An economic analysis of retirement decisions in Taiwan

Hung, Wen-Shai

Chapter 4

Retirement Decisions

4.1 Introduction

This chapter investigates the factors contributing to the retirement decisions of the middle-aged and elderly in Taiwan. As previously mentioned, life expectancies in Taiwan have been increasing in the last decade and are expected to continue rising. The post-war baby boom generation has gradually reached retirement age and, as a result, retirement issues will become more important in the next decade, especially for policy makers. For instance, if workers retire earlier and also live longer than in the past, more resources will be required to support an extended period of retirement. Some workers may amass enough personal savings to support themselves through old age. Other workers may be forced to retire because of ill health, economic downturn, or reaching a mandatory retirement age. This group will probably need additional resources to support them in their old age, either from younger members of their families or from a social security or pension system. In traditional Taiwanese societies, the elderly live in extended, multigenerational households and rely primarily on their adult children for financial support and personal care. However, this traditional family support system is under pressure from demographic, social, and economic changes. In particular, fertility has been low for decades, the elderly have few adult children to provide support, and many of these children have moved away from their family homes. For instance, the percentage of elderly living with children declined from 67% in 1976 to 51% in 1996 (Hermalin, 2000). Therefore, better government social welfare policies are needed to provide more financial help to elderly people.

This chapter contributes to the literature by focusing attention on the retirement decision of the middle-aged and elderly. Most previous studies on retirement behaviour in Taiwan have focused on living arrangements (Chen, 1994; Chang, 1999), the economic well-being (Hermalin et al, 1999), and health status and health-care utilisation of the elderly (Chen, 1999). Furthermore, Taiwan is a unique country with a different social system from most developed countries. Contemporary labour studies have often focused on the effects of retirement age, health, and pensions on retirement (Gustman and Steinmeier, 1986; Stock and Wise, 1990; Berkovec and Stern, 1991). However, straightforward application of these studies to Taiwan is inappropriate because of the social differences.[1] For example, according to the Taiwanese Labour Standards Law, individuals who worked for more than 25 years could apply for voluntary retirement and obtain occupational pension entitlement even before they reached 65. As such, using retirement age as an explanatory variable here is not always appropriate since it may vary across individuals due to different ages of starting work. Taiwanese workers who started working at an early age retire earlier, and those who start later retire later. Furthermore, until now, Taiwan's government has not provided a public or state pension system for the elderly, even after they were 65 and over. Many relied on personal savings and family welfare. To take into account these considerations, this Chapter uses the concept of employment duration to analyse retirement decisions in Taiwan.

Duration analysis has been developed in the field of bio-statistics to describe the

[1] Due to the limitations of the SHLS data, the different social security system and specific retirement law in Taiwan, standard models such as the life cycle model (Gustman and Steinmeier, 1986), option value model (Stock and Wise, 1990) or stochastic dynamic programming model (Berkovec and Stern, 1991) are not suitable for analysing the retirement decisions of the middle-aged and elderly in Taiwan. In particular, the above models depended on sufficient data on wages, income, pension benefit, and other assets.

timing of events. It has become a subject of increasing interest to labour economists – notably to issues on unemployment, retirement, and absence from the workplace. Most methods for analyzing duration models are based on time as a continuously measured variate. For example, Lancaster (1979) used the continuous-time parametric models to estimate the effects of regressors on the expected duration of individuals' unemployment, Diamond and Hausman (1984) applied the Cox proportional hazard model to analyse the continuous-time hazards of retirement, Barmby, Orme and Treble (1991) considered a dynamic model of worker decisions under the sick-pay scheme and used the discrete panel data and Weibull hazard models to analyse the duration of absence. Although the observed data are discrete, time is obviously continuous. Therefore, this study specifies that employment durations are continuous and uses the continuous-time hazard models, including parametric and semi-parametric approaches to assess individuals' retirement behaviour. The former approach is guided by the assumed distribution for the hazard function while the latter is more flexible with no assumed distribution for the hazard function.

Previous empirical studies have often tested the hypothesis that higher wages lead to delayed retirement, and higher pension benefits from social security or private pensions lead to earlier retirement (Mitchell and Fields, 1981). To judge the adequacy of social security payments, they often compared current benefits to a recipient's previous wages to find the optimal replacement rate[2] (Steuerle et al, 2000). However, the main limitation of the SHLS surveys lies with the fact that insufficient data have been collected on wages and assets, and the response rates to these questions were

[2] According to Wang (2000), the optimal replacement rate can be calculated as follows: the employment duration is divided by the sum of employment duration and retirement duration. For example, if the employment duration is 45 years, and retirement is 15 years, then the optimal replacement rate is 75%. If, other things being equal, the employment duration is longer, then the optimal replacement rate increases.

low (Mete and Schultz, 2002). Furthermore, annual income will vary with age, of course, especially between individuals of working age and those of retirement age. To address these problems, this chapter constructs a predicted earnings variable to approximate the wage level, particularly as an indicator of income from work, and a predicted pension income variable to estimate the pension benefits as an indicator of income during retirement. This approach is not new but it has been infrequently adopted in economic studies. The problem is that the estimated standard errors of such constructed variables typically will under-estimate the true standard errors. Pagan (1984) demonstrated this and showed how the standard error should be adjusted. Empirical studies that employed predicted variables include Diamond and Hausman (1984), Slade (1987), Arulampalam and Stewart (1995), and Buckley et al. (2004), who estimated the effect of predicted income on retirement and unemployment duration studies. Diamond and Hausman (1984), for example, noted that predicted pension income has a positive effect on the hazard rate of retirement. Slade (1987) found that predicted earnings have a significant negative effect on labour force exit, and Arulampalam and Stewart (1995) stated that predicted earnings can increase the probability of entering full-time work but this effect falls with age. Buckley et al. (2004) suggested that predicted lifetime income or wealth has a positive effect on good health among older people.

Consequently, two predicted variables, namely predicted earnings as an indicator of income from work and predicted pension income as an indicator of income during retirement, are used to test the following hypotheses on the duration models: If workers have higher predicted earnings, they have a lower hazard rate of retirement. If people have higher predicted pension income, they have a higher hazard rate of retirement. In general, two alternative specifications for the duration models are adopted in this Chapter. (1) Case 1: the hazard function excludes predicted earnings and predicted pension income as explanatory variables, (2) Case 2: the hazard function includes predicted earnings and predicted pension income as explanatory variables. Other explanatory variables, such as Age, Gender, and Health, are also considered to examine the determinants of retirement behaviour.

Further, if the study did not consider the influence of unobserved heterogeneity on estimated duration model, they might produce incorrect results (Lancaster, 1990). The theoretical literature also suggested that the non-frailty model might over-estimate the degree of negative duration dependence in the true baseline hazard, and under-estimate the degree of positive duration dependence in the true baseline hazard (Jenkins, 2005). Hence, this chapter considers the effects of unobserved heterogeneity into the duration model for testing the above theoretical suggestions on retirement behaviour.

The rest of this chapter proceeds as follows. Section 4.2 presents the theoretical framework, including related literature, and outlines a model of the retirement decision. Section 4.3 describes the empirical specification, including specifying employment durations by continuous-time, rather than discrete-time hazard models, the distribution of employment duration, and the estimation of the hazard function. Parametric and semi-parametric approaches, including the exponential, Weibull, and Cox proportional hazard models, are used to estimate the hazard rate of retirement. Section 4.4 presents the data sources and variables used in this analysis. This is followed by the major empirical results in Section 4.5. In particular, the effects of unobserved heterogeneity on retirement decision are also tested by using duration models. Lastly, Section 4.6 discusses the results and section 4.7 concludes the chapter.

4.2 Theoretical Framework of Retirement Decisions

This section develops a simple retirement framework for the middle-aged and elderly. First we review some literature on retirement issues and secondly build up an employment duration model to account for the independent variables and data structure. Several hypotheses are proposed relating to the retirement hazard.

4.2.1 Literature Review on Employment Duration Models

This section reviews some literature on retirement behaviour using the employment duration model, including Diamond and Hausman (1984), Antolin and Scarpetta (1998), and An et al. (1999). Diamond and Hausman (1984) were the first to employ the regression-type hazard model to examine the determinants of individual retirement and savings. Their specification of the model solved three problems: censoring, dynamic regressor variables, and dynamic self-selection. For instance, to resolve the sample-censoring problem, hazard models are used instead of the more traditional regression-type models. They divided their sample into three groups, including left censoring (individuals who retired before the beginning of the sample period), right censoring (individuals who do not retire during the sample period), and event or failure time (individuals who retire during the sample period). From these three sets of individuals, they calculated the likelihood of retirement. The data in their empirical study was obtained from the US National Longitudinal Survey (NLS) over the period 1966-1976. Diamond and Hausman (1984) found that the presence of pension and social security benefits had a significant positive effect on retirement duration. They also pointed out that the proposed increase in the minimum age for receipt of social security benefits could slow down the retirement rate but had not stopped the trend of decreased male labour force participation in the US over the past 20 years, and that individuals with larger permanent incomes or with high earnings capacity were less likely to retire. Diamond and Hausman further found that the demographic variable with the largest effect was poor health, and single males were much more likely to retire earlier. The effects of education and number of dependents were negligible.

Antolin and Scarpetta (1998) analysed the determinants of retirement decisions in Germany using micro data from the German Socio-Economic Panel (GSOEP) over the period 1985-1995. Their non-parametric estimates suggested that the incentive structure generated by different social security schemes played a powerful role in individual retirement decisions. Their semi-parametric analysis was conducted using a piece-wise constant hazard model with multiple destinations and time-varying covariates, and they found that socio-demographic factors such as poor health have a strong impact on retirement decisions contributing to early withdrawal from the labour market. Financial incentives offered in the pension system are powerful influences in shaping the age profile of retirement. In particular, they used pension wealth and an estimate of retirement option values to allow for the planning behaviour of individuals. They found that older people tended to maximise the net value of their pension wealth and retire as soon as the option value of postponing retirement became small. Finally, the results of the hazard model were used to simulate the effects of a reform towards an age-neutral pension system. The results of this simulation suggested a significant shift in the age profile of retirement, with the average retirement age rising by about one year.

An et al. (1999) introduced the Cox proportional hazard model to analyse the retirement behaviour of married couples in the US. Their model generalised the traditional univariate duration analysis to include a family-wide joint retirement process that induced both spouses to retire at the same time. The data was obtained from the 1969 US Retirement History Survey (RHS), which contained observations on 978 couples aged between 58 and 63. They confirmed the asymmetric effects between a husband's and a wife's income on their retirement hazards. Regarding the cross-effects of health, their empirical analysis differed from that obtained using a traditional univariate analysis. The univariate analysis suggested that the husband tended to stay longer at work if the wife was in poor health, possibly to cover the high health costs, whereas the wife tended to retire earlier if the husband was in poor health, possibly to take care of him. On the other hand, results from their model indicated symmetry across genders: both the wife's and the husband's retirement hazard was lowered when their spouse was in poor health.

4.2.2 A Model of the Retirement Decisions

The number of studies devoted to understanding when and why workers retire in Western developed countries is enormous.[3] The theoretical framework in these studies conceptualises retirement as a trade-off between work and leisure within the constraints of health and economics. For example, studies have examined how retirement decisions are influenced by individual factors, such as age, gender, race, education, and health (Diamond and Hausman, 1984), others have used family factors, such as marital status, the number of children, and family size (Chen, 1994; Chang 1999), and yet another group of researchers focused on social factors such as the pension and social security systems (Gustman and Steinmeier, 1986; Stock and Wise, 1990; Berkovec and Stern, 1991).

