Annuitization and Retirement Timing Decisions

Zhen Shi†
Tilburg University

- Job Market Paper -

November 12, 2008

Abstract

This paper analyzes retirement timing decisions of DC pension plan members. In a first step, a model of optimal retirement timing decisions is proposed which incorporates the optimal time to annuitize the DC pension wealth. An individual obtains utility from leisure, labor income before retirement and pension benefits after retirement. These benefits include the income from an annuity which is bought at an optimal time. The optimal annuitization time depends on current and future expected financial market performance. Based on the model, a forward looking retirement likelihood measure is derived which describes the probability that an individual retires within the next few years. In a second step, the retirement likelihood measure is used to predict transitions into retirement which took place between the first and second wave of the English Longitudinal Study of Ageing (ELSA). It turns out that the retirement likelihood measure has strong predictive power for actual retirement timing decisions. More precisely, for various groups of individuals the correlation between model predictions and actual retirement outcomes reaches 94% and the root mean square error can be as low as 5%.

Keywords: Retirement, Annuitization, Optimal Stopping Time
JEL Classification Codes: J26, G12

*I thank Frank de Jong, Rik Frehen, Kai Li, Roel Mehlkopf, Bertrand Melenberg, Theo Nijman, Michael Orszag, Eytan Sheshinski, Jonathan Skinner, Antony Webb, Bas Werker and seminar participants at Tilburg University, the FMA European Meeting Doctoral Student Consortium (Prague), a Netspar Workshop in Den Haag and the German Finance Meeting in Muenster for useful comments and discussions.

†CentER, Netspar, and TILEC, Department of Finance, Tilburg University, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Email: z.shi@uvt.nl.
I Introduction

In recent years, there has been a significant shift from Defined Benefit (DB) to Defined Contribution (DC) pension plans in a number of countries, including the U.S., the U.K. and Australia. In the U.S., the number of DB plans has declined sharply in recent years, from 112,208 in 1985 to about 29,600 in 2004 (FDIC 2006). In the U.K., DC plans started widely about two decades ago. At 2002, approximately a third of pension schemes in the U.K. are DC and the trend away from DB funds is expected to accelerate in coming years (Ross and Wills 2002). This shift makes it increasingly interesting to understand determinants of DC pension plan participants’ retirement decisions.

Retirement decisions of individuals with DC plans are jointly influenced by many factors, for example, expected and realized investment returns, the individuals’ risk aversion, the mortality rate, the subjective valuation of leisure, the labor income and its expected growth rate. DC pension plans generally provide benefit in the form of a lump-sum payment. In some countries, for example, the U.S., there are no obligations to annuitize DC wealth, while in others, for example, the U.K., there are obligations to do so. The seminal paper of Yaari (1965) argues that, in the absence of a bequest motive, all retirement wealth of risk-averse individuals should be annuitized. There are two reasons supporting this view. One is that without annuitization there is a risk that the retirees might consume too much so that they will exhaust their retirement resource before they die. The other one is that some retirees might consume too less while they are alive. These individuals could have consumed more to have better life quality. Thus, an important part of annual DC pension income should be annuity income, especially in countries, like the U.K., where there are obligations to annuitize DC wealth.

In reality, individuals with DC plans do not have to annuitize their DC wealth immediately after retirement. The freedom in choosing the annuitization time allows individuals to benefit from better financial market performance after retirement. Therefore it should have a large impact on the retirement timing. Without this freedom, it would be better for individuals to retire when the financial market performance is favorable. However, with this freedom, individuals do not have to wait for favorable financial market performance. It could be optimal for individuals to retire even when the financial market performance is sluggish because they could continue investing their DC wealth in the financial market after retirement and annuitize the DC wealth when the market performs better. The decision to retire is actually a decision to optimally exercise a compound real option. Once the individual retires, he gets the option to annuitize his DC pension wealth. The optimal retirement decision depends on the expected outcome of the annuitization option.
This paper aims to analyze retirement timing decisions of DC pension plan participants, taking into account the optimal annuitization timing decision. To do so, I will first set up a retirement decision model and develop a forward looking retirement likelihood measure from this model. The retirement likelihood measure describes the probability that an individual will retire within the next few years. In the model, the individual obtains utility from leisure, labor income before retirement and pension benefit after retirement. The DC pension benefit is the income from the annuity which is bought at the optimal annuitization timing.

The retirement likelihood measure is then tested with the English Longitudinal Study of Ageing (ELSA) data. The most important reason why I choose U.K. data is that there is an obligation to annuitize pension wealth before age 75. ELSA is a biannual panel survey among those aged 50 and over (and their younger partners) living in private households in England. For all the individuals who are full-time employed in the wave 1 interviews (conducted in 2002-3), the probabilities of retiring by wave 2 interviews (conducted in 2004-5) are evaluated based only on the information available at wave 1 interviews. The model predictions are compared with the actual retirement ratios and the predictions implied by a Probit model where age, gender, education level and DC wealth are explanatory variables used to explain the retirement decisions reported at the wave 2 interviews. The performance of the retirement likelihood measure, in terms of the correlations with the actual retirement ratios and the roots of Mean Square Errors, are comparable to the performance of the Probit regression. This result gives strong support to the option model setup in this paper because the prediction from the option model is out of sample while the prediction from the Probit regression is in-sample. This paper also quantifies the economic benefits of having annuitization timing freedom. The economic benefit is defined as the percentage difference between the certainty equivalent wealth obtained from optimally choosing the annuitization time on the one hand and annuitizing at the retirement time on the other hand. I show that annuitization timing freedom on average leads to a gain of 1.8% for an individual which retires in one year.

This paper is related to the literature focusing on the determinants of retirement decisions. A first line of research in this area has investigated the retirement incentives of DB pension plan participants. The seminal paper by Stock and Wise (1990) presents an option value model to describe the retirement decisions of DB plan participants. Their model is very close in spirit to the stochastic dynamic programming model of Rust (1987). Stock and Wise (1990) apply their model to data from a large company. They find that their model can explain very well the actual retirement ratios in that company. They argue that pension wealth is a significant determinant of the retirement probability.
Samwick (1998) applies the option model to a national-wide dataset. His research confirms and strengthens the results of Stock and Wise (1990). Sundaresan and Zapatero (1997) links the option value to the lifetime marginal productivity schedule which, given their assumption, is increasing at the beginning of the working life and then starts decreasing. They argue that people will retire when the ratio of DB pension benefit and the current wage reaches certain threshold value. This paper extends the option value model of Stock and Wise (1990) to the DC plan participants’ retirement decision.

A second line of research focuses on differences between impacts of DB and DC pension plans on the retirement decision and pension income. Friedberg and Webb (2005) studies the Health and Retirement Survey data and found that workers with DC plans are retiring significantly later compared with the ones with DB scheme. Samwick and Skinner (1998) investigates whether DC plans, compared to DB plans, are adequate in providing for a comfortable retirement pension. Their results show that DC plans can strengthen the financial security of the retirees.

A third line of research looks at the interactions among wealth, investment strategies and the retirement decisions. Gustman and Steinmeier (2002) and Coronado and Perozek (2003) study the effect of a positive shock in household wealth including private savings and savings through DC accounts on household members’ retirement decision making. These two papers investigate the period in the late 1990s when the stock market was booming in the U.S.. Both papers find that the extraordinary high returns in the stock market increase retirement in the United States. Lachance (2003), Choi and Shim (2006), Farhi and Panageas (2007) and Liu and Neis (2002) study the issue of retirement decision and its implication on the investment choice. Choi and Shim (2006) shows that the individual consumes less and invests more in risky assets when he has an option to retire than he should in the absence of such an option. Farhi and Penagear (2007) finds that investing for early retirement tends to increase savings and reduce an agent’s effective relative risk aversion, thus increasing his stock market exposure.

A fourth line of research investigates the interaction between the optimal retirement age and a number of key factors like the availability of annuity products, life expectancy and wages. Sheshinski (2008) provides a comprehensive analysis on these issues.

This paper is also related to the literature on optimal annuitization timing. The literature on this topic is relatively small but growing. Milevsky and Young (2002) develops a normative model of when the individuals should annuitize their wealth. Their model is built on Merton (1971) and solved by standard continuous-time technology. Milevsky and
Young (2007) argues that in the U.S. annuitization prior to age 65-70 was not optimal even in the absence of any bequest motives.

The main contribution of this paper is to incorporate the optimal annuitization timing decision into a normative model explaining the optimal retirement decision making of DC plan participants. There is no doubt that the annuitization timing has large impact on the size of the DC pension benefit. Therefore, rational individuals with DC plans should take this into account while making their retirement decision. Incorporating the optimal annuitization decision making improves the comprehensiveness of a normative model for optimal retirement timing decision. The empirical findings of this paper suggest that in reality at least some individuals recognize the value of the freedom in choosing the annuitization timing and incorporate it into their retirement decision making.

