents an empirical analysis of how Sweden's public old age pension system affects the retirement decision. We focus on male blue-collar workers whose dominant income source as retired comes from the public old age pension system. We develop a dynamic programming model using the rules for the public pension system. In addition to the effects of economic incentives through the pension systems the DP model also measures the effect of the mandatory retirement age of 65—which applies to most parts of Sweden's labour market—on retirement behaviour. The estimated model fits within-sample retirement patterns remarkably well. A simulation of a hypothetical reform, where all retirement incentives in the pension schemes are delayed by three years, shows that economic incentives affect retirement behaviour.

1. INTRODUCTION AND MOTIVATION

In this paper we develop a dynamic programming model for the retirement decision in Sweden. Earlier empirical work on retirement in Swedish data has used static structural models, or the so-called option value model, see e.g. Karlstrom *et al.* (2003). Theoretically more sound dynamic programming or life-cycle models, that simultaneously model consumption and labour supply, have been proposed by, for instance, Rust *et al.* (2000). One purpose of this paper is to demonstrate that a simple dynamic programming model can explain the retirement pattern of blue-collar male workers in Sweden, even with very few estimated parameters. Our model is similar to the one proposed by Rust *et al.* (2000). Similar work also includes Heyma (2001), who studies retirement behaviour in the Netherlands, and Knaus (2002), who calibrates a life-cycle model using German data.

The model is estimated on panel data for a random sample of 4638 male blue-collar workers born between 1927 and 1940. The retirement behaviour is observed between 1983 and 1997 and the models are estimated under the rules for the current national old age pension system. The model is evaluated by means of a simulation experiment. We compare the retirement behaviour under the current system with the out-of-sample predictions under a hypothetical reform which delays the eligibility and normal retirement ages as well as the actuarial adjustment for early or delayed withdrawal by three years.

Although the most common exit from the labour force is through the disability insurance programme (see Palme and Svensson, 2004), we have excluded those who use this programme

* Correspondence to: Marten Palme, Department of Economics, Stockholm University, SE-10351 Stockholm, Sweden. E-mail: marten.palme@ne.su.se

Received 25 July 2003
Revised 9 September 2003

from our sample. There are two reasons for doing that. First, it simplifies the specification and estimation of the dynamic programming model since we do not need to consider the rules for the disability insurance. We also restrict the sample to blue-collar workers, since it limits the occupational pension plans that need to be considered in the calculations. Second, the empirical focus will be on the effect on those who retire through the old age pension system. We also limit the analysis to those who retire after the early retirement age of 60, which is the case for a large majority of those who finance their retirement through the old age pension system.

We estimate two versions of the model: one version which contains a flexible specification for changes in preference towards retirement by age and one restricted version where the preferences are constant over different ages. The models fit observed retirement patterns well. In particular, the restricted model, considering that it only contains three estimated parameters, predicts the retirement behaviour remarkably well. The results from the simulation of the hypothetical reform show that the model predicts that economic incentives do affect retirement behaviour. This result applies also to the predictions from the unrestricted model. However, the large discrepancy between the out-of-sample predictions from the restricted and the unrestricted model indicates that labour market institutions in a broad sense may be important for the retirement decision and/or difficulties in identifying the effects of economic incentives from labour market institutions.

The rest of the paper is organized as follows. The public old age pension system is briefly described in Section 2 and the data is presented in Section 3. The dynamic programming model is developed in Section 4. The results from the estimation and evaluation of the model are presented in Section 5, and the results from the simulations of the hypothetical reform are reported in Section 6. Section 7 concludes.

2. SWEDEN'S PUBLIC OLD AGE PENSION SYSTEM

Blue-collar workers in Sweden are covered by two national pension programmes, the basic and supplementary pension (ATP). Both these programmes have recently been subject to major reforms. Below we will, however, give a brief description of the pre-reform systems which apply to the workers in our data set.

All Swedish citizens are entitled to the *basic pension*, which is unrelated to previous earnings. As all social insurance programmes in Sweden, the basic pension is indexed by the basic amount (BA), which follows the CPI very closely. The level of the basic pension is 96% of a BA for a single pensioner and 78.5% for married. In the year 2003 the level of one BA was 38,600 SEK.[1]

Individuals with no, or low, ATP are entitled to a special supplement. This supplement is independent of marital status and has grown from 42% of the BA in 1983 to 55.5% as of 1993. The special supplement is reduced on a one-to-one basis against the supplementary pension. Thus a single old age pensioner with only a basic pension and a special supplement receives 151.5% of the BA.