The general framework of this chapter follows Mitchell and Fields (1981) and

[3] For example, see, Lazear (1986), Lumsdaine and Mitchell (1999) for a survey on this area.

assumes that the individual can choose the optimal labour supply path or optimal duration in employment for their lifetime. An individual's income possibilities are constrained by lifetime earning capacity and pension opportunities (including both employer-provided benefits and social security payments). This life cycle framework leads directly to a structure in which the optimal labour supply path or employment duration (D) is a function of the lifetime streams of Earnings(E_i), Pensions(P_i), and other pertinent explanatory variables (X_i) such as Age, Gender, Race, Education, Health, Marital and Residence Status.

$$D = D\,(E_i, P_i, X_i) \qquad i = 1, 2, \cdots, n \tag{4.1}$$

where n is the number of observations. (D) represents the optimal labour supply path or employment duration and is a dependent variable. The explanatory variables in equation (4.1) must be expressed for their lifetime and not just in terms of their current levels. As mentioned in Chapter 2, the definition of retirement in Taiwan is as follows: workers who stopped working full time permanently or government employees who have received pension benefits. This definition and an employment duration model are used to examine the influences on individual retirement behaviours.

Specific hypotheses about the above factors are described as follows:

(i) Personal factors: except for retirement law, the coefficients for Age variables from ages 50 to 70 are expected to have a positive sign for retirement, because as workers become older their productivity declines and fewer job opportunities are available to them. For instance, as previously mentioned and shown in Table 3.1, the labour force participation rates declined with age from 65.3 percent by aged 50-54, to

55.1 percent by aged 55-59, 39.7 percent by aged 60-64, and 8.9 percent by aged 65 and over in 1996. This implies that as workers become older their productivity declines and they gradually lose their jobs. In addition, the marginal utility of leisure time might be an increasing function of age, so that even if productivity and wage rates do not decline with age individuals may be more likely to retire as they become older.

Further, for the Gender variable, female workers have a higher hazard rate of retirement than males. Given the tradition that females typically bear the greater share of the burden of household duties, this might imply that females place a higher valuation on the marginal unit of leisure than their male counterparts. Hence, the reservation wage of females is likely to be higher than that for males and thus faced with the same earnings opportunities in paid employment females may be more likely to retire than males.

For the Race variables, due to the differences in cultural attitudes to work, family support arrangements, and the special political situation in Taiwan, the coefficient of the Mainlander variable is expected to have a positive sign with a negative sign for other Taiwanese ethnic groups. Since most Mainlanders moving from China in 1949 have been in the army or working in the government sectors, they have a regular income and more security benefits for their retirement. Hence, they have a higher hazard rate of retirement than other Taiwanese. In contrast, other Taiwanese, being farmers, self-employed, or working in the private sectors, have an irregular income. Hence, they need to work hard and maintain family support arrangements, so they have a lower hazard rate of retirement, especially Hakka people (Shih, 1999).

For the Health variable, if a worker's health is poor, he or she tends to have reduced ability and opportunities to work, and perhaps a reduced desire or need for employment. Hence, the coefficient of poor health is expected to have a positive sign for retirement; when their health declines, workers are more likely to retire. In contrast, since workers with better education have a higher productivity and more employment opportunities, so the coefficients of education are expected to have a negative sign. The hazard rate of retirement would be lower and they are less likely to retire.

(ii) Family factors: For Marital Status, if a man is married, he has a greater responsibility and tends to work to earn money for his family. Thus, married male workers are expected to have a lower hazard rate of retirement. However, this marriage effect might be different for females; they might depend on their family and have a higher hazard rate of retirement.

(iii) Employment opportunity: For Residence Status, in urban areas employment opportunities for the middle-aged and elderly are more limited than in rural areas. One might therefore expect that states having a larger proportion of the population living in urban areas might have a higher retirement rate. On the other hand, rural workers tend to help out by feeding livestock and poultry and doing other farm chores, so there are greater self-employment opportunities and more low skilled jobs for the middle-aged and elderly. Hence, rural workers are expected to have a lower hazard rate of retirement.

(iv) Economic factors: This is a measure of the incentive to retire or not retire. Workers with higher earnings are expected to have a lower hazard rate of retirement, while workers with higher pension income are expected to have a higher hazard rate of retirement. For instance, Diamond and Hausman (1984), Gustman and Steinmeier (1986) and Stock and Wise (1990) noted that workers eligible for a pension have a higher hazard rate of retirement. However, the pension system in Taiwan only provides occupational pensions, specifically for government employees and workers in large private companies. Section 2.5 presented some conditional results about pensions, that is, if the employment duration is less than 35 years, then workers have a lower hazard rate of retirement, and if greater than 35 years, they will have a higher hazard rate of retirement. Therefore, the above hypothesis will be likely to be true if the employment duration is greater than 35 years.

4.3 Empirical Specifications

In order to examine the determinants of individual retirement behaviour, this study uses continuous-time parametric models for analysis. This method may have some differences with the results of Narendranathan and Stewart (1993) and Nolan (2000), which use discrete-time models for estimation. For instance, Narendranathan and Stewart (1993) focus on the time in weeks and unemployment durations are only observed in terms of completed whole weeks. Further, they estimate the parameters of the models only for those who were unemployed for at least four weeks. Therefore, they use the discrete-time, grouped hazard models, and binary response models to examine the determinants of unemployment duration. Moreover, Nolan (2000) employs weeks and days of length for examining absence durations. He also considers the effect of censoring of the weekend on Barmby, Orme and Treble's (1991) estimate of duration dependence in these absence spells. His focus is a grouped hazard specification for multiple-spells.

In contrast, this study specifies that employment durations are continuous to estimate the retirement hazard. In particular, the definition of length of employment durations is deducted from the answer to the question from the SHLS data: for the current employees, "When did you start your present job?" and the retirees, "When did you start your last job?" and "When did you stop doing your last job?" The employment durations can be calculated by years and months. Therefore, this study can get the precise working life durations of individuals from 1 to 55 years, with a mean of 21.876 years as shown in Table 4.3.

4.3.1 Distribution of Employment Duration

Let us start by assuming the dependent variable of interest to be continuous employment duration, the length of time to retirement from the start of work. Let the duration distribution function (or failure function) $F(t)$ represent the probability of retirement from employment by time t, defined as:[4]

$$F(t) = \Pr(T_i < t) = \int_0^t f(v)dv. \tag{4.2}$$

Basically, (4.2) specifies the probability that the random variable T_i is less than some value t. The corresponding density function is:

$$f(t) = dF(t)/dt \tag{4.3}$$

The probability of survival $S(t)$ in employment to at least time t is:

[4] See Klein and Moeschberger (1997, Chapter 2) and Lancaster (1990, p.7-10).

$$S(t) = \Pr(T_i \geq t) = 1 - F(t). \tag{4.4}$$

The basic building block in duration models is the hazard function, denoted $h(t)$, at time t. In this study, the continuous-time hazard rate represents the instantaneous retirement rate from employment at time t. The probability that an individual who has been in employment until time t retires in a short interval of length dt after t is $\Pr(t \leq T_i \leq t + dt \mid T_i \geq t)$, giving an average probability of retirement per unit of time within the short interval dt of $\Pr(t \leq T_i \leq t + dt \mid T_i \geq t)/dt$. The hazard rate, $h(t)$, is defined as the time limit of this expression:[5]

$$h(t) = \lim_{dt \to 0} \frac{\Pr(t \leq T_i \leq t + dt \mid T_i \geq t)}{dt}. \tag{4.5}$$

Applying conditional probabilities to (4.5) yields:[6]

$$\begin{aligned} h(t) &= \lim_{dt \to 0} \frac{\Pr(t \leq T_i \leq t + dt \mid T_i \geq t)}{dt} \\ &= \frac{1}{S(t)} \lim_{dt \to 0} \left\{ \frac{F(t + dt) - F(t)}{dt} \right\} \\ &= \frac{f(t)}{S(t)}. \end{aligned} \tag{4.6}$$

Changes in the hazard function over time give information about the duration dependence of an underlying stochastic process. If $dh(t)/dt > 0$, then the process

[5] In the discrete-time or grouped hazard model, t may be measured in years and durations only observed in terms of whole years completed, an observed duration of t whole years indicates a duration on the continuous time scale of t and $t+1$ years. Then, the discrete-time hazard function can be rewritten as $h(t) = P(T_i < t + 1 \mid t \leq T_i)$. See, Narendranathan and Stewart (1993).

[6] See Fleming and Harrington (1991, p.3).

exhibits positive duration dependence, the hazard rate increases over the time of duration in employment. If $dh(t)/dt < 0$, then the process exhibits negative duration dependence, the hazard rate decreases over the time of duration in employment.

4.3.2 Estimation of Hazard Function

Previous studies have focused on the use of either non-parametric or semi-parametric approaches for the analysis of retirement behaviour (Diamond and Hausman, 1984; Antolin and Scarpetta, 1998; An et al, 1999). However, a parametric model has some advantages over a nonparametric approach, so this chapter uses both parametric and semi-parametric approaches to estimate the hazard rate of retirement.

First of all, parametric models make explicit assumptions about the distribution of the hazard or survival functions, so full-information maximum likelihood (FIML)[7] can be used to estimate the relevant parameters. Furthermore, a semi-parametric approach analyses duration data where no parametric form of the hazard or survival functions is specified. The effects of covariates are parameterised to alter the baseline hazard or survival function (for which all covariates are equal to zero) in a certain way. For simplification, assume the values of the explanatory variables do not vary with survival time. That is, there are no time-varying covariates. This assumption, however, is changed in Chapter 5, where the effects of time-varying covariates for retirement decisions are analysed.

4.3.2.1 Parametric Approach

According to Hosmer and Lemeshow (1999), a parametric approach has the following advantages: (1) Full information maximum likelihood may be used to

[7] For an introduction to the FIML, see Hayashi (2000, Section 8.5).

estimate the parameters, (2) The fitted values from the model can provide estimates of survival time, (3) The residuals can be computed from the differences between the actual and predicted values of time, and (4) It uses a continuous distribution, rather than a discrete grouped hazard. This section presents some of the specific functional forms, including exponential and Weibull distributions.

4.3.2.1.1 Exponential Model

Exponential distributions are widely used as models for duration data. The hazard function is constant and it has no duration dependence. One of the main advantages of this approach is that it is relatively simple to calculate a constant hazard rate. For this exponential model, the hazard function of employment duration is specified as

$$h(t \mid x_i) = \lambda = e^{(\beta_0 + \beta_i x_i)}. \tag{4.7}$$

where h is the hazard rate, t is an employment duration, x_i is the explanatory variables. That is, let the retirement occurring in a continuous time purely random process of hazard function, so that the probability of an offer in the short interval t, $t+dt$ is $h(t)dt$, and let the hazard function, survival function, probability density function, failure function and expected mean employment duration are specified in Table 4.1.