The organization of the paper is as follows. Section 2 describes the retirement decision model. Second 3 discusses the empirical investigation of the model prediction. Section 4 concludes.

II The Retirement Decision Model

The aim of this section is to model the optimal retirement decision of an individual participating in a DC plan, taking into account the optimal annuitization timing. This model will also account for the DB and the state pension plans existing next to the DC pension plan. Currently, we are at time 0 and the individual’s current age is $F$, where $50 \leq F < 75$. He is working full time at time 0. He can retire between time 1, 2, 3,...and time $T$ where time $T$ is the time when this person turns 75 years old. The oldest age the individual could reach is assumed to be $T_{\text{max}}$ and $T_{\text{max}} > T$. His current DC wealth is $W_0$. The individual does not have to annuitize his retirement wealth immediately after retirement unless he retires at time $T$. If he retires before time $T$, he could annuitize his pension wealth between the retirement date, say $t$, and $T$.

Assume that the individual retires at time $t$, where $t$ could be any time between 1 and $T$ and annuitizes at time $t_a$, which could be either at or between time $t$ and $T$. His subsequent pension income, $P(t, t_a)$, consists of annuity income, $A(t, t_a)$, after the individual annuitizes his DC wealth, the amount, $Q(t, t_a)$, withdrawn from his DC wealth before annuitization, the income from current and past DB plans, $CDB(t)$ and $PDB(t)$, 

5
and the state pension, \( SP(t) \), that is,

\[
P(t, t_a)_j = \begin{cases} 
  A(t, t_a) + CDB(t)_j + PDB(t)_j + SP(t)_j, & \text{for } T_{\text{max}} \geq j \geq t_a, \\
  Q(t, t_a) + CDB(t)_j + PDB(t)_j + SP(t)_j, & \text{for } t_a > j \geq 1,
\end{cases}
\]

\[
P(t, t_a)_j = DC(t, t_a) + CDB(t)_j + PDB(t)_j + SP(t)_j, \quad \text{(1)}
\]

where

\[
DC(t, t_a) = \begin{cases} 
  A(t, t_a), & \text{for } T_{\text{max}} \geq j \geq t_a \\
  Q(t, t_a), & \text{for } t_a > j \geq 1.
\end{cases}
\]

For any given pairs of \( t \) and \( t_a \), \( Q(t, t_a) \) is constant over time \((t, t_a)\) and \( A(t, t_a) \) is constant over time \((t_a, T)\). The DB and state pension benefits, \( CDB(t)_j \), \( PDB(t)_j \) and \( SP(t)_j \), are indexed to inflation after retirement. The pension benefits, \( A(t, t_a), Q(t, t_a), CDB(t), PDB(t) \) and \( SP(t) \) will be discussed below in more detail.

**The Financial Market**

In this section, the asset universe available to the DC pension plan member for investment purposes will be introduced. There are one stock index and one bond available in the financial market. The diffusion processes of the short term interest rate and the stock index are as follows,

\[
dr_t = \kappa_r (\bar{r} - r_t) \, dt + \sigma_r dZ_{1t}
\]

\[
dS_t = (r_t + \lambda_s \sigma_s) S_t \, dt + \sigma_s S_t \, dZ_{2t},
\]

where \( \lambda_s \) is the Sharpe Ratio of stock price, \( \sigma_r \) and \( \sigma_s \) are volatilities of short-term interest rate and stock price, and \( \bar{r} \) is the long-term average of the short-term interest rate. The Vasicek process (2) is mean reverting. When the short-term interest rate falls below the long-term average, \( \bar{r} \), the short-term interest rate tends to increase towards \( \bar{r} \) in the future. When the short-term interest rate is above the long-term average, the short-term interest rate tends to fall towards the long-term average in the future. \( \kappa_r \) determines the speed of the such an adjustment process. \( Z_1 \) and \( Z_2 \) are two standard Brownian Motions supported by a probability space \((\Omega, \mathcal{F}, P)\) over the finite time horizon \((0, T)\) with correlation coefficient \( \rho \). All stochastic processes introduced in this paper are assumed to be measurable with respect to the augmented filtration \( \mathcal{F}_t : t \in (0, T) \).

From the Vasicek model, we can get the price of the zero-coupon bond at time \( t \) with time to maturity \( h \)

\[
B_t^{(h)} = e^{-a(h) - b(h)r_t},
\]

6
where

\[ a(h) = \left( \tau - \frac{\lambda_r \sigma_r}{\kappa} - \frac{\sigma_r^2}{2\kappa^2} \right) (h - b(h)) + \frac{\sigma_r^2}{4\kappa} b(h)^2, \]

\[ b(h) = \frac{1}{\kappa} \left( 1 - e^{-\kappa h} \right), \]

and \( \lambda_r \) is the interest rate price of risk. The yield of a zero-coupon bond with time to maturity \( h \), \( Y(h) \), is

\[ Y(h) = \frac{a(h) + b(h) r_t}{h}. \]  

(5)

By Ito’s lemma, the dynamics of any arbitrary bond prices can be described by

\[ dB_t = B_t \left[ \left( r_t + \lambda_r \sigma_{B,t} \right) dt + \sigma_{B,t} dZ_t \right], \]  

(6)

where \( \sigma_{B,t} = \sigma_r D(r,t) \) and \( D(r,t) = -\left( dB_t/dr \right) / B_t \) is the elasticity of the bond price with respect to the short interest rate. The elasticity is referred to as the duration of the interest rate contingent claim. Following Munk, et al (2003), it is assumed that the bond available for the investor has a constant duration \( D > 0 \). This can be thought of as reflecting the duration of the aggregate portfolio of bonds outstanding, or a bond index, where bonds that expire are always substituted with new longer term bonds.

The DC Income

As we have seen before, the DC income, \( DC(t,t_a) \), consists of the amount the individual withdraw before annuitization, \( Q(t,t_a) \), and annuity income after the annuitization, \( A(t,t_a) \). The DC income is jointly affected among other factors by the investment returns, the amount of contributions made to the DC plan and the annuity rates.

Let \( W(t,t_a)_j \) denote the individual’s DC portfolio wealth at time \( j \), \( j \in [t,t_a] \), if the individual retires at time \( t \) and annuitizes at time \( t_a \). Assume that the total amount of contributions paid by the individual and his employer to the DC plan is \( C \) per year. After retirement, the individual will withdraw \( Q(t,t_a) \) per year from his DC wealth before annuitization. A fraction \( \alpha \) of his DC assets is invested in the stock index and \( 1 - \alpha \) in the bond. As in Samwick and Skinner (2003), the investment portfolio will be rebalanced annually to keep the weight of the stock and bond at \( \alpha \) and \( 1 - \alpha \). The optimal annuitization and retirement dates will be described below. For every possible combination of retirement and annuitization dates, that is, \( 0 \leq t \leq T \) and \( t \leq t_a \leq T \),
the individual’s DC wealth can be described as follows

\[
W(t, t_a)_j = \begin{cases} 
\left( \frac{a \times (W(t, t_a)_{j-1} + C)}{S_{j-1}} \right) S_j + \left( \frac{(1-a) \times (W(t, t_a)_{j-1} + C)}{B_{j-1}} \right) B_j, & 1 \leq j \leq t, \\
\left( \frac{a \times (W(t, t_a)_{j-1} - Q(t, t_a))}{S_{j-1}} \right) S_j + \left( \frac{(1-a) \times (W(t, t_a)_{j-1} - Q(t, t_a))}{B_{j-1}} \right) B_j, & t < j \leq t_a.
\end{cases}
\]

(7)

The upper part of equation (7) describes the wealth process before retirement and the lower part describes the wealth process after retirement. Before the individual retires, the total amount of DC wealth available for investing is the sum of the previous DC wealth and the new contribution. After the individual retires but before the individual annuitizes his DC wealth, the total amount of DC wealth available for investing is the difference between the previous DC wealth and the amount withdrawn by the individual.

If the individual retires at time \( t \) and annuitizes his DC wealth at time \( t_a \), the annuity income, \( A(t, t_a) \), which he will receive immediately after annuitization until he dies depends, among others, on the term structure and the amount of DC wealth at the annuitization date, \( t_a \). \( A(t, t_a) \) is determined as follows,

\[
W(t, t_a)_a = A(t, t_a) \left[ 1 + \sum_{j=1}^{T_{\text{max}} - t_a} \left( \frac{1}{1 + r_{t_a}^{(j)}} \prod_{k=1}^{j} M_k \right) \right] (1 + p).
\]

(8)

In eq.(8), \( p \) is a load factor which is greater than or equal to zero, obtaining a measure of the “money’s worth” of the annuity. If the load factor is zero, then the annuity contract is actuarially fair and the “money’s worth” equals one. Empirical evidence by Mitchell et. al. (1999) illustrates that the load factor varies between 8% and 20% depending on different assumptions about discounting and mortality tables. \( M_k \) denote the probability that the individual is alive at time \( k \), conditional on being alive at time \( k - 1 \) and \( M_1 \equiv 1 \). \( r_{t_a}^{(j)} \) is the \( j \)-year interest rate at the time of annuitization.