The normal retirement age for this pension is 65, but it can be claimed from age 60 with a permanent actuarial reduction of 0.5% for each month of early withdrawal. If the pension is claimed beginning after age 65, the level is permanently increased by 0.7% for each month of delayed withdrawal up to age 70.

[1] In 2003 the exchange rate was $1 ≈ 8.1 SEK.

The *supplementary pension* (ATP) is related to the worker's earnings history and the benefit is calculated in three steps. The first step involves determining pension-rights income for each year from the age of 16. Pension-rights income is calculated on the basis of income from labour reported in an individual's annual tax return and is the share of the income exceeding 1 BA and below the social security ceiling at 7.5 BA.[2] It is set to zero if annual income from labour does not exceed 1 BA. Besides earnings and income from self-employment, transfer payments from social insurance, such as income from sickness and unemployment insurance, the parental cash benefit and the partial retirement pension are included in pension-rights income. Three years of pension-rights income greater than zero between the ages of 16 and 64 are required to receive an old age pension from the ATP scheme.

In the second step, average pension points are calculated by dividing pension-rights income by the corresponding year's BA to obtain the pension points for each year. Thus, due to the social security ceiling at 7.5 BA, the maximum number of pension points an individual may receive in any given year is 6.5. Average pension points comprise the average of an individual's 15 best years of earnings.

The final step is to calculate an individual's ATP benefit (Y_i) by applying the formula

$$Y_i = 0.6 \cdot \mathrm{AP}_i \cdot \min\left(\frac{N_i}{30}, 1\right) \cdot \mathrm{BA}$$

where AP_i is individual average pension points, BA is the basic amount, and N_i is the number of years the individual has recorded covered pension right income greater than zero. The average of pension points is calculated as the average of annual earnings below the social security ceiling of 7.5 BA of the worker's 15 best years. The normal retirement age for the supplementary pension is 65. The actuarial adjustment for early and delayed withdrawal is the same as for the basic pension.[3]

3. DATA[5]

We use the Longitudinal Individual Data (LINDA) panel. LINDA is a pure register sample. It contains data from Statistic Sweden's Income and Wealth register, which is a register containing data from the income tax returns for the entire Swedish population; the Population Census, which is data primarily on occupation and housing conditions from mailed questionnaires made every five years to the entire population; and the National Social Insurance Board registers, which contain data on contributions to the public pension schemes.

The sample size of LINDA is about 300,000 individuals. Detailed income components are available from 1983. Data on earnings below the social security ceiling are obtained back to 1960 from the pension register.

We have selected men born between 1927 and 1940. We have excluded individuals younger than age 50. Since, e.g., the youngest cohort, born in 1940, are just 43 years old in 1983, we exclude the first seven observations for each individual from this cohort. Finally, we restrict the sample to blue-collar workers. The final sample size in the cross-section is 4638 individuals and

[2] The proportional payroll tax used to finance the ATP scheme is also paid on the share of income exceeding 7.5 BA.
[3] Both the basic pension and the supplementary pension contain a survivor's benefit that we do not take into account in the estimation. For a description of the survivor's benefit, see Palme and Svensson (2004).
[4] For a more detailed description of the data, see Palme and Svensson (2004).

since the panel structure of the data set allows us to observe each individual more than once, the total number of observation used in the estimation is 46,734.

4. THE DYNAMIC PROGRAMMING MODEL

4.1. Specification of the Dynamic Programming Model

The individual's preferences are specified by the single period utility functions $U_t(\mathrm{s}_t, d_t, \theta_u)$, where s_t is a vector of state variables at age t, and θ_u are parameters that are to be estimated, but known to the individual; d_t is the individual's control variables. In our application, $d_t = 1$ if the individual retires at age t and zero if the individual continues to work. The individual will make this choice each year starting from age 50 and can remain in the workforce until age 70. Following the framework of Rust (1987, 1994) and Rust and Phelan (1997), we partition the state variables into $(\mathrm{x}, \varepsilon)$, where x is a vector of state variables observed by both the econometrician and the individual, and ε is a vector observed only by the individual. We also impose the assumption that x and ε are conditionally independent, see Rust (1994).