4.3.2.1.2 Weibull Model

Extending the exponential distribution, the Weibull distribution has two parameters, $\lambda > 0$ and $\alpha > 0$. The hazard function is defined to be

$$h(t \mid x_i) = \alpha t^{\alpha-1} \cdot \lambda = \alpha t^{\alpha-1} \cdot e^{(\beta_0 + \beta_i x_i)}. \tag{4.8}$$

Empirically, the parameters λ and α in the exponential distribution and Weibull distribution can be estimated by maximum likelihood. The parameter λ depends on the explanatory variables x_i, thus providing us with a more flexible hazard function. For example, the hazard function is increasing if $\alpha > 1$, decreasing if $\alpha < 1$, and constant if $\alpha = 1$. The last case corresponds exactly to the exponential distribution. The survival function, density function and distribution function are specified in Table 4.1. For observed duration data, $t_1, t_2, \ldots, t_n$ the log-likelihood function can be formulated and maximized to include censored and uncensored observations. Combining these duration models into a general parametric likelihood yields:

$$L(\beta) = \prod_{i=1}^{n} \left\{ \left[f(t_i | x_i, \beta) \right]^{c_i} * \left[S(t_i | x_i, \beta) \right]^{1-c_i} \right\}. \tag{4.9}$$

where $\beta = (\lambda, \alpha)$, and $c_i = 1$ represents uncensored observations, $c_i = 0$ represents right-censored observations (Cleves, Gould, and Gutierrez, 2004). To obtain the maximum likelihood with respect to the parameters of interest, β, then maximise the log-likelihood function:[8]

$$\ln L(\beta) = \sum_{i=1}^{n} \left\{ c_i \ln\left[f(t_i | x_i, \beta) \right] + (1 - c_i) \ln\left[S(t_i | x_i, \beta) \right] \right\}. \tag{4.10}$$

The procedure to obtain the values of maximum likelihood estimation requires taking

[8] Since the log function is monotone, maxima of (4.9) and (4.10) occur at the same value of β; however, maximizing (4.10) is computationally simpler than maximizing (4.9).

derivatives of $\ln L(\beta)$ with respect to β, the unknown parameters, setting these equations equal to zero, and solving for β.[9]

Table 4.1 Hazard and Survival Functions

	Exponential Model	*Weibull Model*
Hazard Function	$h(t \mid x_i) = \lambda$ $= e^{(\beta_0+\beta_i x_i)}$	$h(t \mid x_i) = \alpha\, t^{\alpha-1}\lambda$ $= \alpha\, t^{\alpha-1} \cdot e^{(\beta_0+\beta_i x_i)}$
Survival Function	$S(t \mid x_i) = \exp(-\lambda t)$ $= \exp\left[-e^{(\beta_0+\beta_i x_i)}\, t\right].$	$S(t) = \exp\left[-(\lambda t)^{\alpha}\right]$
Probability Density Function	$f(t) = h(t)S(t)$ $= \lambda \cdot \exp(-\lambda t)$	$f(t) = h(t) \cdot S(t)$ $= \alpha t^{\alpha-1}\lambda \cdot \exp\left[-(\lambda t)^{\alpha}\right]$
Failure Function	$F(t) = 1 - S(t)$ $= 1 - \exp(-\lambda t)$	$F(t) = 1 - S(t)$ $= 1 - \exp\left[-(\lambda t)^{\alpha}\right]$
Mean $E(t)$	$E(t) = \frac{1}{\lambda}$	$E(t) = \frac{\Gamma(1+1/\alpha)}{\lambda^{1/\alpha}}$

Source: Klein and Moeschberger (1997), and Hosmer and Lemeshow (1999).

[9] See Klein and Moeschberger (1997, Appendixes A and B), for a description of the numerical methods for implementing multivariate Newton-Raphson methods.

4.3.2.1.3 Weibull Model with Unobserved Heterogeneity

It is well known that duration analysis produces incorrect results if unobserved heterogeneity is ignored (Lancaster, 1990). However, most previous literature on individual retirement issues by duration analysis only focused on the influencing factors of socio-demographic status, health, marital status, pension system, and economic status; but never considered the effect of unobserved heterogeneity on estimated hazard rate of retirement (see, for example, Diamond and Hausman, 1984; Antolin and Scarpetta, 1998; and An et al., 1999). Therefore, this chapter uses the Weibull model and considers unobserved heterogeneity to fill this gap (see, Cleves, et al, 2004; Collett, 1994; Hosmer and Lemeshow, 1999).

According to Cleves et al. (2004), the unobserved heterogeneity models include unshared frailty and shared frailty models. The former are referred to as an over-dispersion/heterogeneity model, and the latter is a random-effects model where the frailties are common among groups of individuals or spells and are randomly distributed across groups (Gutierrez, 2002). Suppose the SHLS data belong to a random sample, and a "shared frailty" component is included in the model. Then, the shared frailty hazard function in the Weibull model may be written as:

$$h(t_{ji} \mid x_{ji},\, u_j) = u_j h(t_{ji} \mid x_{ji}). \tag{4.11}$$

for data consisting of n groups with the jth group comprised of n_j observations. The index j denotes the group ($j = 1,\ldots,n$) and i denotes the observation within group, $i = 1,\ldots,n_j$. The frailties, u_j, are shared within each group and are assumed to follow either a gamma or inverse-Gaussian distribution. The frailty variance, θ, is estimated from the data and measures the variability of the frailty across groups.

For instance, for men and women eligible for a pension, not only might their observed employment duration be different, but under this model their individual retirement hazard functions could also be different. As previously mentioned in Chapter 2, for the sample as a whole, workers eligible for a pension have a higher survival rate of employment duration before completing 35 years in work and a lower survival rate of employment duration after working for 35 years. However, male workers eligible for a pension generally have a lower survival rate of employment duration after working for 17 years, whereas women have a higher survival rate of employment duration before completing 47 years in work and a lower survival rate of employment duration after working for 47 years. Hence, men may be more "frail" than women due to unobserved factors accounting for individual level differences in their retirement hazard functions. These unobserved factors may contribute an extra layer of heterogeneity, leading to greater variability in the time of employment duration than might be expected under the model without the frailty component.

4.3.2.2 Semi-Parametric Approach

Based on Klein and Moeschberger (1997), this chapter is also concerned with comparing two or more groups of continuous-time of employment duration. If the groups are similar, the nonparametric method described in Chapter 2 may be used directly. More often than not, the subjects in the groups have some additional characteristics, such as age, gender, and other socioeconomic status that may affect their hazard rate of retirement. Hence, the semi-parametric approach can be used to estimate continuous-time of employment duration between groups being less biased and more precise than a simple comparison.

According to Cox (1972), the semi-parametric approach quantifies the relationship between employment duration and a set of explanatory variables, often called the proportional hazard model. In particular, this model does not need to make any assumptions about the shape of the baseline hazard function. Therefore, the continuous-time hazard function can be parameterised as[10]

$$h_i(t \mid x_i) = h_0(t)\exp(\beta' x_i). \tag{4.12}$$

where $h_0(t)$ is the baseline hazard function depending on t and not x_i, β' is the parameter vector, (the prime mark (') denotes transposition) and x_i is the time-constant covariate vector. Simply, $h_0(t)$ summarises the pattern of "duration dependence" common to all persons, $\exp(\beta' x_i)$ is the relative hazard function and a non-negative function of covariates x_i, which does not depend on t by construction. Hence, in the Cox model, $h_0(t)$ is simply left unparameterised, and through conditioning on failure times, estimates of β' are obtained anyway.

In general, the integrated continuous-time hazard function in a Cox proportional hazard model is produced by

$$\begin{aligned} H(t \mid x_i) &= \int_0^t h(v, x_i)\,dv \\ &= \exp(\beta' x_i)\int_0^t h_0(v)\,dv \\ &= \exp(\beta' x_i) H_0(t) \end{aligned} \tag{4.13}$$

---

[10] If time is measured in terms of whole years completely, the discrete-time hazard function can be rewritten as $h(t) = P(T_i < t+1 \mid t \le T_i) = 1 - \exp\left[-\int_t^{t+1} h_0(t)\exp(\beta' x_i)dt\right]$. See, Narendranathan and Stewart (1993).

As with the integrated hazard, the baseline survival function $S_0(t)$ is the survival function evaluated with all the covariates equal to zero. So the survival function can be derived as follows

$$\begin{aligned} S(t \mid x_i) &= \exp\{-H(t \mid x_i)\} \\ &= \exp\{-\exp(\beta' x_i) H_0(t)\} \\ &= S_0(t)^{\exp(\beta' x_i)} \end{aligned} \tag{4.14}$$

Cox (1972) proposed a method for estimating β without having to specify any functional form for $h_0(t)$, and this method uses partial likelihood.[11] In general, the Cox partial likelihood function is

$$L(\beta) = \prod_{i=1}^{k} \left[\frac{\exp(\beta' x_i)}{\sum_{j \in R(t_{(i)})} \exp(\beta' x_j)} \right]. \tag{4.15}$$

where k represents observed failure times and $R(t_{(i)})$ is the set of all observations still under study at the time just prior to $t_{(i)}$. Treat (4.15) as the usual likelihood function, and maximise it to give an estimate of β' that is asymptotically normal with mean β' and variance-covariance matrix equal to the inverse of the negative Hessian.[12]

[11] Partial likelihood works in terms of the ordering of events by contrast with the focus in maximum likelihood on spells. See Klein and Moeschberger (1997), pp.231-234.

[12] See Greene (2000) for an introduction to ML estimation and the necessary conditions that give maximum likelihood estimates.

4.4 Data Description

4.4.1 Data Source

The data are from the Survey of Health and Living Status of the Middle Aged and Elderly in Taiwan (SHLS), a joint survey conducted by the Taiwan Provincial Institute of Family Planning and the Population Studies Centre, University of Michigan. The total sample has 2462 observations aged 50 to 70, and their spouses. The survey questionnaire contains eight distinct sections: (i) Background information, marital status, and living situation; (ii) Family structure, general circumstances, and living with kin; (iii) Health, use of medical services, and hygiene habits; (iv) Social support and exchange of support; (v) Employment history; (vi) Leisure, activities, and general attitudes; (vii) Economic status; (viii) Livelihood plans. The SHLS survey data are fairly comprehensive and thus allow for a detailed discussion of the retirement decisions of the middle-aged and elderly in Taiwan.

4.4.2 Variables Specifications

4.4.2.1 Dependent Variable

According to the SHLS data, the sample consists of two groups, namely the current workers and retirees. The former group did not retire during the sample period and are known as "right-censored" spells. The latter group retired during the sample period, and the date on which an individual started their last job and the exact age at which they retired were observed. These are known as the "uncensored" spells. Therefore, employment duration includes the date when an individual started working to when they completely retired for "uncensored" spells, and they continue working for "right-censored" spells. This variable can be categorized as a dependent variable. The uncensored variable is coded 1 for retirement and 0 otherwise.

4.4.2.2 Explanatory Variables

The explanatory variables recorded in the SHLS data include (1) Personal factors: age, gender, race, educational attainment, marital status, health status, and residence status. (2) Economic factors: eligibility for a pension, predicted earnings as an indicator of income from work and predicted pension income as an indicator of income during retirement. The details of the above variables are described below.

First, age is explored. As noted in Chapter 2, the effect of ageing alone is important in explaining why people retire. From the SHLS data, age can be categorised in four groups: Age1 (aged 50 to 54), Age2 (aged 55 to 59), Age3 (aged 60 to 64), and Age4 (aged 65 to 70). This corresponds to the normal retirement ages at 50, 55, 60, or 65 years old in the Labour Standard Law in Taiwan.[13] Second, the Gender variable is coded 1 for female and 0 for male. Third, the Race variable has four groups, namely Race1 (Fujianese), Race2 (Hakka), Race3 (Mainlander), and Race4 (Aboriginal). Fourth, for the Education variable, Gordon and Blinder (1980) suggested that people who acquire more schooling remain in the labour force longer to recoup the costs of their education investments. Holding other variables equal, the more educated people are, the less prone they are to retire. The education variable is divided into four levels of schooling, namely Edu1 (informal schooling), Edu2 (primary level: 1 to 6 years), Edu3 (high school level: 7 to 12 years) and Edu4 (college level: 13 to 17 years). Fifth, marital status includes married, single, divorced, separated, and widowed. The Married variable is coded 1 for married and 0 for otherwise. Sixth, for health assessment, the SHLS survey identifies five levels including excellent, good, average, not so good, and poor. The Health variable is

[13] According to the Labour Standard Law (Chapter 6 Retirement, Article 53 and 54) in Taiwan, workers can choose their retirement ages at 50, 55, 60 or 65.

coded 1 for poor health, including "not so good" and "poor" health, and 0 for otherwise. Finally, for the residence area status, Gunderson (1977) suggested that the residence factor could reflect employment opportunities and living environments for people in urban (Resid1), town (Resid2) and rural (Resid3) areas.