I assume that the amount, \( Q(t, t_a) \), the individual withdraws after retirement but before annuitization equals the amount of annuity income he could get if he annuitizes immediately after retirement, that is,

\[
Q(t, t_a) = A(t, t).
\]

The DB and state pension incomes are introduced in the following part of this section.
The DB and State Pension Income

If the person retires at time $t$, where $t$ could be any time between 1 and $T$, his income from current and past DB plans, $CDB(t)$ and $PDB(t)$, are determined by, among others, the accrual rate, years of membership and labor income, that is,

$$CDB(t) = \text{acc\_rate} \times n_t \times Y_t$$

$$PDB(t) = \text{acc\_rate} \times n_{\text{past}} \times Y_{\text{last year}} \times \exp(\pi (t - t_{\text{last year}})),$$

where $\text{acc\_rate}$ is the accrual rate, $n_t$ is the number of membership years in the current DB scheme at time $t$, $n_{\text{past}}$ is the number of years in the past DB scheme, $\pi$ is the annual inflation rate, $t_{\text{last year}}$ is the last year in the past DB plan, $Y_{\text{last year}}$ is the individual’s annual gross income during his last year in the past scheme and $Y_t$ is the person’s annual gross income at time $t$. Thus, the DB plan is of a final salary type and the DB income after retirement is indexed to inflation which is required by law in the U.K. (see Blake 2003). This means if the individual retires at time $t$, his income afterwards is,

$$CDB(t)_j = CDB(t) \exp(\pi (j - t)), \text{ for } j = t...T_{\text{max}},$$

$$PDB(t)_j = PDB(t) \exp(\pi (j - t)), \text{ for } j = t...T_{\text{max}}.$$

The state pension is also indexed to inflation, therefore, we have

$$SP(t)_j = SP(t) \exp(\pi (j - t)), \text{ for } j = t...T_{\text{max}}. \quad (11)$$

The Optimal Retirement and Annuitization Timing

The utility function is closely related to Stock and Wise (1990). At time 0, the individual is full time employed. The individual can retire between time 1 and $T$. Looking ahead, he will receive his labor income as long as he keeps working. Once he retires he receives pension income and enjoys the leisure until he dies. At time $t$, $1 \leq t \leq T$, if the individual retires, his utility of retirement, $U_t$, is the sum of the utility from labor income, pension benefit and leisure, that is,

$$U_t = \sum_{s=1}^{t-1} \exp(\beta (t - s)) \frac{Y_s^{1-\gamma}}{1-\gamma} + \sum_{s=t}^{T_{\text{max}}} \exp(\beta (t - s)) \frac{(K P(t, t_{a})^s)_{1-\gamma}}{1-\gamma} \prod_{k=t}^{s} M_k, \quad (12)$$

where $\beta$ stands for the subjective discount factor and the parameter $K$ takes into account the disutility of work. $Y_s$ stands for labor income which is deterministic and $P(t, t_{a})_s$ is
the pension income which is explained in (1). \( \prod_{k=t}^{s} M_k \) is the cumulative survival probability from time \( t \) to \( s \) with \( M_t = 1 \). The first term of (12) is the accumulation of the utility from labor income at time \( t \) and the second term is the sum of the discounted utility from pension and leisure at time \( t \). As in Stock and Wise (1990), the parameter \( K \) has two specifications. In the first specification, \( K \) is a constant. In the second specification, \( K \) is a convex function of current age, \( F \), and \( K = k_0 \left( \frac{F}{k_2} \right)^{k_1} \) where \( k_0, k_1 \) and \( k_2 \) are constants.

For each of the possible retirement stopping times, the DB and state pension income is determined by (9), (10) and (11). But as we have seen before, the DC pension income, \( DC(t, t_a) \), depends not only on when the individual retires but also on when DC wealth is annuitized. This makes the retirement option a compounded real option. Once the individual retires, he obtains the right to exercise his annuitization option. But the retirement decision depends on the expected outcome of the annuitization option. Therefore, in order to find a solution to (12), we first have to find the optimal annuitization timing and thus, the optimal DC pension income, \( P(t, t_a^*) \), for all the possible retirement times from year 1, 2, 3 to year \( T \). After that, we could attempt to solve for the optimal retirement timing for eq.(12).

The retirement timing decision is an example of optimal stopping problems with fixed horizon. The optimal stopping problem describes the problem of choosing a time to stop a certain action based on sequentially observed random variables in order to maximize the expected payoff or utility. A random variable \( \tau \) defined on \( \Omega \) and taking values in the time set is called a stopping time if the event \( \{ \tau \leq t \} \) belongs to \( \mathcal{F}_t \) for all \( t \in (1, T) \). In other words, for \( \tau \), to be a stopping time, it should be possible to decide whether or not the event \( \{ \tau \leq t \} \) has occurred based on the knowledge that are known at time \( t \), i.e., the knowledge in the information set \( \mathcal{F}_t \). The stopping time for retirement decisions is called retirement stopping time. The retirement problem can be formulated as finding an optimal retirement stopping time, \( \tau^*_r \), from all retirement stopping times, \( \tau_r \), with values in \((1, T)\), that maximizes the expected discounted utility of retirement at time 1, i.e.,

\[
\sup_{1 \leq \tau_r \leq T} E_1 \left[ \exp \left( -\beta (\tau_r - 1) \right) \left( \prod_{k=1}^{\tau_r} M_k \right) U_{\tau_r} \right],
\]

(13)

where \( \prod_{k=1}^{\tau_r} M_k \) is the cumulative surviving probability from time 1 to \( \tau_r \) with \( M_1 = 1 \).

The annuitization timing decision is also an example of optimal stopping problems.
with fixed horizon. The stopping time for annuitization decisions is called annuitization stopping time. The annuitization time, $\tau_a$, must be between retirement time and the deadline for annuitization, that is, $\tau_a \in (\tau_r, T)$. The optimal annuitization stopping timing, $\tau^*_a$, is the stopping time that maximizes the expected discounted utility of pension income at retirement time $\tau_r$, with $\tau_r \in (1, T)$, that is,

$$
\sup_{\tau_r \geq \tau_a \leq T} E_{\tau_r}[\exp(-\beta(\tau_a - \tau_r)) B(\tau_r, \tau_a)],
$$

where $B(\tau_r, \tau_a) = \sum_{s=\tau_r}^{T_{\text{max}}} \exp(\beta(\tau_a - s)) \left( \prod_{k=\tau_r}^{s} M_k \right) \frac{DC(\tau_r, \tau_a)^{1-\gamma}}{1-\gamma}$ and the product, $\exp(-\beta(\tau_a - \tau_r))B(\tau_r, \tau_a)$, is the sum of the discounted utility of pension income at retirement time $\tau_r$.

The Least Square Monte Carlo (LSM) valuation algorithm developed by Longstaff and Schwartz (2001) is adopted to numerically solve the optimal stopping problem. The LSM algorithm follows the dynamic programming principle and provides a pathwise approximation to optimal stopping rules. At time 0, the stock price and interest rate at time 0 are known but future prices and interest rates are unknown. For each of the exercise dates, 1, ..., $T$, $N$ paths of stock prices and short-term interest rates are simulated. At time $t$ and path $i$, where $t$ could be any time between 1 and $T$ and $i$ could be any of the simulated paths, the individual would retire or annuitize if the utility of retiring or annuitizing at time $t$ and path $i$ is larger than the expected utility of retiring or annuitizing later conditional on the information available at time $t$ and path $i$. The LSM algorithm guarantees that there will be one and only one optimal stopping time for each path which solves (13) for the retirement option and (14) for the annuitization option.\(^1\) Details of the numerical solution and the LSM algorithm are provided in Appendix B.

### The Retirement Likelihood Measure

The probability estimated at time 0 of retiring before time $k$, $k$ could be any time between 1 and $T$, can be computed as follows. Let $\tau^*_{\tau_i}$ denote the optimal retirement time for path $i, i = 1, 2, ..., N$. Let $H$ be a $N \times T$ matrix, where the rows correspond to the simulated paths and the columns correspond to time. The matrix $H$ records the optimal retirement decisions of the individual. If $H(i, j) = 1, j$ is the optimal retirement time for path $i$.