Assuming an additive random utility model, the single period utility functions are given by $U_t = u_t(\mathrm{x}_t, d_t, \theta_u) + \varepsilon_t\ (d_t)$. In the following we will assume that the unobserved components $\varepsilon_t(d_t)$ are i.i.d. Gumbel distributed. The individual chooses the action that maximizes the lifetime utility. The value function is given by

$$V_t(\mathrm{x}_t, \varepsilon_t, \theta_u) = \max_{d \in D(\mathrm{x}_t)} [v_t(\mathrm{x}_t, d_t, \theta_u) + \varepsilon_t(d_t)] \tag{1}$$

where $D(\mathrm{x}_t)$ denotes the choice set available to the individual in state x_t, and v_t is the *expected value function*, defined recursively by

$$\begin{aligned} v_t(\mathrm{x}_t, d_t, \theta) &= u_t(\mathrm{x}_t, d_t, \theta_u) \\ &+ z_{t+1}\beta \int \left\{ \log \sum_{d_{t+1} \in D(\mathrm{x}_{t+1})} \exp[v_{t+1}(\mathrm{x}_{t+1}, d_{t+1}, \theta)] p_t(d\mathrm{x}_{t+1}|\mathrm{x}_t, d_t, \theta_p) \right\} \end{aligned} \tag{2}$$

where β is the discount factor and z_{t+1} is the survival probability from age t to $t+1$. The person dies with probability 1 at age 102. The finite horizon DP model is therefore easily solved with backward recursion, and we assume $u_t = 0$ at $t = 102$.

Thus, the individual's behaviour depends on the discount factor β, preferences u_t and beliefs p_t, which are specified by parameters θ partitioned into (θ_u, θ_p).

Integrating out the unobserved state variables, we arrive at *conditional choice probabilities*

$$P_t(d_t|\mathrm{x}_t, \theta) = \frac{\exp(v_t(\mathrm{x}_t, d_t, \theta))}{\sum_{d' \in D(\mathrm{x}_t)} \exp(v_t(\mathrm{x}_t, d', \theta))} \tag{3}$$

The conditional choice probability is used to estimate the model and we use maximum likelihood (ML) as estimation method. As noted in the Introduction, similar models have until recently been calibrated, rather than estimated. However, recently French (2001) and, for instance, Heyma (2001) have provided econometric estimates of (full) life-cycle models. French uses GMM as estimation method, which will yield statistically efficient and asymptotically identical estimates as ML, when using optimal weighted moments.[5]

We will use (partial information) maximum likelihood to estimate the model. Given our panel decisions d_t^i for each individual i at age $t = T_i^1, T_i^1 + 1, \ldots, T_i^2$, and observed state variables x_t, we form the likelihood function

$$L(\beta, \theta_u, \theta_p) = \prod_{i=1}^{N} \prod_{t=T_i^1}^{T_i^2} P_t(d_t^i | x_t^i, \theta_u) p_t(\mathrm{x}_t^i | \mathrm{x}_{t-1}^i, d_{t-1}^i, \theta_p) \tag{4}$$

The beliefs are specified by $p_{t+1}(\mathrm{x}_{t+1}|\mathrm{x}_t, \theta_p)$. To reduce the computational burden, we will follow a two-stage estimation procedure (see Rust, 1987 or Rust and Phelan, 1997). In short, we will estimate θ_p by defining a first-stage partial likelihood using only p_t, and then use these parameters to estimate the remaining parameters in θ_u. Although this strategy does not yield fully efficient estimates, Rust and Phelan (1997) indicate that efficiency loss may be rather small. However, we have not examined the potential loss of efficiency in this application. Finally, since the log-likelihood function is not necessarily convex for all parameter values, we have used simulated annealing to attain the estimates.

Having specified the general structure of the dynamic programming model, we now turn to the specification of preferences and beliefs.

4.2. State Variables and Beliefs

We have five different state variables $\mathrm{x}_t = (t, w_t, ap_t, ra_t, m)$, where t is the person's age, w_t is wage earnings, ap_t is the average pension points, ra_t is the retirement age and m is the martial status.[6]

As indicated above, we will use a two-stage estimation approach. The updating rules of age and retirement age are trivial, given decisions. We will also simplify our model by assuming that marital status is known in advance.

Following Rust *et al.* (2000) we assume that wage earnings follow a stochastic process, and use a deterministic updating rule for average pension points. The beliefs are decomposed, and we write the transition probabilities as a product $\tilde{p}_t(ap_{t+1}, w_{t+1}|\mathrm{x}_t) = p_1(ap_{t+1}|ap_t, w_t, t)p_2(w_{t+1}|w_t, t)$. We will consider these factors in turn.