The economic factors include eligibility for a pension and other economic variables, but a limitation of the SHLS is the lack of data on the interviewees' wages and their assets. Predicted earnings (P-Earnings) as an indicator of income from work and predicted pension income (P-Pension) as an indicator of income during retirement are constructed to facilitate further analysis of retirement decisions. This approach is not new but it has been infrequently adopted in economic studies (see, Diamond and Hausman, 1984; Slade, 1987; Arulampalam and Stewart, 1995; and Buckley et al., 2004). However, there are likely to be some selection problems[14] at work in the equations estimated in order to generate values of predicted earnings and predicted pension income for the individuals in the sample. The eligibility for a pension dummy variable, which includes retirees who have received pension benefits and workers who expect to receive pension benefits, is coded 1 for those with benefits and 0 otherwise.

4.4.2.2.1 Predicted Earnings as Income from Work

The regression model uses 983 observations[15] on those who are currently working to predict the effective sample of 2052 observations, except for the never

[14] The selection problems arise in the context of the regression equations for predicted earnings and predicted pension income because the individuals for whom we observe earnings and pension income are drawn non-randomly from the overall data sample. It is not easy to incorporate an allowance for selection effects in the models. In particular, there are no exact pension income data from the questionnaires of the SHLS survey; actual pension income is assumed to be equal to major sources of income from pension, retirement fund and insurance.

[15] According to the questionnaire of 1996 SHLS (E4a: how much did you earn last year?), the effective sample included 1072 observations working full-time and 124 observations working part-time. Only 983 observations presented their actual earnings, 161 observations didn't know or found it hard to estimate, 29 observations refused to answer, and there were 23 missing for calculation.

worked group (410 observations). The estimated model is as follows (the standard errors are in parentheses):

$$\begin{aligned}\hat{E} = {} & \underset{(5.835)}{43.929} - \underset{(2.909)}{4.117}*Age2 - \underset{(3..304)}{8.673}*Age3 - \underset{(5.953)}{13.031}*Age4 - \underset{(2.969)}{15.176}*Gender \\ & - \underset{(3.185)}{4.328}*Race2 - \underset{(5.312)}{15.799}*Race3 - \underset{(9.157)}{1.211}*Race4 - \underset{(3.555)}{1.391}*Edu2 \\ & + \underset{(4.254)}{19.051}*Edu3 + \underset{(5.616)}{51.052}*Edu4 + \underset{(3.801)}{5.369}*Married - \underset{(3.593)}{2.696}*Health \\ & + \underset{(2.958)}{2.442}*Pension - \underset{(3.238)}{7.639}*Town - \underset{(3.044)}{13.925}*Rural\end{aligned} \tag{4.15}$$

$$F(15, 967) = 22.67 \qquad R^2 = 0.260$$

It is clear from the $F-statistic$ that we strongly reject the null hypothesis that all the regression coefficients except the constant term are zero. Comparing this with the standard errors presented above, we note that Age2, Race2, Race4, Edu2, Married, Health, and Pension are individually insignificant. Therefore, if the restricted model only considers a significant effect on actual earnings, the resulting equation is as follows:

$$\begin{aligned}\hat{E} = {} & \underset{(2.707)}{45.598} - \underset{(3.061)}{7.349}*Age3 - \underset{(5.724)}{11.610}*Age4 - \underset{(2.712)}{15.626}*Gender \\ & - \underset{(5.089)}{15.133}*Race3 + \underset{(3.038)}{21.139}*Edu3 + \underset{(4.468)}{53.773}*Edu4 \\ & - \underset{(3.209)}{7.784}*Town - \underset{(2.953)}{15.064}*Rural\end{aligned} \qquad (4.15)'$$

$$F(8, 974) = 41.58 \qquad R^2 = 0.255$$

The coefficients of Age3 (aged 60 to 64), Age4 (aged 65 to 70), Gender, Race3 (Mainlander), Town, and Rural variables are found to have a negative value and a statistically significant effect on predicted earnings in Table 4.2.1. Specifically, older workers, female workers, Mainlander workers, and workers living in town or rural areas have lower predicted earnings. In contrast, the coefficients of the Edu3 (7 to 12 years schooling) and Edu4 (13 to 17 years schooling) variables have a positive value and a statistically significant effect on predicted earnings. This means that workers with better education have higher predicted earnings. In general, workers with better education have higher productivity and better job prospects available to them in the government sector or large companies than people with lower education. Consequently, they would have higher earnings than others. The F value and R^2 value are 41.58 and 0.255 respectively.[16] The average actual earnings of current workers are NT$408,410 (equal to £9,523),[17] and the average predicted earnings are NT$355,370 (equal to £8,287) – the latter are about 13.0% lower than the former as shown in Table 4.3. The average predicted earnings are lower than actual earnings because they took into account the proportion of people who retired or became unemployed, and who might have a lower income.

4.4.2.2.2 *Predicted Pension Income as Income during Retirement*

Due to the lack of exact pension income data from the questionnaires of the SHLS survey, actual pension income is assumed to be equal to major sources of income from pensions, retirement funds and insurance. The survey contained 222 of these cases. These individuals' incomes can be used to predict pension income for the effective sample, except for the never worked group.

[16] Predicted earnings of the effective sub-sample were only based on the particular values of the independent variables in equation (4.15)'. Some insignificant variables such as Age2, Race4, Edu2, Married, Health, and Pension in equation (4.15)' were dropped. See Slade (1982, p.9).

[17] The rate of foreign exchange was about NT$42.883 equals £1 in 1996. The website is (in Chinese): http://investintaiwan.nat.gov.tw/zh-tw/env/stats/exchange_rates.html.

$$\hat{P} = \underset{(8.026)}{16.309} - \underset{(6.071)}{4.659^*} Age2 - \underset{(5.830)}{6.596^*} Age3 - \underset{(5.681)}{10.073^*} Age4 + \underset{(4.716)}{4.642^*} Gender$$
$$+ \underset{(5.177)}{1.644^*} Race2 + \underset{(4.374)}{2.342^*} Race3 - \underset{(9.409)}{7.834^*} Race4 + \underset{(4.228)}{.790^*} Edu2$$
$$+ \underset{(4.777)}{16.800^*} Edu3 + \underset{(6.719)}{35.508^*} Edu4 + \underset{(3.979)}{11.653^*} Married - \underset{(3.625)}{5.144^*} Health \qquad (4.16)$$
$$+ \underset{(4.665)}{9.631^*} Pension - \underset{(3.238)}{4.718^*} Town + \underset{(8.026)}{16.309^*} Rural$$

$$F(15, 206) = 6.76 \qquad R^2 = 0.330$$

Further, for comparison of the effective sub-sample, this thesis uses the same set of explanatory variables in the equation for predicted earnings and predicted pension income. The estimated model is as follows (the standard errors are in parentheses):

$$\hat{P} = \underset{(5.175)}{27.964} - \underset{(4.451)}{6.371^*} Age3 - \underset{(4.341)}{9.425^*} Age4 - \underset{(4.124)}{0.261^*} Gender$$
$$+ \underset{(4.162)}{2.677^*} Race3 + \underset{(3.791)}{20.805^*} Edu3 + \underset{(5.980)}{41.393^*} Edu4$$
$$- \underset{(3.915)}{2.244^*} Town - \underset{(3.928)}{0.158^*} Rural \qquad (4.16)'$$

$$F(8, 213) = 10.29 \qquad R^2 = 0.279$$

The coefficients of Edu3 (7 to 12 years schooling), and Edu4 (13 to 17 years schooling) variables are found to have a significant positive effect on average predicted pension income as shown in Table 4.2.2. That means workers with better education have a higher predicted pension income. Workers with better education are more likely to work in the government sectors and large private companies and thus be eligible for pensions. Therefore, educated people have a higher predicted pension income. In contrast, the coefficient of the Age4 (aged 65 to 70) variable has a significantly negative effect on average predicted pension income, which implies that elderly workers have a lower predicted pension income. Finally, the F value and R^2 value are 10.29 and 0.279 respectively. The average actual pension income is found to be NT$339,510 (equal to £7,917), and the average predicted pension income is NT$319,930 (equal to £7,461) – the latter is about 5.8% lower than the former from Table 4.3. The average predicted pension income is lower than actual pension income because it took into account the proportion of current workers who might not receive their pension benefits.

4.4.3 Summary Statistics

This chapter uses the data and variables described above to analyse retirement decisions in Taiwan. A full definition of the variables and summary statistics of the sample are given in Table 4.3.

Table 4.2.1 Estimated Results for Predicted Earnings

Model	Distribution of Sub-sample		General Regression		Restricted Regression for Earnings	
Variables	Mean	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age1	.393	(.489)	-		-	
Age2	.343	(.475)	-4.117	(2. 909)	-	
Age3	.208	(.406)	-8.673**	(3.304)	-7.349**	(3.061)
Age4	.056	(.230)	-13.031**	(5.953)	-11.610**	(5.724)
Gender	.300	(.456)	-15.176***	(2.969)	-15.626***	(2.712)
Race1	.716	(.451)	-		-	
Race2	.191	(.393)	-4.328	(3.185)	-	
Race3	.076	(.266)	-15.799***	(5.312)	-15.132***	(5.089)
Race4	.016	(.127)	-1.211	(9.157)	-	
Edu1	.195	(.397)	-		-	
Edu2	.467	(.499)	-1.391	(3.555)	-	
Edu3	.244	(.430)	19.051***	(4.254)	21.138***	(3.038)
Edu4	.094	(.291)	51.052***	(5.616)	53.773***	(4.468)
Married	.874	(.332)	5.369	(3.801)	-	
Health	.136	(.343)	-2.696	(3.593)	-	
Pension	.322	(.468)	2.442	(2.958)	-	
Urban	.385	(.487)	-		-	
Town	.242	(.428)	-7.639**	(3.238)	-7.784***	(3.209)
Rural	.373	(.484)	-13.925***	(3.044)	-15.064***	(2.953)
Constant			43.929***	(5.835)	45.598***	(2.706)
N	983		983		983	
F (15, 967)			22.67		41.58	
Prob > F			0.000		0.000	
R-squared			0.260		0.255	

Notes:

1. The F distribution of restricted regression for earnings is F (8, 974).

2. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

Table 4.2.2 Estimated Results for Predicted Pension Income

Model	Distribution of Sub-sample		General Regression		Restricted Regression for Pension Income	
Variables	Mean	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age1	.117	(.322)	-		-	
Age2	.162	(.369)	-4.659	(6.071)	-	
Age3	.302	(.460)	-6.596	(5.830)	-6.371	(4.451)
Age4	.419	(.494)	-10.073*	(5.681)	-9.425**	(4.341)
Gender	.468	(.500)	4.642	(4.716)	-.261	(4.124)
Race1	.513	(.501)	-		-	
Race2	.122	(.328)	1.644	(5.177)	-	
Race3	.333	(.472)	2.342	(4.374)	2.677	(4.162)
Race4	.032	(.175)	-7.834	(9.409)	-	
Edu1	.284	(.452)	-		-	
Edu2	.320	(.467)	.790	(4.228)	-	
Edu3	.302	(.460)	16.800***	(4.777)	20.805***	(3.791)
Edu4	.095	(.293)	35.508***	(6.719)	41.393***	(5.980)
Married	.775	(.419)	11.653***	(3.979)	-	
Health	.288	(.454)	-5.144	(3.625)	-	
Pension	.504	(.501)	9.631**	(4.665)	-	
Urban	.441	(.498)	-		-	
Town	.279	(.449)	-4.718	(3.238)	-2.244	(3.915)
Rural	.280	(.450)	1.317	(3.044)	-.158	(3.928)
Constant			16.309**	(8.026)	27.964***	(5.175)
N	222		222		222	
F (15, 206)			6.76		10.29	
Prob > F			0.000		0.000	
R-squared			0.330		0.279	