\(^1\)Clément et al. (2002), Egloff (2005) and Moreno and Navas (2003) provide proofs for the convergence of the LSM algorithm.
otherwise, $j$ is not the optimal retirement time for path $i$, that is,

$$H(i, j) = \begin{cases} 1 & \text{if } j = \tau^*_i \\ 0 & \text{otherwise} \end{cases}.$$  \hspace{1cm} (15)

By construction, there will be only one "1" in each row.

From the optimal decision matrix, $H$, we can derive an estimator of the probability of retiring before and including time $k$, $k > 0$. The notation, $PR_{0}^{OptionModel}$, denotes the retirement probability and

$$PR_{0}^{OptionModel} = \frac{1}{N} \sum_{j=1}^{k} \sum_{i=1}^{N} H(i, j).$$  \hspace{1cm} (16)

At time 0, the probability that the individual will retire before and including time $k$ is the percentage of the paths where the optimal retirement times occur no later than time $k$. This probability is referred to as the retirement likelihood measure.

### III The Retirement Decision in the U.K.

In this section, the economic benefit of annuitization freedom will be evaluated and the likelihood measure developed in the previous section will be tested empirically. The empirical investigation is based on data from the English Longitudinal Study of Ageing (ELSA).

**Data**

ELSA is a biannual panel survey among those aged 50 and over (and their younger partners) living in private households in England. The field work for ELSA wave 1 is conducted in 2002-3 and for wave 2 in 2004-5. There are 12,100 individuals interviewed in wave 1. 1,659 individuals are employed full time (not less than 30 hours per week) and are interviewed again in wave 2. Among them, 518 persons participate in DC plans and provides complete information about their DC accounts. The sample consists of these 518 individuals. Detailed information about the sample selection is given in table 1.

In this sample, 29 persons retired by the wave 2 interviews of ELSA. None of the 29 persons report that their main reason of retirement is due to the sickness of themselves or their family members. 69.5% of the individuals are contracted out which means that
Table 1: Sample Selection. This table shows detailed information about the sample selection, including the reasons why individuals are removed from the sample and the number of individuals removed. Except reason (5), the other removal reasons are based on information reported at wave 1 interviews. There are 12,100 individuals participated in the ELSA wave 1 interviews. The sample in this paper consists of 518 individuals with DC pension plans.

<table>
<thead>
<tr>
<th>Removal Reasons</th>
<th>Number of Individuals Removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Younger than 50 or older than 90</td>
<td>673</td>
</tr>
<tr>
<td>(2) Not employed</td>
<td>8375</td>
</tr>
<tr>
<td>(3) Work less than 30 hours per week</td>
<td>873</td>
</tr>
<tr>
<td>(4) Incomplete information about education</td>
<td>148</td>
</tr>
<tr>
<td>(5) Do not participate in wave 2 interviews</td>
<td>372</td>
</tr>
<tr>
<td>(6) Do not have DC pension plan</td>
<td>1131</td>
</tr>
<tr>
<td>(7) The size of the DC scheme is not known</td>
<td>10</td>
</tr>
</tbody>
</table>
they cannot get retirement income from the second pillar state pensions\(^2\). In addition to the DC schemes, 31.27% of the individuals in the sample also have past DB plans and 11% of the individuals have current DB plans.

Our sample consists of 374 men and 144 women. 18.3% of the individuals have higher education or equivalent degrees. 30.5% of them didn’t receive high school education. The summary statistics of the DC plan participants’ age, gross income, DC wealth, asset income, benefit income, gross household wealth and debt are presented in table 2. The average age of the sample members is 55. The average annual gross income is about £24,400 and the average DC wealth is £33,122. Overall, the size of the average DC plan is small compared with the gross income. The small size could be caused by the short contribution records and the contributions to parallel pension plans, for example, DB plans and state pensions. DC pension plans started widely in the U.K. in the 1990’s, which means that the individuals in our sample started to contribute to the DC plan in their 40’s. Asset income, benefit income, gross household wealth excluding the primary housing and debt are at household level. Asset income consists of interest income, dividend income and the rent from second house, etc. Benefit income refers to state benefits, for example, Minimum Income Guarantee (MIG), Child Benefit and Disable Benefit. Gross household wealth is the household’s overall wealth excluding the house where they live.

\(^2\)A brief introduction to the U.K. pension system is provided in Appendix A.
Table 3: Summary Statistics for the Short-Term Interest Rate, 10 - Year Government Bond Yield and the Stock Market Return in the U.K. 1984 - 2002

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-Month UK T-Bill Rate</td>
<td>8.35%</td>
<td>3.00%</td>
</tr>
<tr>
<td>10 - Year Zero - Coupon UK Government Bond Yield</td>
<td>8.45%</td>
<td>2.37%</td>
</tr>
<tr>
<td>Annual Return of the U.K. Stock Index</td>
<td>15.78%</td>
<td>18.10%</td>
</tr>
</tbody>
</table>

I use quarterly data covering the period January 1984 to December 2002 for the variables describing the financial market. For the short rate, I use U.K. Treasury Bill data from Datastream. I obtain the yield of a 10-year U.K. government zero-coupon bond from the Bank of England. For the return on stocks I use the total return (including distributions) on a broad U.K. stock market index constructed by Datastream. The summary statistics for these variables are provided in table 3. The average return on stocks is 15.78%. The average yield on the 10-year zero-coupon bond is 8.45%. The average short rate is only slightly smaller (8.35%) during this particular sample period.

I use the Euler-Maruyama method to discretize the diffusion processes of the short rate, bond price and stock index. The parameters of these diffusion processes are estimated from the U.K. data discussed above. The estimation method is introduced briefly in Appendix C. The estimation results are as follows, $\kappa_r = 0.0232$, $\tau = 0.0129$, $\sigma_r = 0.0019$, $\lambda_s = 0.1639$, $\sigma_s = 0.0923$ and $\lambda_r = -0.0992$. The negative interest rate risk premium is in line with the literature (c.f. Brennan and Xia 2002). The interest rate risk premium is negative because investors are averse towards increases in interest rates while concerning stocks, investors are averse towards decreases in stock prices (De Jong, Schotman and Werker 2008).

**Projected Annual Incomes**

Information on past and future gross incomes is necessary to calculate the state pension and the DB pension income. The past and future gross income is projected based on the following variables: a gender dummy, experience which is defined as current age less
Table 4: Projection of Labor Income. The past and future gross incomes are projected based on the following variables: a gender dummy, experience which is defined as current age minus the age the individual started to work divided by 10, dummies for education degrees and years of schooling. High Education is a dummy, which equals to 1 if the individual has higher education or equivalent degree. Low Education is a dummy, which equals to 1 if the individual has educational degree lower than high school. The dependent variable is the log of the current gross annual income. Two stars means significance at 5 percent level.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>9.5517 **</td>
</tr>
<tr>
<td>Female</td>
<td>-0.3806 **</td>
</tr>
<tr>
<td>Experience</td>
<td>0.0818</td>
</tr>
<tr>
<td>Experience²</td>
<td>-0.028</td>
</tr>
<tr>
<td>Low Education</td>
<td>-0.2094 **</td>
</tr>
<tr>
<td>High Education</td>
<td>0.3424 **</td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>0.0519 **</td>
</tr>
</tbody>
</table>

The age the individual started to work divided by 10, dummies for education degrees and years of schooling. The (log) current gross annual income is regressed on the above mentioned variables and the square term of experience. The sample for testing the retirement likelihood measure consists of 518 individuals who work full time and have DC plans with complete information. But this analysis is based on the 1659 individuals who are working full time as reported at wave 1 interviews in order to make the projection more precise. The regression results are presented in table 4.

The regression results show that female workers earn significantly less than male workers. Individuals with high education degree (higher education degree or equivalent) earn significantly more than individuals with low (lower than high school degree) and medium education (high school degree) degrees. Income also increases with years of schooling. Experience and its square term have correct signs but they are both insignificant which could due to the fact that the individuals in the sample are of similar age.

In this paper the inflation rate, \( \pi \), is assumed to be constant at 2% level. The projected
past or future labor incomes for individual \(i\), \(Y_{projected, i}\), is

\[
Y_{projected, i} = EY(\theta_i) \exp(\pi(\theta - F_i)),
\]

where \(F_i\) is the individual \(i\)'s current age, \(\theta\) stands for individual \(i\)'s future age, \(\theta > F_i\), or past age, \(\theta < F_i\), and \(EY(\theta_i)\) denote the projected labor income of individual \(i\) at wave 1 interviews if he is \(\theta\) years old at that time which is derived from the regression reported in table 4.

**The State Pension**

The amount of state pension the individuals can receive depends on, among others, whether they are contracted out of the second-pillar state pension system and how long they have contributed to the state pension. The individual cannot receive their state pension until his State Pension Age is reached. If the individual delays receiving the state pension, the amount of pension is increased, at present, by approximately 7.5 per cent per year of delay in return. The maximum reward for deferment is 37 per cent, which is achieved by deferring for five years.