First, the average pension points, ap_t, is a key variable in the model. It serves as a 'sufficient statistic' for the income history, allowing us to reduce the dimension of the problem to be solved. We define ap_t as $m \cdot T_{\max}$, where m is the average point according to ATP and $T_{\max}$ is the number of years with ATP points, but constrained to a maximum of 30 since, as we explained in Section 2.1, contributions beyond 30 years do not affect the ATP benefit. In reality, ap_t is, as follows from the description in Section 2.1, inherently a deterministic function of past earnings history. However, as in Rust *et al.* (2000) and Knaus (2002), it turns out that the average pension points can be closely modelled by an autoregressive process, i.e.

$$\log\,(ap_{t+1}) = \gamma_1 + \gamma_2 \log\,(ap_t) + \gamma_3 t + \gamma_4 t^2 + \eta_t \tag{5}$$

[5] Our model is more simple than French's model in the sense that we only have one control variable and estimation by maximum likelihood is tractable.

[6] Note that retirement status r_t is implicitly given by ra_t and t.

Table I. Estimation results for wages and average pension points regression equations

Parameter	Point estimate	s.e.
Average pension points		
γ_0	-0.2446	0.0303
γ_1	0.9393	0.0005
γ_2	0.0230	0.0004
γ_3	0.0034	0.0010
γ_4	-4.45×10^{-5}	8.95×10^{-6}
R^2	0.9972	
$\hat{\sigma}_\eta^2 = 0.0003$		
Wages		
α_0	1.0386	0.2381
α_1	0.8876	0.0026
α_2	0.0117	0.0082
α_3	-0.0001	7.09×10^{-5}
R^2	0.8030	
$\hat{\sigma}_\zeta^2 = 0.0429$		

Number of observations = 27, 704.

where η_t is i.i.d. $N(0, \sigma_\eta^2)$. The estimated regression has a very high R^2, and the estimated standard error is very small, see Table I. This indicates that, with high accuracy, we have a deterministic, Markovian updating formula for the average pension points

$$ap_{t+1} = \exp\left\{\hat{\gamma}_1 + \hat{\gamma}_2 \log(ap_t) + \hat{\gamma}_3 t + \hat{\gamma}_4 t^2 + \frac{\hat{\sigma}_\eta^2}{2}\right\} \tag{6}$$

where $\hat{\sigma}_\eta^2$ is the estimated variance of η_t and $\hat{\gamma}_1, \ldots, \hat{\gamma}_4$ are the estimated parameters. The value of ap_t is frozen if the individual enters retirement, i.e. $ap_s = ap_t, \ \forall s > t$.

Second, turning to wage earnings, future earnings are not known to the individual in advance. Hence, it is modelled as a stochastic process. We model wage earnings by a simple autoregressive process

$$\log(w_{t+1}) = \alpha_1 + \alpha_2 \log(w_t) + \alpha_3 t + \alpha_4 t^2 + \zeta_t \tag{7}$$

where t denotes age and ζ_t are i.i.d. $N(0, \sigma_\zeta^2)$. The quadratic specification allows for an age income profile.

Table I shows the results from the estimation of this model. R^2 is fairly high also for this regression, although not as high as for the regression of the average pension points, equation (6). The standard error of the regression of the wage equation is also much larger than the regression of ap_t. Therefore, it is indeed not appropriate to model wage as a deterministic process.

The wage equation is estimated conditional on working the next year. At age 65, many individuals in our sample do not have the option of remaining in the work force, i.e., they can only choose retirement. Whether or not an individual will have the choice of remaining in the work force available at age 65 or not is another important source of uncertainty for the individual. We will model this by setting the wage to zero,[7] if the individual does not have this option available.

[7] Note that the single period utility approaches $-\infty$ as consumption approaches zero. Hence, no one will choose working if the wage is zero. For numerical reasons, we set the wage to a very small number.

Thus, the transition probabilities of the wage variable are adjusted at age 65, by assigning a probability p_{65} of receiving a positive wage, if continuing working. This probability is assumed to be the same for all individuals. If the individual does in fact continue to work beyond age 65 (and consequently had the option available at the age of 65), we will assume that the individual will continue to have the option available with probability one for all ages ≥ 65. That is, for the individual, the uncertainty is dissolved at 65. We will estimate the mixing probability p_{65} simultaneously with the preference parameters, assuming rational expectations.

4.3. Preferences

Workers are assumed to receive utility from consumption and leisure. *A priori*, we expect the utility function to be increasing and concave in consumption. The deterministic component of the single period utility function is given by

$$u_t(\mathrm{x}, d_t, \theta_u) = \tilde{u}_t(c_t, f_t) = \alpha \log(c_t) + \phi_t(\theta) \log(f_t) \tag{8}$$

where f_t is the amount of leisure time. It is assumed to be $f_W = 1$ if the individual is not working (i.e. is retired, $r_t = 1$) and $f_R = 1 - 0.45$ if the individual is working ($r_t = 0$). c_t is consumption, which is assumed to equal the retirement benefits b_t if the worker is retired and labour earnings w_t otherwise. The only source of intertemporal substitution is thus through the public pension system.