Notes:

1. The F distribution of restricted regression for pension income is F (8, 213).

2. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

Table 4.3 Descriptive Statistics of Variables

Variables	Description	Mean	Std Err
DURATION	1-55 Years.	21.876	(14.079)
CENSOR	1 = Uncensored (retirement), 0 = Otherwise.	.352	(.478)
AGE1	1= Aged 50 to 54, 0 = Otherwise.	.312	(.463)
AGE2	1 = Aged 55 to 59, 0 = Otherwise.	.319	(.466)
AGE3	1 = Aged 60 to 64, 0 = Otherwise.	.255	(.436)
AGE4	1 = Aged 65 to 70, 0 = Otherwise.	.114	(.318)
GENDER	1 = Female, 0 = Male.	.389	(.487)
RACE1	1 = Fujianese, 0 = Otherwise.	.724	(.447)
RACE2	1 = Hakka, 0 = Otherwise.	.174	(.380)
RACE3	1 = Mainlander, 0 = Otherwise.	.085	(.280)
RACE4	1 = Aboriginal, 0 = Otherwise.	.016	(.126)
EDU1	1 = Informal schooling, 0 = Otherwise.	.249	(.433)
EDU2	1 = 1 to 6 years of schooling, 0 = Otherwise.	.473	(.499)
EDU3	1 = 7 to 12 years of schooling, 0 = Otherwise.	.209	(.407)
EDU4	1 = 13 to 17 years of schooling, 0 = Otherwise.	.069	(.253)
MARRIED	1 = Married, 0 = Otherwise.	.841	(.366)

Table 4.3 (Continued)

HEALTH	1 = Poor health, 0 = Otherwise.	.231	(.421)
PENSION	1 = Eligible for a pension, 0 = Otherwise.	.267	(.443)
URBAN	1 = Live in urban areas, 0 = Otherwise.	.380	(.486)
TOWN	1 = Live in town areas, 0 = Otherwise.	.237	(.426)
RURAL	1 = Live in rural areas, 0 = Otherwise.	.383	(.486)
EARNINGS	Average actual earnings from current workers.	40.841	(43.344)
P-EARNINGS	Average predicted earnings for working-time income.	35.537	(20.789)
PENSION INCOME	Average actual pension income from pension, retirement fund and insurance.	33.951	(29.964)
P-PENSION	Average predicted pension income for retirement-time income.	31.993	(13.404)

Note:

1. The effective sample of duration model only has 1732 observations, including 610 retirees (complete observations) and 1122 continuing work (right-censored observations).

2. The units of EARNINGS, P-EARNINGS, PENSIONS, and P-PENSION are NT$10,000. The rate of foreign exchange was about NT$42.883 equals £1 in 1996.

3. However, the variables of Race3 and Edu4 are dropped due to collinearity with predicted variables in the models.

4.5 Empirical Results

The hazard rates of retirement are estimated by continuous-time parametric models, including Exponential, Weibull, and Cox proportional hazard approaches. All models are estimated for the 1996 SHLS survey, considering two cases in each of the approaches. The gender effects and unobserved heterogeneity on retirement decisions are also tested. The details are described as follows.

4.5.1 Exponential Models

The empirical results of the Exponential models, tabulated in Tables 4.4.1 and 4.4.2, present the maximum likelihood estimates (hereafter, MLE) of the parameters. Simply, the estimated coefficient β can reflect the effect on retirement hazard. If $\beta > 0$, the retirement hazard increases. If $\beta < 0$, the retirement hazard decreases. If $\beta = 0$, then there is no effect on retirement hazard. In particular, the retirement hazard is constant and it has no duration dependence.

4.5.1.1 Case 1: Without Predicted Earnings and Predicted Pension Income

Table 4.4.1 presents the estimated coefficients (with standard errors in parentheses) for Exponential model specifications, the case without predicted earnings and predicted pension income variables, and the dependent variable being continuous-time of employment duration. Based on these estimates, the hazard rate of retirement can be calculated for the benchmark individual and for other individuals with different demographic circumstances.

For the benchmark individual, all explanatory variables take a value of zero. That is, the benchmark individual in all cases is an unmarried Fujianese man aged 50 to 54, who is in good health, is not eligible for a pension, and who lives in an urban area.

This benchmark case is reflected by the constant term in the estimation. For example, in Table 4.4.1, the benchmark estimates lead to the hazard rate of retirement estimates in the exponential model of

$$h(t;x_i) = \lambda = e^{(\hat{\beta}_0+\hat{\beta}_1\cdot 0)} = e^{\hat{\beta}_0} = \exp(-4.771) = 0.008.$$

The effects on retirement hazard can be calculated for different demographic circumstances. How, for example, other things being equal, does the retirement hazard change for workers in Age2 (aged 55 to 59), Age3 (aged 60 to 64), and Age4 (aged 65 to 70)? The first situation presents the hazard rate of retirement changes for workers in Age2 (aged 55 to 59):

$$h(t;Age2) = \lambda = e^{(\hat{\beta}_0+\hat{\beta}_i\cdot 1)} = \exp(-4.771+0.307) = 0.012.$$

The second situation presents the hazard rate of retirement changes for workers in Age3 (aged 60 to 64):

$$h(t;Age3) = \lambda = e^{(\hat{\beta}_0+\hat{\beta}_i\cdot 1)} = \exp(-4.771+0.591) = 0.015.$$

The third situation presents the hazard rate of retirement changes for workers in Age4 (aged 65 to 70), other variables being constant:

$$h(t;Age4) = \lambda = e^{(\hat{\beta}_0+\hat{\beta}_i\cdot 1)} = \exp(-4.771+0.834) = 0.020.$$

Therefore, if the coefficient, $\beta > 0$, the hazard rate of retirement will increase and be higher. So, older workers have a higher hazard rate of retirement than benchmark workers.

In general, the coefficient of the Gender variable is significantly positive (at the 1 percent level of significance). This means that female workers have a higher hazard rate of retirement. The variable Race3 (Mainlander) has a significantly positive effect on the retirement hazard, but the coefficients for Race2 (Hakka) and Race4

(Aboriginal) variables are negative and insignificant. This implies that Mainlander workers have a higher hazard rate of retirement, particularly if they retired from the army and government sectors, which might offer a number of pension, retirement fund and insurance benefits. In contrast, Hakka and Aboriginal workers have a lower hazard rate of retirement because they might have fewer pension benefits and need to work longer. These are consistent with the results reported in Shih (1999).

Moreover, the coefficients for Edu3 (7 to 12 years of schooling) and Edu4 (13 to 17 years of schooling) variables are significantly negative. This means that workers with better education have a lower hazard rate of retirement and are less likely to retire. This is similar to the results reported in Zimmer and Liu (1999), and Chang (1999). Similarly, the Married variable has the expected negative sign, but only significant at the 10% level. This means that married workers might have a higher economic burden from their family and lower hazard rate of retirement. Hence, they are more likely to retire later.

Table 4.4.1 also notes that the Health variable is statistically significant in the model, in keeping with a priori expectations of a positive effect on retirement hazard. An explanation for workers with poor health is that this might reduce their productivity, time and ability to work and, consequently, they might quit their job to improve their health. Hence, workers with poor health have a higher hazard rate of retirement. This finding is entirely consistent with the results reported in Diamond and Hausman (1984), An et al. (1999), and Mete and Schultz (2002). For instance, Diamond and Hausman (1984) used the NLS survey to examine the determinants of individual retirement and savings behaviour and found that the demographic variable with by far the largest effect is bad health.

For the economic factors, the variable Pension has a positive effect on retirement behaviour, which implies that workers eligible for an occupational pension have a higher hazard rate of retirement, but the coefficient is insignificant. Finally, for Residence Status, the variable Rural is negatively significant. Rural workers have a lower hazard rate of retirement. This is because most rural workers are engaged in agricultural work or are self-employed. Their earning capacities and opportunities are relatively lower and more spasmodic, and so they need to work longer and retire later.

A further analysis for gender effects is estimated in Table 4.4.1. Some estimated effects of male and female retirement decisions are similar, such as Age groups and Health variables. Older people and workers with poor health have a higher hazard rate and are more likely to retire. However, other influencing factors of retirement have a different effect for men and women. For instance, the Race2, Edu3, and Married variables for men show a significantly negative effect on retirement duration, but insignificant for women. This implies that male Hakka workers, males with high school education and married male workers have a lower hazard rate of retirement. Further, the Town and Rural variables for women have a significantly negative effect on retirement duration, but insignificant for men. This means that female workers living in town and rural areas have a lower hazard rate of retirement. Furthermore, the Pension variable for men has a significantly positive effect on retirement, but negative effect for women. That is, males with eligible pension have a higher hazard rate, but females with eligible pension have a lower hazard rate. In particular, actual sample proportions from the 1996 SHLS data, 34.4% of male workers are eligible for a pension, but for females only 14.8% have a pension. Thus, the incentives of pension for women are stronger than men.

Table 4.4.1 Exponential Models without Predicted Variables

Sample	Overall		Male		Female	
Variables	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.307**	(.130)	.408*	(.241)	.364**	(.158)
Age3	.591***	(.126)	.989***	(.225)	.400**	(.159)
Age4	.834***	(.141)	1.334***	(.242)	.471**	(.190)
Gender	.977***	(.091)	-		-	
Race2	-.181	(.119)	-.419**	(.203)	-.041	(.148)
Race3	.596***	(.149)	.369*	(.193)	-.011	(.321)
Race4	-.123	(.311)	-.385	(.724)	-.150	(.349)
Edu2	-.002	(.099)	-.232	(.162)	.136	(.125)
Edu3	-.267*	(.144)	-.441**	(.203)	.116	(.213)
Edu4	-.486**	(.239)	-.479	(.296)	-.280	(.452)
Married	-.169*	(.099)	-.371**	(.161)	-.006	(.126)
Health	.522***	(.086)	.677***	(.131)	.319***	(.114)
Pension	.011	(.115)	.259*	(.154)	-.652***	(.211)
Town	-.020	(.106)	.191	(.155)	-.286*	(.148)
Rural	-.473***	(.103)	-.225	(.156)	-.712***	(.136)
Constant	-4.771***	(.186)	-4.979***	(.295)	-3.596***	(.197)
No. of subjects	1732		1059		673	
No. of retirees	610		271		339	
Log likelihood	-1292.152		-592.035		-672.271	
LR chi2 (15)	311.58***		165.93***		59.48***	

Notes:

1. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

4.5.1.2 Case 2: With Predicted Earnings and Predicted Pension Income

The main limitation of the SHLS is the lack of data on the interviewees' wages and their pension benefits. This section follows Slade (1987) to construct predicted earnings and predicted pension income variables in equation (4.15) and equation (4.16) respectively. Then these two predicted variables can be added into the hazard function (4.7) to observe how lifetime income affects individual retirement decisions as shown in Table 4.4.2.

Table 4.4.2 shows that most estimated coefficients are similar to those estimated by the model without predicted variables in Table 4.4.1. For instance, the estimated coefficients of Age2, Age3, Age4, Gender, and Health variables have a significantly positive effect on retirement duration, and the Married, Town, and Rural variables have a significant negative effect. This implies that older workers, female workers, and workers with poor health have a higher hazard rate of retirement; and married workers, workers who live in town and rural areas have a lower hazard rate of retirement. However, the Race3 and Edu4 variables are dropped due to collinearity, since they might have some relationship when calculating predicted earnings and predicted pension income variables.