For the individuals who contracted out (in) in the wave 1, I assume that they contracted out (in) throughout their working life. Before 2002, the second pillar state pension is called SERPS. After 2002, the SERPS is replaced by S2P. But since S2P is only introduced in 2002, the individuals’ contribution records to S2P are very short. Therefore, this reform does not have big impact on the individuals’ pension income at 2004. Thus, in this paper, this reform is ignored. Department of Work and Pensions (2005) gives a very detailed description about the calculation of the first pillar state pension income (BSP) and the second pillar state pension income (SERPS) which is adopted for the calculation of state pension in this paper.

**The Economic Benefits of Annuitization Freedom**

I use the certainty equivalent gain to quantify the economic benefit of annuitization. The certainty equivalent gain is defined as the percentage difference between the certainty equivalent wealth at time 0 of retiring in the future with and without annuitization freedom. For the case without annuitization freedom, the individual has to annuitize at the retirement date. In the sample, the average discounted certainty equivalent gains are about 1.8%, 1.2% and 1.0% for individuals which retire at year 1, year 2 and year

---

3 Please see Appendix A for an introduction to the U.K. pension system.
The distributions of the certainty equivalent gains are shown in figure 1. The certainty equivalent gain distributions are almost identical for the two leisure parameter specifications. Assuming that all 518 individuals in the sample retire at year 1, certainty equivalent gains between 0% and 2% (the first column) are experienced by 320 individuals. There are about 100 individuals whose certainty equivalent gains are between 2% and 4% (the second column). The later the retirement date the smaller the certainty equivalent gains due to time preference, risk aversion and the shortening of the annuitization option life.

The Estimated Retirement Probability

It is assumed that the interviews of wave 1 are conducted at the end of 2002 and the interviews of wave 2 are conducted at the end of 2004. For the individuals who are reported to be retired at wave 2 interviews, the exact retirement years are not known. Based on the information available at wave 1 interviews, the probabilities of retiring by wave 2 interviews, \( PR_{2002}^{OptionModel} \), are estimated for every individuals in the sample from eq.(16).

I assume that during 2003 and 2004, at the beginning of each year the individual has a chance to consider retirement. The stock prices and bond prices for the years 2003 and 2004 are simulated from the diffusion processes (3) and (6). The value of the parameters of these diffusion processes are estimated from the market prices before and up to the end of 2002. 2000 paths for future stock and bond prices are simulated. The subjective discount factor, \( \beta \), is set to 0.03, the risk aversion parameter, \( \gamma \), is 5, and the preference for leisure, \( K \), equals to 1.5 in specification 1. In specification 2, the preference for leisure \( K \) equals to \( \left( \frac{F}{5} \right)^5 \). 70% of the portfolio assets are stocks and 30% are bonds, that is, \( \alpha = 0.7 \). The mortality rates are obtained from the U.K. Government Actuary’s Department (GAD). The maximum age an individuals can live is assumed to be 100.

Table 5 reports actual and average predicted percentages of individuals who retire during 2003 and 2004 for the whole sample (518 individuals) and two subsamples. Subsample 1 consists of individuals who were retired by wave 2 and subsample 2 consists of individuals who were not yet retired by wave 2. The actual percentage of retirement is 5.6%. The predicted percentage of retirement is 7.75% for specification 1 and 6.14% for specification 2. The predicted percentage of retirement for subsample 1 is 20.57% for specification 1 and 33.10% for specification 2. For subsample 2, the predicted percentage of retirement is 6.99% for specification 1 and 4.54% for specification 2.
Figure 1: The Certainty Equivalent Gain of Annuityization Freedom. This figure shows the certainty equivalent gain of annuityization freedom. The certainty equivalent gain is defined as the percentage difference between the certainty equivalent wealth at time 0 of retiring in the future with and without annuityization freedom. For the case without annuityization freedom, the individual has to annuitize at the retirement date. The certainty equivalent gain scale 1 refers to the interval $[0\%, 2\%]$, scale 2 refers to $[2\%, 4\%]$, scale 3 refers to $[4\%, 6\%]$, scale 4 refers to $[6\%, 8\%]$, scale 5 refers to $[8\%, 10\%]$, and scale 6 refers to $[10\%, +\infty]$. In specification 1, the leisure parameter is a constant, $K = 1.5$, and in specification 2 the leisure parameter is age dependent, $K = (F/55)^5$. Panel a, b and c show the certainty equivalent gains under the assumption that all 518 individuals retire at year 1, year 2 and year 3, respectively.
Table 5: The predicted retirement probability evaluated at the end of 2002 is an indicator measuring how likely the individuals will retire between 2003 and 2004. It is evaluated with the method discussed in section 2. This table reports the mean of the estimated retirement probability for the whole sample, the subsample consisting of individuals retired at wave 2, and the subsample consisting of individuals who are not retired at wave 2. There are two specifications for parameter K. In model specification 1, the disutility of work parameter K is a constant which equals to 1.5. In model specification 2, the disutility of work parameter is age dependent.

<table>
<thead>
<tr>
<th>Actual Percentage of Retirement</th>
<th>Predicted Retirement Probability $PR_{2002}^{OptionModel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K = 1.5$</td>
</tr>
<tr>
<td>Whole Sample</td>
<td>5.60%</td>
</tr>
<tr>
<td>Subsample 1: Retired</td>
<td>100.00%</td>
</tr>
<tr>
<td>Subsample 2: Not Retired</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
The Proxy of Retirement Incentive

In order to check whether the retirement likelihood measure, $PR_{2002}^{OptionModel}$, is significant in explaining and predicting the retirement decision making in reality, the retirement likelihood measure is treated as a proxy for retirement incentives. A Probit analysis is applied to test the significance of this proxy. The dependent variable is the sample individuals’ retirement decisions reported at wave 2 interviews which takes value 1 if the individual is reported to be retired and 0 if not. The variables, Asset Income (AI), Benefit Income (BI), Gross Household Wealth (GH) and Debt, which are not used for calculating $PR_{2002}^{OptionModel}$ are also included in the analysis. The results are presented in table 6. For both leisure parameter specifications, the proxy of retirement incentives, $PR_{2002}^{OptionModel}$, is positive and significant at 1% level no matter whether the other four variables are included or not. This analysis shows that the retirement likelihood has significant explanatory power in explaining and predicting the retirement decision in reality. It also means that financial incentives are important to the DC plan participants when they are making their retirement decision.

The Model Fit

The model fit is analyzed by comparing the actual retirement probability at wave 2 interviews, the predicted retirement probability from the option model based on wave 1 interview information, $PR_{2002}^{OptionModel}$, and the predicted retirement probability from a Probit model, $PR_{2002}^{Probit}$, where the regressors are variables such as, age, gender, education dummies, gross income and DC wealth, which are used for evaluating the retirement likelihood measure $PR_{2002}^{OptionModel}$ and the dependent variable is the retirement decision at wave 2. The probability of retiring by wave 2 interviews computed from this Probit model is actually an in-sample prediction. By contrast, the prediction from the option model is out of sample.

The probit regression reported in table 7 shows the impact of these variables on the individuals’ retirement decision in the sample. The results are very intuitive. Older individuals are significantly more likely to retire than younger ones. Women are significantly more likely to retire than men. This is because in the U.K., the State Pension Age for women at 2002 is lower than that for men. Age and gender are significant at 5% level. DC wealth, gross income and education dummies have expected signs, but they are insignificant. From the Probit model in table 7, for each individual we can compute the

---

4 The estimation error of the estimated retirement likelihood measure, $PR_{2002}^{OptionModel}$, is not taken into account when estimating the standard deviation of the estimated slope coefficient of this variable.
Table 6: Retirement Likelihood Measure As a Proxy for Retirement Incentive. This table reports the results of the Probit regression where the retirement probability derived from the option model is cheated as a retirement incentive. The dependent variable equals to 1 when the individual is reported to be retired at wave 2 and 0 otherwise. Asset income (AI) consists of interest income, dividend income and the rent from second house, etc. Benefit income (BI) refers to the state benefits, for example, Minimum Income Guarantee (MIG), Child Benefit and Disable Benefit. Gross household wealth (GH) is the household’s overall wealth excluding the house where they live. There are two specifications for parameter K. In model specification 1, the disutility of work parameter K is a constant which equals to 1.5. In model specification 2, the disutility of work parameter is age dependent. Panels A and B report the results from specification 1 and 2, respectively. One star stands for significance at 10 percent level and two stars stand for significance at 5 percent level and three stars stand for significance at 1 percent level.