As we explained in Section 2, the old age pension benefit is determined by the three state variables of retirement age, average pension points and, through the basic pension, marital status, i.e., $b_t = b(ra_t, ap_t, m_t)$. The benefit is known by the individual. Since retirement is an absorbing state, lifetime utility after retirement is a deterministic, discounted sum

$$v_t(\mathrm{x}_t, d_t = 1, \theta) = \sum_{s=t}^{102} \beta^{s-t} \prod_{k=1}^{s-t} s_{t+k} \tilde{u}_t(b_t, f_R) \tag{9}$$

Hence, calculating the value function given retirement is easy.

Note that the specification of the single period utility function (8) implies that preferences are not dependent on marital status *per se*, but only through monetary incentives from the rules of how the retirement benefit is calculated.

$\phi_t(\theta)$ is a function given by several parameters θ that allow for an age-varying profile of leisure preferences. One motivation for allowing a flexible profile for leisure by age is that health status may decline with age. This may be an important feature since we have no separate measure of health status in our data. In calibrated models for the USA (Rust *et al.*, 2000) and Germany (Knaus, 2002), it has been found appropriate to use a preference profile for leisure that is rather fast increasing in age. We follow Knaus (2002) and let the parameter capturing the preferences for leisure have a conventional logit specification, i.e.

$$\phi_t(\theta) = \exp(\theta_1) + \exp(\theta_2) \cdot \frac{\exp\left(\dfrac{t - \theta_3}{\theta_4}\right)}{1 + \exp\left(\dfrac{t - \theta_3}{\theta_4}\right)} \tag{10}$$

This specification allows for a flexible specification to capture changing leisure preferences by age, while ensuring that the model's ability to fit our data is not a result of age-specific dummy variables.

We have not been able to get an estimate of the discount factor β. It is therefore set to 0.97. The preference parameters to be estimated are therefore $\theta_u = (\alpha, \theta_1, \ldots, \theta_4)$.

4.4. Timing of Events

As we described above, the only decision variable for the individual is whether to retire or to continue to work full-time. We assume that this decision is taken at the beginning of each year. That is, at the beginning of the year, the individual will contemplate whether to retire or work yet another year. The information available for the individual at that time is the wage earnings the previous year, and the average pension points (and of course the age and present labour status). The individual is not certain about future earnings, and consequently also uncertain about future average pension points and pension benefits. Given the available information, the individual knows the distribution of next year's wage earnings, given by equation (7). From the distribution of next year's wage earnings, a probability distribution of next year's average pension points can also be inferred (given by equation (6)). Only the current year's idiosyncratic taste variables are known to the individual, but not future years'.

With all the available information, the individual can calculate the expected utilities for the two decisions (retire or not retire), and choose the one with the highest expected utility. However, since we as researchers cannot observe the idiosyncratic taste variables, we can only calculate the probability that an individual will choose to retire.

5. ESTIMATION

5.1. Implementation

In the numerical implementation of the model, we have used discretization of the continuous earnings from labour variable. In the reported results we have used 400 discretization points. The estimation results were, however, robust with much smaller number of points. Although an efficient way to solve the one-dimensional integrals is using quadrature rules, such as e.g. Gauss–Legendre used in Knaus (2002), we have chosen to use simulation instead. The reason for doing so is that we are working on extending the model to include more state variables. For instance, in an ongoing project we model health status and include the disability insurance (DI) system. For such models it is necessary to use simulation techniques in the estimation and therefore it is the method of choice also in this paper. Consequently, we have also used a random grid for the discretization points, and used linear interpolation to calculate the value functions.[8] The number of simulation draws was set such that both the estimated parameters and the log-likelihood value was stable to at least six digits.[9]

5.2. Results

We estimate two different versions of the model. The first version is a full model allowing for a flexible age-differentiated preference for leisure, where we estimate all parameters in equation (10).

[8] A randomized grid is more efficient than a uniform grid to solve multiple integrals. For a similar study in a dynamic programming context, see Rust (1997).

[9] The resulting simulated maximum likelihood simulator is not unbiased with finite number of draws, due to 'Jensen's inequality'.

This model will, in the following, be labelled M1. In the second version, we impose the restriction of not allowing for changing preferences for leisure by age. That is, we estimate only θ_1 in (10). This model will be labelled M2.