In particular, Table 4.4.2 highlights that the P-Earnings variable is statistically significant in the models, in keeping with a priori expectations of a negative effect on retirement hazard. An explanation is that higher predicted earnings might induce people to work longer and earn more; consequently, they are more likely to continue working and have a lower hazard rate of retirement. This is consistent with the results reported in Slade (1987), who used data from the Longitudinal Retirement History Study (LRHS) to examine the determinants of retirement status and state dependence in the US, and found that predicted earnings had a significant negative effect on labour force exit and change in predicted earnings had a significant negative effect. Furthermore, Arulampalam and Stewart (1995) also used a predicted earnings variable to examine the determinants of individual unemployment duration. Using data from the Department of Health and Social Security (DHSS) in the UK, they found that the probability of entering full-time work falls with age, and increases with predicted earnings in employment. In particular, they demonstrated a more generalised result that the younger groups (aged under 20 and 25 to 34) are likely to have increasing earnings, and the middle-aged and old groups (ages 35 to 44, 45 to 54, and 55 to 64) are likely to have decreasing earnings. Consequently, the estimated result of Taiwanese older workers with higher predicted earnings face two opposing effects on retirement hazard, since age has a positive link and predicted earnings have a negative link.

On the other hand, Table 4.4.2 also indicates that the P-Pension variable has a significant positive effect on retirement hazard. This result highlights that workers with higher predicted pension incomes have a higher hazard rate of retirement and are more likely to retire. This is consistent with the result reported in Diamond and Hausman (1984). However, they focused on social security benefits such as public or state pensions rather than occupational pensions, which workers continue receiving after 65. Workers with higher predicted pension income had a higher hazard rate of retirement. Reduction in pension income caused by early retirement had a small effect. In particular, the predicted pension income in Taiwan is only based on occupational pensions, and only about 25% of retirees receive this pension benefit at their retirement age. However, this is somewhat different from the results reported in Slade (1987), who shows that the present values of social security benefits have a significant negative effect on labour force exit, but the change in social security benefits was insignificant. This implied that a decrease in benefits would not discourage labour force exit by workers, but both effects were small. Hence, workers with higher predicted pension income have better economic status and they are more likely to retire earlier.

Finally, the gender effects are estimated in Table 4.4.2. Most estimated effects are similar to those estimated by the model without predicted variables as shown in Table 4.4.1. For both men and women, older people and workers with poor health have a higher hazard rate and are more likely to retire, but rural workers have a lower hazard rate and are less likely to retire. In addition, there are some different results for gender. For instance, male Hakka workers, married male workers and workers with higher predicted earnings have a significantly lower hazard rate of retirement, but this is insignificant for females. The Pension variable for men has a significantly positive effect on retirement duration, but a negative effect for women. Furthermore, male and female workers with higher predicted earnings have a lower hazard rate of retirement and are less likely to retire. In the meantime, male workers with higher predicted pension incomes have a higher hazard rate of retirement and are more likely to retire, and this shows a negative effect for women, but all insignificant for males and females.

Table 4.4.2 Exponential Models with Predicted Variables

Sample	Overall		Male		Female	
Variables	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.307**	(.130)	.408*	(.241)	.364**	(.158)
Age3	.547***	(.132)	.936***	(.231)	.358**	(.175)
Age4	.745***	(.154)	1.238***	(.255)	.408*	(.219)
Gender	.458***	(.144)	-		-	
Race2	-.181	(.119)	-.419**	(.203)	-.041	(.148)
Race3	-		-		-	
Race4	-.123	(.311)	-.385	(.724)	-.150	(.349)
Edu2	-.002	(.099)	-.232	(.162)	.136	(.125)
Edu3	-.221	(.161)	-.326	(.203)	.255	(.294)
Edu4	-		-		-	
Married	-.169*	(.099)	-.371**	(.161)	-.006	(.126)
Health	.522***	(.086)	.677***	(.131)	.319***	(.114)
Pension	.011	(.115)	.259*	(.154)	-.652***	(.211)
Town	-.211*	(.113)	.061	(.166)	-.303*	(.169)
Rural	-.957***	(.153)	-.546**	(.219)	-.719**	(.282)
P-Earnings	-.034***	(.008)	-.022**	(.010)	-.001	(.017)
P-Pension	.032***	(.011)	.016	(.015)	-.006	(.023)
Constant	-4.129***	(.283)	-4.456***	(.405)	-3.410***	(.387)
No. of subjects	1732		1059		673	
No. of retirees	610		271		339	
Log likelihood	-1292.152		-592.035		-672.271	
LR chi2 (15)	311.58***		165.93***		59.48***	

Notes:

1. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.

2. The variables of Race3 and Edu4 are dropped due to collinearity with predicted variables in the models.

3. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

4.5.2 Weibull Models

The estimations of the Weibull models are tabulated in Tables 4.5.1 and 4.5.2. The estimated coefficient β can reflect the effect on retirement hazard. If $\beta > 0$, the retirement hazard increases. If $\beta < 0$, the retirement hazard decreases. If $\beta = 0$, then there is no effect on retirement hazard. Furthermore, the hazard function increases in duration if $\alpha > 1$, decreases if $\alpha < 1$, and remains constant if $\alpha = 1$. The last, equality, is exactly the same as the exponential case.

4.5.2.1 Case 1: Without Predicted Earnings and Predicted Pension Income

For the Weibull model with the hazard function (4.8) lacking predicted earnings and predicted pension income variables, most of the parameter values resemble the results reported in the exponential model in Table 4.4.1. For the benchmark individual, *ceteris paribus*, the hazard rate of retirement estimates can be derived from the Weibull model of

$$
\begin{aligned}
h(t; x_i) &= \alpha t^{\alpha-1}\lambda = \alpha t^{\alpha-1} \cdot e^{(\hat{\beta}_0 + \hat{\beta}_i \cdot 0)} \\
&= 1.537 * t^{0.537} * \exp(-6.456) > \exp(-6.456).
\end{aligned}
$$

The Weibull model lacks a constant hazard rate of retirement, in particular, $\alpha = 1.537 > 1$ and $1 \leq t \leq 54$, which indicates the hazard rate has positive duration dependence, $\dfrac{dh(t)}{dt} > 0$. And $-\ln(1/\alpha) = 0.430$, the estimate suggests that the hazard rate is increasing over time. As employment duration gets longer, the hazard rate increases and workers are more likely to retire.

Table 4.5.1 indicates that the estimated coefficients of those in Age2 (aged 55 to

59), Age3 (aged 60 to 64), and Age4 (aged 65 to 70) are positive and statistically significant and have higher hazard rates *ceteris paribus* (i.e. higher conditional retirement rates and hence shorter employment survival times). In the meantime, the hazard ratios for age groups are larger than one in the hazard-ratio representation: $\text{Age2}: \exp(0.224) = 1.251$, $\text{Age3}: \exp(0.424) = 1.528$, and $\text{Age4}: \exp(0.642) = 1.900$. This implies that older people are more likely to retire. Workers in Age2 (aged 55-59) are associated with a 25.1% higher hazard rate than Age1 (aged 50-54), and Age3 (aged 60 to 64) with 52.8% and Age4 (aged 65 to 70) with 90% higher hazard rates, *ceteris paribus*. Table 4.5.1 also presents other results in the Weibull model that female workers, Mainlander workers, and workers with poor health have a higher hazard rate of retirement. In contrast, workers with better education, married workers, and workers living in rural areas all have a lower hazard rate of retirement.

For gender effects, most estimated coefficients in Table 4.5.1 are similar to those estimated by the Exponential model shown in Table 4.4.1. For both men and women, older people and workers with poor health have a higher hazard rate and are more likely to retire, but rural workers have a lower hazard rate and are less likely to retire. Moreover, there are some different results for gender. For instance, male Mainlander workers have a significantly higher hazard rate of retirement, but insignificant for females. The Pension variable has a significantly positive effect on retirement for men, but a negative effect for women. In addition, the estimated values of the α parameters of men and women are significantly above one, indicating that they all have positive duration dependence. In particular, $\alpha_M = 2.044$ and $\alpha_w = 1.338$, which indicates that the rate at which the retirement hazard rate of men increases with employment duration over time is faster than that of women.

Table 4.5.1 Weibull Models without Predicted Variables

Sample	Overall		Male		Female	
Variables	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.224*	(.131)	.221	(.242)	.326**	(.158)
Age3	.424***	(.127)	.640***	(.229)	.297*	(.161)
Age4	.642***	(.142)	.967***	(.245)	.346*	(.191)
Gender	1.079***	(.091)	-		-	
Race2	-.193	(.119)	-.486**	(.204)	-.031	(.148)
Race3	.721***	(.149)	.540***	(.196)	.030	(.320)
Race4	-.173	(.312)	-.460	(.728)	-.206	(.350)
Edu2	-.017	(.101)	-.249	(.163)	.162	(.126)
Edu3	-.276*	(.145)	-.424**	(.206)	.116	(.215)
Edu4	-.521**	(.241)	-.462	(.301)	-.333	(.456)
Married	-.188*	(.099)	-.443***	(.163)	-.006	(.126)
Health	.534***	(.086)	.673***	(.132)	.322***	(.114)
Pension	.012	(.117)	.349**	(.158)	-.720***	(.216)
Town	-.066	(.106)	.138	(.156)	-.341**	(.149)
Rural	-.593***	(.104)	-.339**	(.157)	-.836***	(.139)
Constant	-6.456***	(.250)	-8.280***	(.454)	-4.566***	(.265)
$/\ln\_\alpha$	.430***	(.034)	.715***	(.052)	.291***	(.045)
α	1.537***	(.052)	2.044***	(.106)	1.338***	(.060)
$1/\alpha$	.651***	(.022)	.489***	(.025)	.747***	(.033)
No. of subjects	1732		1059		673	
No. of retirees	610		271		339	
Log likelihood	-1223.892		-518.975		-653.442	
LR chi2 (15)	336.29***		158.63***		68.52***	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

4.5.2.2 Case 2: With Predicted Earnings and Predicted Pension Income

Table 4.5.2 shows that most estimated coefficients are similar to those estimated by the Exponential model with predicted variables shown in Table 4.4.2. The estimated coefficients of Age2, Age3, Age4, Gender, and Health variables show significantly positive effects on retirement duration, and the Married, Town, and Rural variables have significant negative effects. This means that older workers, female workers, and workers with poor health have a higher hazard rate of retirement; and married workers, workers who live in town and rural areas have a lower hazard rate of retirement.

Table 4.5.2 also highlights that the P-Earnings variable is statistically significant in the models, in keeping with a priori expectations of a negative effect on retirement hazard. On the other hand, the P-Pension variable has a significant positive effect on retirement hazard. A possible reason for this is that higher predicted earnings might induce people to work longer and earn more; consequently, they are more likely to continue working and have a lower hazard rate of retirement. In contrast, workers with higher predicted pension income might have more opportunities to enjoy and manage their later life; so they are more likely to make their decision to retire.

Finally, for the gender effects in Table 4.5.2, most estimated effects are similar to those estimated by the Exponential model as shown in Table 4.4.2. In particular, male and female workers with higher predicted earnings have a lower hazard rate of retirement. Furthermore, male workers with higher predicted pension incomes have a higher hazard rate of retirement and are more likely to retire. But the analyses reported here did not find a statistically significant response for women.