<table>
<thead>
<tr>
<th></th>
<th>Submodel 1</th>
<th>Submodel 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Specification 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.6753***</td>
<td>-1.6612***</td>
</tr>
<tr>
<td></td>
<td>(0.0991)</td>
<td>(0.1193)</td>
</tr>
<tr>
<td>$PR_{2002}^{OptionModel}$</td>
<td>0.7288***</td>
<td>0.7554***</td>
</tr>
<tr>
<td></td>
<td>(0.2820)</td>
<td>(0.2895)</td>
</tr>
<tr>
<td>Asset Income (AI) / £1000</td>
<td></td>
<td>-0.0075</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0177)</td>
</tr>
<tr>
<td>Benefit Income (BI) / £100</td>
<td>0.0045</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0064)</td>
</tr>
<tr>
<td>Gross Household Wealth (GH) / £10000</td>
<td>0.0024</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0061)</td>
</tr>
<tr>
<td>Debt/£1000</td>
<td>-0.0168***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0181)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Submodel 1</th>
<th>Submodel 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel B: Specification 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.7709***</td>
<td>-1.7586***</td>
</tr>
<tr>
<td></td>
<td>(0.1049)</td>
<td>(0.1249)</td>
</tr>
<tr>
<td>$PR_{2002}^{OptionModel}$</td>
<td>1.3700***</td>
<td>1.4211***</td>
</tr>
<tr>
<td></td>
<td>(0.2687)</td>
<td>(0.2794)</td>
</tr>
<tr>
<td>Asset Income (AI) / £1000</td>
<td></td>
<td>-0.0115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0210)</td>
</tr>
<tr>
<td>Benefit Income (BI) / £100</td>
<td>0.0061</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0064)</td>
</tr>
<tr>
<td>Gross Household Wealth (GH) / £10000</td>
<td>0.0010</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0068)</td>
</tr>
<tr>
<td>Debt/£1000</td>
<td>-0.0013***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0017)</td>
<td></td>
</tr>
</tbody>
</table>
Table 7: Comparison Probit Regression. This table reports the results of the Probit regression of age, gender and other variables related to individual retirement decisions. The dependent variable equals to 1 when the individual is reported to be retired at wave 2 and 0 otherwise. The gender dummy equals to 1 for woman and 0 for man. The high education dummy equals to 1 for the individuals with higher education degree or equivalent. The low education dummy equals to 1 for the individuals with degree lower than high school degree. One star stands for significance at 10 percent level and two stars stand for significance at 5 percent level and three stars stand for significance at 1 percent level.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-11.0969 ***</td>
</tr>
<tr>
<td></td>
<td>(1.6676)</td>
</tr>
<tr>
<td>Age</td>
<td>0.1642 **</td>
</tr>
<tr>
<td></td>
<td>(0.0281)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.5301 **</td>
</tr>
<tr>
<td></td>
<td>(0.2331)</td>
</tr>
<tr>
<td>High Education</td>
<td>-0.5086</td>
</tr>
<tr>
<td></td>
<td>(0.3671)</td>
</tr>
<tr>
<td>Low Education</td>
<td>0.0728</td>
</tr>
<tr>
<td></td>
<td>(0.2242)</td>
</tr>
<tr>
<td>DC Wealth/1000</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td>(0.0021)</td>
</tr>
<tr>
<td>Gross Income/1000</td>
<td>-0.0025</td>
</tr>
<tr>
<td></td>
<td>(0.0067)</td>
</tr>
</tbody>
</table>
(in-sample) probability of retiring by wave 2 interviews, $PR_{\text{2002}}^{\text{Probit}}$.

As in Stock and Wise (1990), I divide the sample into several age groups and then compare the actual retirement ratio in each age group with the predictions from the option model, $PR_{\text{2002}}^{\text{Option\_Model}}$, and from the Probit model, $PR_{\text{2002}}^{\text{Probit}}$. The results are shown in figure 2 and table 8. It can be seen from figure 2, that the actual retirement probability increases with age. The predictions from the option model catches this trend very well especially the one from specification 2. The correlations between the option model probabilities and the actual retirement probabilities are 0.89 for model specification 1 ($K = 1.5$) and 0.94 for model specification 2 ($K = (\frac{F}{55})^5$). The correlation between the (in-sample) Probit model probabilities and the actual probabilities is 0.94. Furthermore, the option model probabilities have roughly the same roots of Mean Square Errors (MSEs) as those from the Probit analysis. The root of the MSE is 6% for the Probit model, 7% for the option model specification 1 and 5% for the option model specification 2.

The sample was also divided by the DC wealth level. Level 1 includes the individuals with DC wealth smaller than £5,000. Level 2 includes individuals with DC wealth larger than £5,000 but smaller than £10,000 and so on until level 7 which is the highest level and includes the individuals with DC wealth larger than or equal to £150,000. The results are reported in figure 3 and table 9. Overall, the actual retirement probability is increasing with the DC wealth level. The correlation coefficient between the actual retirement ratio (column 2 in table 9) and the predicted retirement ratio from the Probit model is 0.78. The correlation coefficient between the actual retirement ratio and the predicted retirement ratios from the option models are about 0.72 for specification 1 and 0.67 for specification 2. The root of the MSE of the Probit model is 2%. The roots of the MSEs of option model specifications 1 and 2 are 27% and 19%, respectively. The relatively large MSEs are due to the prediction errors for the very wealthy groups (group 6 and 7). There are 39 individuals in these two groups. The roots of the MSEs of option model specifications 1 and 2 without these individuals are about 3%.

Generally speaking, the performance of the option model, especially using the model specification where the leisure parameter is age dependent, in terms of correlations with the actual retirement probabilities and the roots of the Mean Square Errors, are comparable to the performance of the in-sample Probit predictions.
Figure 2: Actual and Predicted Retirement Ratios By Age Groups. This figure shows the actual and predicted retirement ratios from the Probit model and the option model which are reported in table 8.
Table 8: Actual and Predicted Retirement Ratios By Age Groups. This table shows the actual and predicted retirement ratios from the Probit model presented in table 7 and the option model described in section 2. The actual percentage of retirement measures the percentage of individuals retired by the end of 2004 for each age group. In the option model, there are two specifications for the leisure parameter, K. In specification 1, the disutility of work parameter K is a constant which equals to 1.5. In model specification 2, the disutility of work parameter is age dependent.

<table>
<thead>
<tr>
<th>Age</th>
<th>No. of Obs</th>
<th>Actual Percentage of Retirement</th>
<th>In-Sample Probit Model Prediction</th>
<th>Out-of-Sample Option Model Prediction</th>
<th>K = 1.5</th>
<th>K = (F/55)^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>51</td>
<td>50</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>52</td>
<td>53</td>
<td>0.02</td>
<td>0.01</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>53</td>
<td>56</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>54</td>
<td>59</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>55</td>
<td>58</td>
<td>0.03</td>
<td>0.03</td>
<td>0.06</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>56</td>
<td>50</td>
<td>0.00</td>
<td>0.04</td>
<td>0.10</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>57</td>
<td>25</td>
<td>0.12</td>
<td>0.06</td>
<td>0.08</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>58</td>
<td>36</td>
<td>0.08</td>
<td>0.07</td>
<td>0.12</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>59</td>
<td>33</td>
<td>0.03</td>
<td>0.10</td>
<td>0.11</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>60</td>
<td>21</td>
<td>0.05</td>
<td>0.12</td>
<td>0.10</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>61</td>
<td>18</td>
<td>0.17</td>
<td>0.17</td>
<td>0.23</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>62</td>
<td>12</td>
<td>0.25</td>
<td>0.22</td>
<td>0.16</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>&gt;=63</td>
<td>18</td>
<td>0.50</td>
<td>0.33</td>
<td>0.31</td>
<td>0.56</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Corr. Coef. With Column 3  
Root of Mean Squared Error
Figure 3: The Actual and Predicted Retirement Ratios By DC Wealth Group. This figure shows the actual and predicted retirement ratios from the Probit model and the option model. The actual percentage of retirement measures the percentage of individuals retired by the end of 2004 for each wealth group. The Probit model is described in table 7. The option model described in section 2. Group 1 are the individuals with DC wealth less than £5,000, Group 2 includes the individuals with DC wealth between £5,000 and £10,000, Group 3 includes the individuals with DC wealth between £10,000 and £25,000, Group 4 includes the individuals with DC wealth between £25,000 and £50,000, Group 5 includes the individuals with DC wealth between £50,000 and £100,000, Group 6 includes the individuals with DC wealth between £100,000 and £150,000, and Group 7 includes the individuals with DC wealth larger or equal to £150,000.
Table 9: Actual and Predicted Retirement Ratio By DC Wealth Groups. This table shows the actual and predicted retirement ratios from the Probit model and the option model. The actual percentage of retirement measures the percentage of individuals retired by the end of 2004 for each wealth group. Probit model is described in table 7 and the prediction from the option model is described in section 2.