Table II shows the estimation results for both models. For the more flexible M1 model, all estimated parameters are significantly different from zero and have the expected signs. Note that the estimated function for preferences for leisure is not very smooth. In fact, the estimated preference profile is essentially a dummy variable for ages >65.

It is, of course, possible that the estimated profile reflects a real change in preferences starting at age 65. However, it seems much more likely that this age profile reflects institutional rules on the Swedish labour market and social norms on retirement at the normal retirement age of 65. As explained in Section 2, many workers automatically lose their job at age 65 due to rules on mandatory retirement age. This means that the estimated dummy variable for ages above age 65 is likely to reflect the institutional constraints, rather than changes in preferences by age. As we will come back to in the discussion of the simulation results in Section 5, it is for this reason useful to estimate a model without allowing for time-differentiated preferences. The estimation shows, as expected, that the goodness of fit in terms of log-likelihood value is not as good for the constrained, M2 model as the more flexible model, see Table II.

Hence, allowing for a flexible preference structure over time clearly leads to better goodness of fit, but we also run the risk of over-parameterization. One way to examine the trade-off between lack of fit and over-parameterization is to compare how well the models predict the actual retirement behaviour. Figure 1 shows the cumulative distribution functions for the actual labour force participation of the workers in the sample, along with the predictions for the full M1 model and the constrained M2 model, respectively.

Let us first look at the predictions of the M1 model. As expected, at early ages the model underestimates the hazard rate into retirement. In our model the individuals have zero pension benefit by retiring before age 60, which probably does not apply to all individuals in the sample.[10] This is, however, a minor problem since very few individuals in our sample retire at that age. In the interesting age span of ages 60 to 65, the estimated model follows the observed retirement pattern quite closely. The hazard rates are somewhat overestimated at ages 63 and 64. At older

Table II. Results from the maximum likelihood estimation of the parameters in the model

Parameter	M1 model		M2 model	
	Point estimate	s.e.	Point estimate	s.e.
α	0.9074	0.0610	1.2307	0.0590
θ_1	0.8504	0.1870	1.8215	0.1703
θ_2	0.4078	0.0835	—	—
θ_3	65.4	—		
θ_4	0.007	—		
p_{65}	0.9730	0.0040	0.9531	0.0029
Log-likelihood	−3625.3		−3692.3	

Note: $N = 4638$; number of observations: 46,734; $\beta = 0.97$.

[10] Also, we do not model savings. However, as argued in Section 2, this is probably not an important shortcoming, at least not for explaining early retirement in this sample.

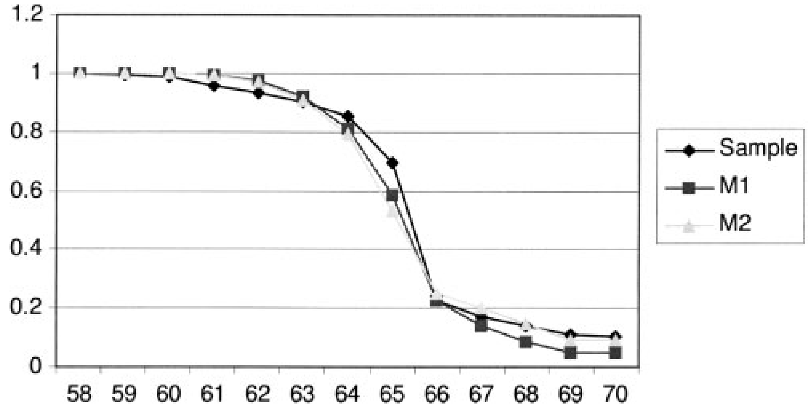

Figure 1. Cumulative distribution of labour force participation. M1 is estimated model with leisure preferences varying by age, and M2 is estimated model where preferences are not age-dependent

ages, the model underestimates the probability of remaining in the work force. Again, this may be explained by unobserved variables and the fact that our maximum likelihood estimator attaches little weight to accurately predicting the retirement behaviour of this age group, since very few observations remain in the sample after age 65. In fact, more than 80% are retired at age 66.

Figure 1 shows that the M2 model, which in addition to the constant only uses measures of economic incentives, also gives a fairly good fit to the observed retirement pattern. At early ages, the models M1 and M2 are almost identical, while the constrained model M2 overestimates the hazard rates somewhat more at age 63 and 64, while it somewhat underestimates the peak at age 65. However, the more parsimonious specification, with only variables measuring economic incentives, gives a satisfactory fit to the observed retirement pattern.