Table 4.5.2 Weibull Models with Predicted Variables

Sample	Overall		Male		Female	
Variables	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.224*	(.131)	.221	(.242)	.326**	(.158)
Age3	.381***	(.127)	.597**	(.234)	.248	(.177)
Age4	.549***	(.142)	.881***	(.256)	.272	(.222)
Gender	.456***	(.091)	-		-	
Race2	-.193	(.119)	-.486**	(.204)	-.031	(.148)
Race3	-		-		-	
Race4	-.173	(.312)	-.460	(.728)	-.206	(.350)
Edu2	.017	(.101)	-.249	(.163)	.162	(.126)
Edu3	-.253	(.145)	-.371*	(.204)	.267	(.295)
Edu4	-		-		-	
Married	-.188*	(.099)	-.443***	(.163)	-.006	(.126)
Health	.534***	(.086)	.673***	(.132)	.322***	(.114)
Pension	.012	(.117)	.349**	(.158)	-.720***	(.216)
Town	-.291***	(.106)	-.036	(.169)	-.373**	(.170)
Rural	-1.197***	(.104)	-.796**	(.222)	-.878***	(.283)
P-Earnings	-.041***	(.008)	-.031***	(.011)	-.003	(.017)
P-Pension	.040***	(.011)	.029***	(.015)	-.004	(.024)
Constant	-5.728***	(.250)	-7.684***	(.528)	-4.539***	(.429)
$/\ln\_\alpha$	.430***	(.034)	.715***	(.052)	.291***	(.045)
α	1.537***	(.052)	2.044***	(.106)	1.338***	(.060)
$1/\alpha$	.651***	(.022)	.489***	(.025)	.747***	(.033)
No. of subjects	1732		1059		673	
No. of retirees	610		271		339	
Log likelihood	-1223.892		-518.975		-653.442	
LR chi2 (15)	336.29***		158.63***		68.52***	

Notes:

1. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.

2. The variables of Race3 and Edu4 are dropped due to collinearity with predicted variables in the models.

3. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

4.5.3 Weibull Models with Unobserved Heterogeneity

In this section a "frailty" component is included in the model. Frailty is a random component designed to account for variability due to unobserved individual-level factors that are otherwise unaccounted for by the other predictors in the retirement model. In particular, suppose the SHLS data belong to a random sample, the shared frailty models can be used for examining the effects of unobserved heterogeneity on retirement behaviour.

4.5.3.1 Case 1: Without Predicted Earnings and Predicted Pension Income

The estimated results of the models without predicted variables are shown in Table 4.6.1. First, without unobserved heterogeneity, most results are similar to the results previously reported in Table 4.5.1. That is, the estimated coefficients of those with Age3, Age4, Race3, Health variables are positive and statistically significant and have higher hazard rates *ceteris paribus*. In contrast, the estimated coefficients for Edu2, Edu3, Edu4, Married, and Rural variables are significantly negative. The estimate for the shape parameter is 1.479 suggesting an increasing hazard over time.

Second, the frailty in the model is assumed to follow a gamma distribution with mean 1 and variance equal to theta(θ). The estimate of theta is 0.262. A variance of zero (theta = 0) would indicate that the frailty component does not contribute to the model. A likelihood ratio test for the hypothesis theta = 0 is shown directly below the parameter estimates and indicates a chi-square value of 129.89 with 1 degree of freedom yielding a highly significant p-value of 0.000.

Notice how all the parameter estimates are altered with the inclusion of the frailty. The estimate for the shape parameter is now 1.536, different from the estimate

1.478 obtained from the model without frailty. The inclusion of frailty not only has an impact on the parameter estimates but also complicates their interpretation. The other estimated coefficients on the regressors Age2, Age3, Age4, and Race3 are slightly larger in magnitude that the corresponding coefficients in the reference model. The Weibull distribution shape parameter α is also slightly larger in the frailty model than in the reference model: 1.536 and 1.478, respectively. The median duration for the person with mean characteristics and the median among the sample as a whole is also slightly larger in the shared frailty model than in the reference model: 37.917 years and 37.674 years, respectively.

4.5.3.2 Case 2: With Predicted Earnings and Predicted Pension Income

The estimated results of the models with predicted variables are shown in Table 4.6.2. There is negligible unobserved heterogeneity - observe the near-zero frailty variances, and the $p-values$ for the likelihood ratio test equal to one. The estimated coefficients on the covariates are almost exactly the same as those in the corresponding model without unobserved heterogeneity. The possible reasons include: the model with predicted variables might be mis-specified to estimate the degree of duration dependence. In particular, most observations have the same predicted earnings and predicted pension incomes, they might easily trade off the effects of unobserved heterogeneity by themselves on retirement behaviour.

Table 4.6.1 Frailty Models without Predicted Variables

	Without Unobserved Heterogeneity		With Gamma-Heterogeneity	
Variables	Coefficient	Standard Error	Coefficient	Standard Error
Age2	.134	(.130)	.223*	(.130)
Age3	.336***	(.127)	.422***	(.126)
Age4	.543***	(.143)	.640***	(.141)
Race2	-.099	(.119)	-.191	(.119)
Race3	.499***	(.151)	.717***	(.148)
Race4	.1473	(.310)	-.168	(.311)
Edu2	-.330***	(.095)	.012	(.101)
Edu3	-.582***	(.139)	-.280*	(.145)
Edu4	-.812***	(.237)	-.524**	(.240)
Married	-.373***	(.099)	-.190*	(.098)
Health	.592***	(.085)	.534***	(.085)
Pension	-.117	(.115)	.010	(.117)
Town	-.150	(.105)	-.067	(.105)
Rural	-.675***	(.102)	-.594***	(.103)
Constant	-5.263***	(.222)	-5.770***	(.432)
$/\ln\_\alpha$	.391***	(.034)	.429***	(.033)
$/\ln\_the$			-1.338	(.972)
α	1.478***	(.050)	1.536***	(.050)
$1/\alpha$	.676***	(.023)	.650***	(.021)
theta			.262	(.254)
No. of subjects	1732		1732	
No. of retirees	610		610	
Log likelihood	-1294.315		-1229.368	
LR chi2(14)	195.45***		171.18***	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. In particular, Log-likelihood ratio test of theta = 0: chibar2 (01) = 129.89, Prob >= chibar2 = 0.000.

Table 4.6.2 Frailty Models with Predicted Variables

	Without Unobserved Heterogeneity		With Gamma-Heterogeneity	
Variables	Coefficient	Standard Error	Coefficient	Standard Error
Age2	.224*	(.130)	.224*	(.130)
Age3	.408***	(.134)	.408***	(.134)
Age4	.567***	(.156)	.567***	(.156)
Race2	-.192	(.119)	-.192	(.119)
Race3	-.555***	(.174)	-.555***	(.174)
Race4	-.173	(.311)	-.173	(.311)
Edu2	.017	(.100)	.017	(.100)
Edu3	-.429***	(.154)	-.429***	(.154)
Edu4	-		-	
Married	-.187*	(.098)	-.187*	(.098)
Health	.533***	(.085)	.533***	(.085)
Pension	.012	(.117)	.012	(.117)
Town	-.436***	(.109)	-.436***	(.109)
Rural	-1.640***	(.130)	-1.640***	(.130)
P-Earnings	-.070***	(.005)	-.070***	(.005)
P-Pension	.078***	(.010)	.078***	(.010)
Constant	-5.451***	(.300)	-5.451***	(.300)
$/\ln\_\alpha$	.429***	(.033)	.429***	(.033)
$/\ln\_the$			-23.319	(635.147)
α	1.537***	(.051)	1.537***	(.051)
$1/\alpha$	.650***	(.022)	.650***	(.022)
theta			7.46e-11	(4.74e-08)
No. of subjects	1732		1732	
No. of retirees	610		610	
Log likelihood	-1223.892		-1223.892	
LR chi2 (15)	336.29***		182.14***	

Notes:
1. The Edu4 variable was dropped due to collinearity with predicted variables in the models.
2. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.
3. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. However, Log-likelihood ratio test of theta = 0: chibar2 (01) = 0.00, Prob>=chibar2 = 1.000.

4.5.4 Cox Proportional Hazard Models

Compared to the parametric approach, the advantage of the semi-parametric Cox model is that it does not make assumptions about $h_0(t)$, and it can estimate β_i more accurately without the constant distribution. This analysis, first considered by Diamond and Hausman (1984), is derived from Cox's (1972) proportional hazard model to examine determinants of retirement behaviour and has become increasingly popular in labour economics research.

Recall the Cox proportional hazard model discussed in Section 4.3.2.2, the formulated analysis of employment duration data where no parametric form of the hazard function is specified, and yet the effects of the covariates are parameterised to alter the baseline hazard function (that for which all covariates are equal to zero) in a certain way (Cleves et al, 2004).

$$h_i(t \mid x_i) = h_0(t)\exp(\beta' x_i). \tag{4.12}$$

The baseline hazard $h_0(t)$ is simply left unparameterised, and through conditioning on failure times, estimates of β are obtained anyway. The Cox proportional hazard model also considers two different cases, including the model (i) without predicted earnings and predicted pension income, (ii) with predicted earnings and predicted pension income, as explanatory variables. The estimated results are shown in Tables 4.7.1 and 4.7.2, which not only estimate the coefficients, but also analyse the hazard ratio for multivariable records. If the estimated coefficient $\beta > 0$, the hazard rate of retirement increases. If $\beta < 0$, the hazard rate of retirement decreases. If $\beta = 0$, then there is no effect on the hazard rate of retirement. If the hazard ratio is greater than one, it indicates that the Cox hazard is greater than the baseline hazard, and the hazard rate of retirement is higher. If the hazard ratio is less than one, the hazard rate of retirement is lower.

4.5.4.1 Case 1: Without Predicted Earnings and Predicted Pension Income

Table 4.7.1a shows that most estimated coefficients are similar to those estimated by the Exponential and Weibull models in Tables 4.4.1 and 4.5.1. The Age3 (aged 60 to 64), Age4 (aged 65 to 70), Gender, Race3 (Mainlander), and Health variables present a strong significant and positive effect on retirement. This means that older workers, female workers, Mainlander workers, and workers with poor health have a higher hazard rate of retirement. In contrast, the Race2 (Hakka), Edu3 (7 to 12 years of schooling), Edu4 (13 to 17 years of schooling), Married, and Rural variables have a significant and negative effect on retirement. This means that Hakka workers, workers with better education, married workers, and workers living in rural areas have a lower hazard rate of retirement.

The hazard ratio results in Table 4.7.1b give the ratio of a variable against its base dummy. The hazard ratio is computed as follows:[18]

$$\hat{h}(t \mid x_i = 0) = \hat{h}_0(t)$$

$$\hat{h}(t \mid x_i = 1) = \hat{h}_0(t)\exp(\beta_i{}' x_i) = \hat{h}_0(t)\exp(\beta_i{}')$$

where $\beta_i{}'$ are the coefficient values. For example, the relevant effect of gender on hazard ratio is calculated from Table 4.7.1a as follows:

$$\hat{h}(t \mid Gender = 0) = \hat{h}_0(t)$$

$$\hat{h}(t \mid Gender = 1) = \hat{h}_0(t) \cdot \exp(1.120).$$

Then, the hazard ratio of Gender is $\exp(1.120) = 3.065$. More exactly, these results

[18] See Cleves et al (2002), Chapter 9.

show that if we constrain the hazard rate of females to a multiplicative constant of the hazard rate of males, then that multiplicative constant is estimated to be 3.065. Furthermore, the hazard ratio of Race3 (Mainlander) is $\exp(0.802) = 2.230$. This shows that, holding other variables constant, the estimated hazard rate of Mainlander workers retiring compared to that of Fujianese workers is 2.230 times greater. Mainlander workers have a higher hazard rate of retiring than Fujianese workers, assuming other variables are constant. Other results include: workers with poor health have a hazard rate of retiring equal to 1.755 times that of workers in good health. However, workers with higher levels of education have a hazard rate of retiring equal to 0.604 times that of workers with informal education. Married workers have a hazard rate of retiring equal to 0.844 times that of unmarried workers. Finally, rural workers have a hazard rate of retiring equal to 0.535 times that of urban workers. This suggests that urban workers have a higher hazard rate of retirement than rural workers.

For gender effects, most estimated coefficients and hazard ratios in Tables 4.7.1a and 4.7.1b are similar to those estimated by the Exponential and Weibull models in Tables 4.4.1 and 4.5.1. For instance, older people and workers with poor health have a higher hazard rate and are more likely to retire, but rural workers have a lower hazard rate and are less likely to retire. In contrast, there are some different results for gender. For example, male Mainlander workers have a significantly higher hazard rate of retirement, but this is insignificant for females. The Pension variable for men has a significantly positive effect on retirement, but a negative effect for women.