<table>
<thead>
<tr>
<th>Level</th>
<th>DC Wealth (in £1000)</th>
<th>No. of Obs</th>
<th>Actual Ret. Percentage</th>
<th>In-Sample Probit Model Prediction</th>
<th>Out-of-Sample Option Model Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt; 5</td>
<td>132</td>
<td>0.05</td>
<td>0.06</td>
<td>0.02</td>
</tr>
<tr>
<td>2</td>
<td>[5, 10]</td>
<td>70</td>
<td>0.04</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>[10, 25]</td>
<td>145</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>4</td>
<td>[25, 50]</td>
<td>56</td>
<td>0.07</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>5</td>
<td>[50, 100]</td>
<td>76</td>
<td>0.08</td>
<td>0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>6</td>
<td>[100, 150]</td>
<td>17</td>
<td>0.12</td>
<td>0.10</td>
<td>0.42</td>
</tr>
<tr>
<td>7</td>
<td>&gt;= 150</td>
<td>22</td>
<td>0.09</td>
<td>0.11</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Corr. Coef with Column 4: 0.78 0.72 0.67
Root of Mean Square Error: 0.02 0.27 0.19
IV Conclusions

This paper models the real options to retire and - conditional on retirement - to annuitize the accumulated wealth in Defined Contribution (DC) pension schemes. It contributes to the extensive literature on the option value of retirement in the tradition of Stock and Wise (1990) but shifts the focus from participants in Defined Benefit (DB) schemes to participants in DC schemes. This accounts for the observation that DB schemes are increasingly being replaced by DC schemes in most industrialized countries including the U.S. and the U.K.. A major contribution of this paper is to recognize and model the sequence of retirement and annuitization options. Since annuitization does not need to occur at the retirement date in many DC schemes, participants in DC schemes can time the financial markets in order to annuitize in an environment of high prices on the stock market which increase the DC wealth and high interest rates which reduce the price of an annuity. In a model where individuals obtain utility from leisure, labor income before retirement and pension income after retirement, I show that the freedom to optimally choose the annuitization time can lead to an increase of certainty equivalent wealth of up to 1.8%. Hence, the embedded annuitization option in the retirement option value is of significant economic value to individuals.

In order to assess the predictive power of my model, I compare retirement likelihoods derived from the theoretical setup with retirement decisions observed at the second wave of the English Longitudinal Study of Ageing (ELSA) for a sample of individuals who were full-time employed at the time of the first wave interviews. It turns out that the theory-motivated retirement likelihood measure is a statistically significant predictor of actual retirement decisions. Moreover, I show that the proposed retirement likelihood measure is highly correlated with observed retirement ratios across groups of individuals defined by age or wealth. The correlation reaches 94% and is not dominated by the predictions of a retirement probit model which in contrast to my proposed retirement likelihood measure is based on in sample information. With a magnitude of 5%, root mean square errors turn out to be small. These empirical results suggest that individuals do take into account the embedded annuitization option when they decide on when to retire.
References


Coile, C. and J. Gruber (2000), Social security incentive for retirement, NBER working paper 7651.


Department of Work and Pension (2005), A guide to the state pension, NP46.


FDIC (2006), The shift away from defined benefit plans, FDIC Outlook.

Friedberg, L. and A. Webb (2003), Retirement and the evolution of pension structure, NBER working paper 9999.


Ross, D. and L. Wills (2002), The shift from defined benefit to defined contribution retirement plans and the provisioning of retirement savings, Pension Institute Discussion Paper PI-0210.


Appendix A

The British Pension System

The U.K. pension system consists of three main pillars. The first pillar, known as Basic State Pension (BSP), is a mandatory, flat rate state pension\(^5\). The second pillar system is provided by the state, employers and private sector financial institutions. In the second pillar, the employees have considerable choices over the type of pension that they can accumulate. The main choices are between: (1) an earnings-related state pension plan\(^6\); (2) an occupational DB plan provided by employers and (3) an occupational DC pension plan. The state pension plan offers a pension that is low relative to average earnings, but is fully indexed to prices after retirement. The occupational DB plan offers a relatively high level of pension to the employees who spend most of their working time with the same employer, but provides poor transfer values between plans on changing jobs. The occupational DC pension plan is fully portable, but the pension income depends on uncertain investment returns (see Blake 2003). The second pillar state pension is by default compulsory to all the employees who earn above a lower threshold set by the state. But individuals are able to contract out of the second pillar state pension into an occupational pension scheme provided that the latter is at least as generous as the second-pillar state pension. The third pillar consists of voluntary private pension plans\(^7\). The third pillar pension arrangements are usually of DC type.

In the U.K., the DC plan participants do not have to annuitize their DC wealth immediately at the retirement date. Up to one-quarter of the value of a pension fund can be taken as a lump sum, but three-quarters must be annuitized before the age of 75 (Finance Act 1995)\(^8\). The obligation to annuitize DC wealth and the freedom in choosing the annuitization time are the most important reasons why U.K. data is selected for the empirical investigation in this paper.

---

\(^5\)The BSP is funded on a pay-as-you-go basis. It is a flat rate benefit. Individuals are entitled to at least some part of the BSP if they have made National Insurance (NI) contributions for at least 25% of their working lives. The BSP benefit in 2006/7 is about £85 per week (Department of Work and Pensions). This benefit is indexed to inflation (Clark and Emmerson 2003).

\(^6\)The second pillar state pension plan was called State Earnings-Related Pension Scheme (SERPS) and replaced by State Second Pension (S2P) in 2002. The second pillar state pension plans are of DB nature (Cocco and Lopes 2004). Both the first and second pillar state pensions are paid by the Department of Pension and Working once the retiree reaches his State Pension Age (SPA). Currently the State Pension Age is 65 for men and 60 for women. By 2020, the SPA for woman will increase gradually to 65.

\(^7\)Employers and individuals can also make additional contributions to a private pension. The state supports the savings in private pension plans through tax relief (see Clark and Emmerson 2003).

\(^8\)Since April 6, 2006, which is after the second wave interviews of ELSA, the individuals at age 75 can also choose to drawdown their DC wealth without annuitization, known as Alternatively Secured Pension (ASP). But tax charges introduced by the government on ASP make this option very unattractive to retirees.
Appendix B

The Solution Method

The optimal annuitization and retirement decisions are very similar to the decision of exercising an American option optimally, in the sense that, like the American option, both the retirement and annuitization decisions can be made at any stopping time between the "purchase" date, in our cases, the time when the individual is allowed to retire/annuitize, and the "expiration" date, in our cases, the time when the individual turns 75 years old.

Let \( n \) be the "purchase" date of an annuitization option or a retirement option. The optimal annuitization and retirement stopping problems can be stated as

\[
V_n = \sup_{n \leq \tau \leq T} E_n \left[ \exp \left( -\beta (\tau - n) \right) Z_{\tau} \right],
\]

(17)

where the function \( Z(\cdot) = \left( \prod_{k=1}^{\tau} M_k \right) U(\cdot) \) for the retirement option, \( Z(\cdot) = B(\cdot) \) for the annuitization option, \( n = 1 \) for the retirement option and \( n = \tau_r \) for the annuitization option.

The standard solution to a optimal stopping problem with finite horizon is to follow the dynamic programming principle (c.f. Peskir and Shiryaev 2006). Let \( J_t \) be the highest attainable expected utility at time \( t \) the individual can achieve if he exercises his option at or later than time \( t \), that is,

\[
J_t = \sup_{t \leq \tau \leq T} \exp \left( -\beta (\tau - t) \right) E \left( Z_{\tau} | \mathcal{F}_t \right).
\]

Here exercising an option means retiring for the retirement option and annuitizing for the annuitization option. At time \( t = T \), the individual has to stop immediately and gains \( J_T = Z_T \). At time \( t = T - \Delta t \), where \( \Delta t \) stands for very short period of time, he can either stop or continue. If he stops, \( \tau = t \) and \( J_{T-\Delta t} \) equals to \( Z_{T-\Delta t} \), and if he continues, \( \tau = T \) and \( J_{T-\Delta t} \) equals to \( \exp \left( -\beta \Delta t \right) E \left( J_T | \mathcal{F}_{T-\Delta t} \right) \). It follows that if \( Z_{T-\Delta t} \geq \exp \left( -\beta \Delta t \right) E \left( J_T | \mathcal{F}_{T-\Delta t} \right) \) then he needs to stop at time \( t = T - \Delta t \); otherwise, he needs to continue at time \( t = T - \Delta t \). This decision rule reflects the fact that the individual’s decision about stopping or continuation at time \( t = T - \Delta t \) must be based on the information contained in \( \mathcal{F}_{T-\Delta t} \) only. For \( t = T - 2\Delta t, ..., n \), the considerations are continued analogously.