To summarize the estimation results, we believe that the estimated models can satisfactorily replicate the observed pattern of retirement. We also think that the more parsimonious specified model, where we only include economic incentives, gives a good fit to the observed pattern. In the next section, we will therefore simulate a policy reform and assess the effects on labour supply using our two estimated models.

6. SIMULATION RESULTS

To further evaluate the empirical implication of the estimated models, we simulate the effect of a hypothetical reform on the retirement behaviour of the workers in the sample. The content of the reform is simply to delay the eligibility ages and the actuarial adjustment of the public pension system by three years. This means that the public pension can be claimed starting from age 63 and that the normal retirement age is changed to 68.

Retirement behaviour is affected by economic incentives, labour market institutions in a broad sense and changes in preferences towards retirement by age. We divide economic incentives into incentives inherent in the old age pension scheme and incentives which follow from rules on mandatory retirement age implying that workers lose their jobs at a particular age. The aim of our overall simulation strategy is to separate out the effect of the two different kinds of economic incentives that we consider on retirement behaviour.

We do three different simulations. In the first one, S1a, we use the full M1 model. We maintain the mandatory retirement at age 65. The only change compared to the baseline is that we change economic incentives *through the old age pension system only*. This implies that the effects of the incentives working through the old age pension system are isolated.

In the second simulation, S1b, we shift the rule of mandatory retirement to apply at age 68 (instead of 65). However, we still maintain the estimated flexible preference structure of model M1, where individuals have more preference for leisure from age 65 onward. The interpretation of this simulation is that we change economic incentives through both pension benefits and rules of mandatory retirement. Thus, this simulation gives the full effect of economic incentives on retirement behaviour.

As we explained in the previous section, the M1 model including time preferences is over-parameterized in the sense that it is likely that unmeasured economic incentives are captured by the time preference specification. In this sense, it is thus likely that the behavioural effect of the reform is underestimated using this simulation and modelling strategy. The S1 simulations can therefore be interpreted as a 'lower bound' prediction on retirement behaviour of the hypothetical reform.

In the third simulation, S2, we use the restricted M2 model. As in simulation S1a, we change pension benefits and delay the rule of mandatory retirement to age 68. This simulation does not include the effects of age-varying preferences. Since such a specification is likely to reflect the effects of institutions, such as the effect of social norms of retirement at age 65, as well as genuine changes in preference in retirement by age, which are not changed in the reform, it may be argued that the S2 simulation overpredicts the behavioural response to the reform. This prediction can therefore serve as an 'upper bound' to the behavioural response to the reform.

Figure 2 shows the cumulative distribution functions for the predicted labour force participation rates under the current system from model M1 and M2 respectively, along with the outcome from the three simulations. The S1a and S1b simulations should be compared to the M1 prediction and the S2 simulation to the M2 prediction. Table III shows the expected retirement age for each prediction and simulation, along with retirement hazard rates.

Several interesting results emerged from this exercise. First of all, the S1a simulation shows that economic incentives operating through the old age pension scheme also have an effect on retirement behaviour in this 'lower bound' simulation strategy. The second form of economic

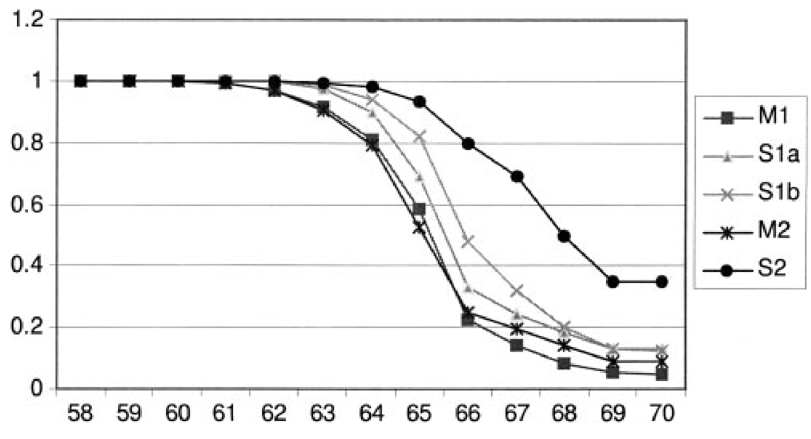

Figure 2. Cumulative distribution of labour force participation for the simulated hypothetical policy reforms

Table III. Hazard rates and expected retirement ages

Age	Sample	M1	M2	S1a	S1b	S2
60	0.03	0.01	0.00	0.00	0.00	0.00
61	0.02	0.01	0.00	0.00	0.00	0.00
62	0.03	0.05	0.06	0.02	0.00	0.00
63	0.04	0.12	0.12	0.07	0.00	0.01
64	0.18	0.27	0.33	0.22	0.01	0.04
65	0.67	0.61	0.53	0.52	0.04	0.14
66	0.23	0.37	0.19	0.25	0.11	0.13
67	0.18	0.39	0.26	0.24	0.34	0.28
68	0.20	0.40	0.37	0.28	0.15	0.29
69	0.08	0.02	0.00	0.01	0.12	0.00
70	0.00	0.00	0.00	0.02	0.07	0.01
Expected retirement age	65.07	64.80	64.95	65.58	66.00	67.59

incentive inherent in the mandatory retirement age has an effect since shifting the mandatory retirement age, again in the ‘lower bound’ simulation, further shifts out the cumulative distribution function for labour force participation in Figure 2.

The large discrepancy between the outcomes of the S1b and S2 simulations shows that there is a large element of uncertainty in the prediction of the effect on retirement behaviour. There is an important effect of the specification of changes in preferences towards retirement by age, which in turn is likely to reflect social norms and labour market institutions in a broad sense affecting retirement behaviour. The importance of this effect is not revealed in the within-sample predictions.

7. DISCUSSION AND CONCLUSIONS

In this study we develop two dynamic programming models to study the retirement behaviour of blue-collar workers in Sweden: one model which allows for a flexible specification of changes in preferences towards retirement by age and one restricted version. Both models, but in particular the restricted one which only contains three parameters, perform remarkably well in the within-sample predictions of labour force participation rates. There are probably two main reasons for this outcome. The first is that we study a fairly homogenous group exposed to the same labour market institutions and economic incentives, both through the pension scheme and the rules on mandatory retirement age. The second reason is that we are able to fairly accurately predict the income of the workers in the sample as retired, since this group relies to a large extent on the public pension system for their old age income security.

In understanding the relation between economic incentives and labour force participation rates for older workers on the Swedish labour market, the present study has several limitations which call for further research. The most important limitation is that we exclude all workers who retire through the DI system. For blue-collar workers in Sweden more than 50% finance their exit from the labour force through the DI system. A related limitation is that we also disregard the influence of other parts of Sweden’s income security system—such as the unemployment and sickness insurance programmes—which have become increasingly important as pathways out of the labour force (see Palme and Svensson, 2004).

The results from the simulations—which are out-of-sample predictions—suggest that economic incentives affect retirement behaviour, both through the public pension system and through rules on mandatory retirement age. However, the large discrepancy between the simulations using the model which allows for changes in preferences towards retirement by age and those of the restricted model shows that labour market institutions in a broad sense, such as social norms on retirement at age 65, may be important and/or problems in identification of the effect of rules on mandatory retirement age. This conclusion is supported by the fact that the economic incentives inherent in the public old age pension scheme are likely to be measured with small errors.

ACKNOWLEDGEMENTS

We are grateful to the editor John Rust for suggestions and encouragement, as well as John Dagsvik for comments. Financial support for this research is provided by the National Social Insurance Board in Sweden.

REFERENCES

French E. 2001. The effects of health, wealth, and wages on labor supply and retirement behavior. Manuscript, Federal Reserve Bank of Chicago.

Heyma A. 2001. Lessons from a structural dynamic analysis of retirement behavior in the Netherlands. Manuscript, University of Amsterdam.

Karlstrom A, Palme M, Svensson I. 2003. The timing of retirement and social security reforms: measuring individual welfare changes. Manuscript, Department of Economics, Stockholm University.

Knaus T. 2002. The retirement decision in Germany: a dynamic programming model. Manuscript, Freie Universität Berlin.

Palme M, Svensson I. 2004. Income security programs and retirement in Sweden. In *Social Security and Retirement Around the World: Microestimation*, Gruber J, Wise D (eds). University of Chicago Press: Chicago.

Rust J. 1987. Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. *Econometrica* **55**: 999–1023.

Rust J. 1994. Structural estimation of Markov decision processes. In *Handbook of Econometrics*, Vol. 4, Engle R, McFadden D (eds). North-Holland: Amsterdam.

Rust J. 1997. A comparison of policy iteration methods for solving continuous-state, infinite-horizon Markovian decision problems using random, quasi-random, and deterministic discretizations. Working paper, Yale University.

Rust J, Phelan C. 1997. How social security and medicare affect retirement behavior in a world of incomplete markets. *Econometrica* **65**: 781–831.

Rust J, Buchinski M, Benitez-Silva H. 2000. Dynamic structural models of retirement and disability. Manuscript, Department of Economics, Yale University.