The method of backward induction just explained leads to a sequence of random
variables, \((J_t)_{n\leq t\leq T}\), defined recursively as follows:

\[
J_t = \begin{cases} 
  Z_T & \text{for } t = T, \\
  \max(Z_t, \exp(-\beta \Delta t) E(J_{t+\Delta t} | \mathcal{F}_t)) & \text{for } t = T - \Delta t, \ldots, n.
\end{cases}
\]

The method also suggests that we consider the following stopping time

\[
\tau_n = \min \{n \leq k \leq T : J_k = Z_k\} 
\]  

as a candidate for optimal stopping time for problem (17). Peskir and Shiryaev (2006) proved that \(\tau_n\) is indeed the optimal stopping time in (17). The proof is provided in Appendix C.

At time \(t, t < T\), the value of immediate exercise, \(Z_t\), is known to the individual. But the value of \(\exp(-\beta \Delta t) E(J_{t+\Delta t} | \mathcal{F}_t)\) is still unknown. The key to solve the optimal stopping problem (17) is therefore, to evaluate the conditional expectations, \(\exp(-\beta \Delta t) E(J_{t+\Delta t} | \mathcal{F}_t)\) for \(t = T - \Delta t, \ldots, n\). Least Square Monte Carlo (LSM) valuation algorithm developed by Longstaff and Schwartz (2001) is adopted to approximate \(E(J_{t+\Delta t} | \mathcal{F}_t)\) and to solve optimal stopping problem numerically. Clément, Lamberton and Protter (2002), Egloff (2005) and Moreno and Navas (2003) proved the convergence of the LSM algorithm.

**The Least Square Monte Carlo (LSM) Algorithm**

The objective of the LSM algorithm is to provide a pathwise approximation to the optimal stopping rules. It is assumed that the option can only be exercised and considered at a finite number of discrete times, \(n, \ldots, t, t + \Delta t, \ldots, T\). For each exercise date, \(n, \ldots, T\), \(N\) paths (scenarios) of stock prices and short-term interest rates are simulated.

The LSM algorithm follows the standard backward induction method as described previously. At the final expiration date, \(T\), the option has to be exercised, the individual gets \(Z_{T,i}\), where \(i\) stands for a simulated path and \(i = 1, 2, \ldots, N\). At exercise dates before the final expiration date, say time \(t\), the individual must choose whether to exercise the option immediately or to keep the option alive and make the exercise decision at the next exercise date. At time \(t\), for any path \(i\), where the utility from immediate exercise, \(Z_{t,i}\), is larger than or equal to the expected utility of continuation conditional on the information available at time \(t\) and path \(i\), \(\exp(-\beta \Delta t) E(J_{t+\Delta t} | \mathcal{F}_{t,i})\), it is optimal to exercise the option. For any paths where the opposite holds, it is optimal to wait.
At time $t$ and path $i$, the value of immediate exercise, $Z_{t,i}$, is known to the individual but the value of waiting, $E(J_{t+\Delta t} | \mathcal{F}_{t,i})$, is unknown. The conditional expectation at time $t$ and path $i$, $(-\beta \Delta t) E(J_{t+\Delta t} | \mathcal{F}_{t,i})$, is approximated by regressing the vector of discounted value of continuation at time $t$, $\exp(-\beta \Delta t) J_{t+\Delta t}$, where $J_{t+\Delta t} = (J_{t+\Delta t,1}, J_{t+\Delta t,2}, \ldots J_{t+\Delta t,N})'$, on the simulated paths of relevant state variables at time $t$, $X'_i$s where $X'_i$s include the utility at time $t$ and the DC wealth at time $t$.

Let $\hat{E}^m (J_{t+\Delta t} | X_{t,i})$ denote the estimated conditional expectation at time $t$ and path $i$. The individual will decide at time $t$ whether to exercise the option or not. For the paths where the value of immediate exercise, $Z_{t,i}$, is larger (smaller) than or equal to the estimated conditional expectation, $\exp(-\beta \Delta t) \hat{E}^m (J_{t+\Delta t} | X_{t,i})$, it is optimal to exercise the option (wait). Proceed these calculations and comparisons recursively backwards until the "purchase" date is reached. The optimal stopping time for each path is then decided by starting from the "purchase" date, moving along each path until the first stopping time. For each path, the first stopping time is the optimal exercise time for that path. Thus, there will be one and only one optimal stopping time for each path.

**Appendix C**

We have

\begin{align}
J_n &\geq E(Z_{\tau_n} | \mathcal{F}_n) \text{ for each } \tau \in (n, T), \\
J_n &= E(Z_{\tau_n} | \mathcal{F}_n).
\end{align}

Taking expectation in (19), we find that $EJ_n \geq E(Z_{\tau_n} | \mathcal{F}_n)$ for all $\tau \in (n, T)$ and hence by taking the supremum over all $\tau \in (n, T)$ we see that $EJ_n \geq V_n$. On the other hand, taking the expectation in (20), we get $EJ_n = E(Z_{\tau_n} | \mathcal{F}_n)$. Since $\tau_n \in (n, T)$ and (17), it holds that $E(Z_{\tau_n} | \mathcal{F}_n) \leq V_n$ and therefore, $EJ_n \leq V_n$. The two inequalities give the equality $V_n = EJ_n$, and since $EJ_n = E(Z_{\tau_n} | \mathcal{F}_n)$, we see $V_n = E(Z_{\tau_n} | \mathcal{F}_n)$ implying that $\tau_n$ is the optimal stopping time to the problem (17).

**Appendix D: The Parameter Estimation**

In this subsection, the parameters of the diffusion processes, (2), (3) and (6), will be estimated. The Euler-Maruyama method is used to derive the discrete-time approximations of these diffusion processes. For the short term interest rate, the discrete-time
approximation is

\[ r_{t+\Delta t} - r_t = \kappa_r (\tau - r_t) \Delta t + u_{r,t+\Delta t}, \quad (21) \]

\[ r_{t+\Delta t} = \alpha + \beta_r r_t + u_{r,t+\Delta t}, \quad (22) \]

where the error term, \( u_{r,t+\Delta t} = \sigma_r \Delta Z_1 \) with \( \Delta Z_1 = Z_{1,t+\Delta t} - Z_{1,t} \), is normal distributed with

\[ E_t(u_{r,t+\Delta t}) = 0, \]

\[ E_t(u_{r,t+\Delta t}^2) = \sigma_r^2 \Delta t, \]

\( \alpha = \kappa_r \tau \Delta t \) and \( \beta_r = 1 - \kappa_r \Delta t \). The discrete-time approximation of the stock index is

\[ S_{t+\Delta t} - S_t = S_t (r_t + \lambda_s \sigma_s) \Delta t + u_{s,t+\Delta t}, \quad (23) \]

where the error term, \( u_{s,t+\Delta t} = \sigma_s \Delta Z_2 \) with \( \Delta Z_2 = Z_{2,t+\Delta t} - Z_{2,t} \), has the properties

\[ E_t(u_{s,t+1}) = 0 \]

\[ E_t(u_{s,t+1}^2) = \sigma_s^2 S_t^2 \Delta t. \]

The distribution of the excess return on stock index can be approximated by a normal distribution with mean \( \lambda_s \sigma_s \Delta t \) and variance \( \sigma_s^2 \Delta t \). For this estimation, \( \Delta t \) is taken to be 1, referring to 1 quarter of a year.

The estimation of the AR(1) model (22) is presented in table 10. The AR(1) term of the short rate, \( \beta_r \), is significant at 1% level. From the estimation reported in table 3, we can get \( \kappa_r = 0.0232 \) and \( \tau = 0.0129 \). The volatility of the short rate, \( \sigma_r \), is derived from the residuals of the two AR(1) process and \( \sigma_r = 0.0019 \). The price of risk for stock index and the volatility of stock index are estimated from the distribution of excess return of stock index. We have \( \lambda_s = 0.1639 \) and \( \sigma_s = 0.0923 \).

Let the yield of a 10-year zero-coupon government bond derived from Vasicek model be \( Y_t \), which is a function of \( \lambda_r \) and let \( Y_t \) stands for the yield in the data sample. \( \lambda_r \) is estimated by minimizing the objective function, \( F(\lambda_r) \),

\[ F(\lambda_r) = \frac{1}{T} \sum_{t=1}^{T} (\hat{Y}_t(\lambda_r) - Y_t)^2. \]

The price of risk for short-term interest rate, \( \lambda_r \), is \(-0.0992\).
Table 10: Estimation of AR(1) Processes for Stock Index and Short-Term Interest. One star means significance at 10 percent level, two stars mean significance at 5 percent level and three stars mean significance at 1 percent level.

<table>
<thead>
<tr>
<th>Short-Term Interest Rates</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
</tr>
<tr>
<td>$\beta_r$</td>
<td>0.9768 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0319)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9278</td>
</tr>
</tbody>
</table>