Table 4.7.1a Cox Proportional Hazard Models without Predicted Variables

Sample	Overall		Male		Female	
Variables	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.176	(.131)	.168*	(.243)	.302*	(.159)
Age3	.286**	(.130)	.445*	(.234)	.233	(.163)
Age4	.457***	(.145)	.770***	(.248)	.193	(.199)
Gender	1.120***	(.092)	-		-	
Race2	-.244**	(.120)	-.551***	(.206)	-.064	(.148)
Race3	.802***	(.149)	.558***	(.196)	.085	(.321)
Race4	-.225	(.311)	-.635	(.731)	-.259	(.350)
Edu2	.022	(.102)	-.275*	(.164)	.210	(.128)
Edu3	-.266*	(.147)	-.442**	(.208)	.163	(.219)
Edu4	-.505**	(.241)	-.453	(.303)	-.323	(.457)
Married	-.169*	(.099)	-.473***	(.164)	.034	(.127)
Health	.562***	(.086)	.699***	(.133)	.358***	(.114)
Pension	.063	(.118)	.465***	(.163)	-.732***	(.217)
Town	-.075	(.106)	.127	(.157)	-.348**	(.150)
Rural	-.626***	(.104)	-.395**	(.159)	-.862***	(.140)
No. of subjects	1732		1059		673	
No. of retirees	610		271		339	
Log likelihood	-3810.248		-1532.121		-1845.918	
LR chi2(15)	337.99***		158.70***		72.45***	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

Table 4.7.1b Cox Proportional Hazard Models without Predicted Variables

Sample	Overall		Male		Female	
Variables	Haz. Ratio	Std. Err.	Haz. Ratio	Std. Err.	Haz. Ratio	Std. Err.
Age2	1.192	(.156)	1.183	(.288)	1.352*	(.215)
Age3	1.331**	(.173)	1.151*	(.365)	1.262	(.206)
Age4	1.579***	(.230)	2.160***	(.535)	1.213	(.241)
Gender	3.065***	(.281)	-		-	
Race2	.783**	(.094)	.576***	(.119)	.938	(.139)
Race3	2.230***	(.333)	1.747***	(.342)	1.089	(.350)
Race4	.798	(.249)	.530	(.388)	.772	(.270)
Edu2	1.022	(.104)	.759*	(.124)	1.233	(.158)
Edu3	.766*	(.113)	.643**	(.134)	1.177	(.257)
Edu4	.604**	(.146)	.635	(.192)	.724	(.331)
Married	.844*	(.084)	.623***	(.102)	1.035	(.131)
Health	1.755***	(.151)	2.012***	(.268)	1.431***	(.163)
Pension	1.065	(.126)	1.592***	(.259)	.481***	(.104)
Town	.927	(.099)	1.135	(.178)	.706**	(.106)
Rural	.535***	(.056)	.674**	(.107)	.422***	(.059)
No. of subjects	1732		1059		673	
No. of retirees	610		271		339	
Log likelihood	-3810.248		-1532.121		-1845.918	
LR chi2 (15)	337.99***		158.70***		72.45***	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

4.5.4.2 Case 2: With Predicted Earnings and Predicted Pension Income

Table 4.7.2a shows that most estimated coefficients are similar to those estimated by the Exponential and Weibull models with predicted variables presented in Tables 4.4.2 and 4.5.2. In particular, the empirical result highlights that the P-Earnings variable is statistically significant in the Cox proportional hazard model, in keeping with a priori expectations of a negative effect on retirement hazard. This implies that workers with higher predicted earnings might be induced to work longer and earn more; consequently, they are more likely to continue working and have a lower hazard rate of retirement. In contrast, the P-Pension variable has a significant positive effect on retirement hazard. This means higher predicted pension income might reduce the incentive for people to work longer and reassure them about financial security in later life; so they are more likely to retire and have a higher hazard rate of retirement.

The other estimated coefficients of Age2, Age3, Age4, Gender, and Health variables have significantly positive effects on retirement duration, and the Married, Town, and Rural variables have significant negative effects. This means that older workers, female workers, and workers with poor health have a higher hazard rate of retirement; and married workers, and workers who live in town and rural areas have a lower hazard rate of retirement. However, the Race3 and Edu4 variables might have some collinearity problems for calculating predicted earnings and predicted pension income variables and are dropped.

For gender effects, firstly, Figure 4.1 shows estimated baseline cumulative hazard in the Cox proportional hazard model without predicted earnings and predicted pension income variables for women and men. The estimated curve of non-baseline cumulative hazard for women is higher than the baseline cumulative hazard curve for men. This indicates that female workers have a higher hazard rate of retirement than males. Secondly, Figure 4.2 also shows estimated baseline cumulative hazard in the Cox proportional hazard model with predicted earnings and predicted pension income variables for women and men. The gap between the two estimated curves in Figure 4.2 gradually decreases and is smaller than the gap in Figure 4.1. This implies that people with higher predicted earnings and predicted pension income have a higher hazard rate of retirement, particularly for men.

Table 4.7.2a Cox Proportional Hazard Models with Predicted Variables

Sample	Overall		Male		Female	
Variables	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.176	(.131)	.168	(.243)	.302*	(.159)
Age3	.250*	(.136)	.405*	(.240)	.188	(.179)
Age4	.370**	(.158)	.687***	(.259)	.123	(.229)
Gender	.431***	(.143)	-		-	
Race2	-.244**	(.120)	-.551***	(.206)	-.064	(.148)
Race3	-		-		-	
Race4	-.225	(.311)	-.635	(.731)	-.259	(.350)
Edu2	.022	(.102)	-.276*	(.164)	.210	(.128)
Edu3	-.277*	(.163)	-.400**	(.204)	.292	(.296)
Edu4	-		-		-	
Married	-.169*	(.099)	-.473***	(.164)	.034	(.127)
Health	.562***	(.086)	.699***	(.133)	.358***	(.114)
Pension	.063	(.118)	.465***	(.163)	-.732***	(.217)
Town	-.321***	(.114)	-.051	(.170)	-.393**	(.172)
Rural	-1.294***	(.155)	-.866***	(.222)	-.948***	(.287)
P-Earnings	-.045***	(.008)	-.032***	(.011)	-.006	(.017)
P-Pension	.046***	(.011)	.030**	(.015)	-.001	(.024)
No. of subjects	1732		1059		673	
No. of retirees	610		271		339	
Log likelihood	-3810.248		-1532.121		-1845.918	
LR chi2 (15)	337.99***		158.70***		72.45***	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. The variables of Race3 and Edu4 are dropped due to collinearity with predicted variables in the models.

3. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

Table 4.7.2b Cox Proportional Hazard Models with Predicted Variables

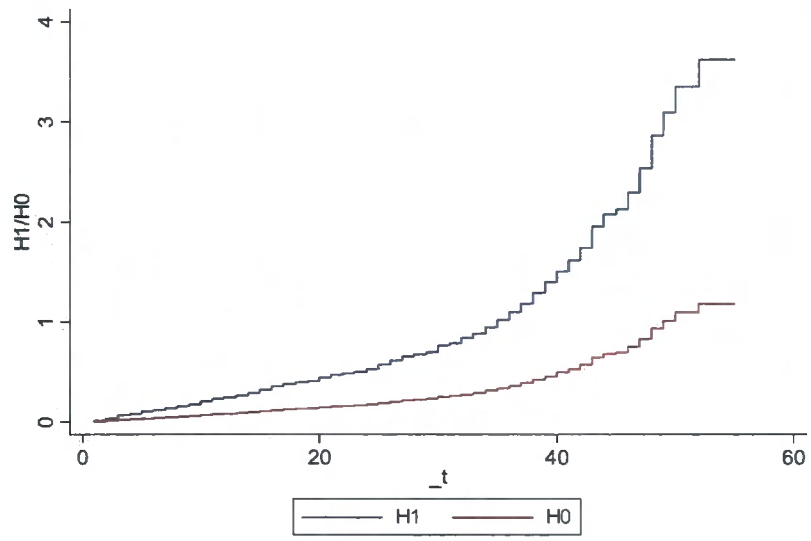
Sample	Overall		Male		Female	
Variables	Haz. Ratio	Std. Err.	Haz. Ratio	Std. Err.	Haz. Ratio	Std. Err.
Age2	1.192	(.157)	1.183	(.288)	1.352*	(.215)
Age3	1.283*	(.175)	1.499*	(.359)	1.207	(.217)
Age4	1.448**	(.229)	1.987***	(.516)	1.131	(.259)
Gender	1.539***	(.220)	-		-	
Race2	.783**	(.094)	.576***	(.119)	.938	(.139)
Race3	-		-		-	
Race4	.799	(.249)	.530	(.388)	.772	(.270)
Edu2	1.022	(.104)	.759*	(.124)	1.233	(.158)
Edu3	.758*	(.123)	.670**	(.137)	1.339	(.397)
Edu4	-		-		-	
Married	.844*	(.084)	.623***	(.102)	1.035	(.131)
Health	1.755***	(.151)	2.013***	(.268)	1.431***	(.163)
Pension	1.065	(.126)	1.592***	(.259)	.481***	(.104)
Town	.725***	(.083)	.950	(.161)	.675**	(.116)
Rural	.274***	(.043)	.421***	(.094)	.388***	(.111)
P-Earnings	.956***	(.008)	.969***	(.010)	.994	(.017)
P-Pension	1.047***	(.012)	1.031**	(.016)	.999	(.024)
No. of subjects	1732		1059		673	
No. of retirees	610		271		339	
Log likelihood	-3810.248		-1532.121		-1845.918	
LR chi2 (15)	337.99***		158.70***		72.45***	

Notes:

1. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.

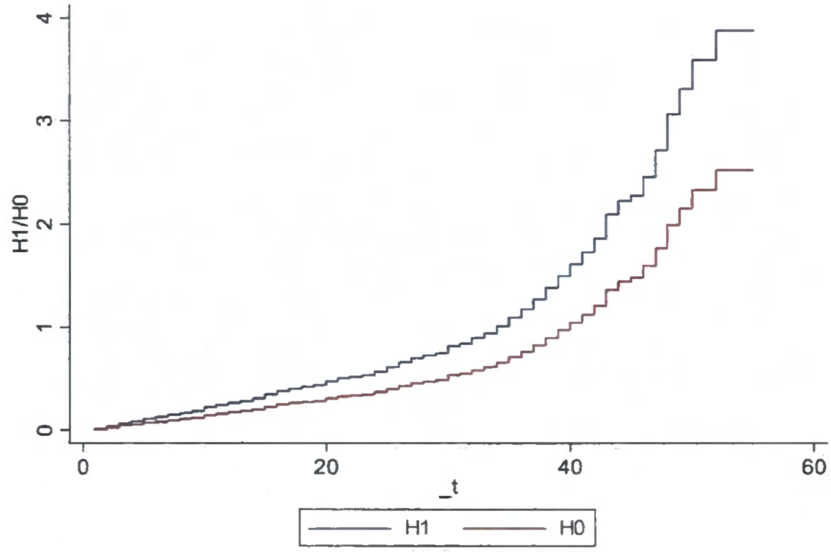
2. The variables of Race3 and Edu4 are dropped due to collinearity with predicted variables in the models.

3. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.



Note: H1 for women and H0 for Men

Figure 4.1 Estimated Cumulative Hazard in the Cox Proportional Hazard Models without Predicted Earnings and Predicted Pension Income



Note: H1 for women and H0 for Men

Figure 4.2 Estimated Cumulative Hazard in the Cox Proportional Hazard Models with Predicted Earnings and Predicted Pension Income

4.6 Discussions

In general, this chapter uses the continuous-time parametric models to assess individual retirement behaviours. The estimated results of three models without considering the effect of unobserved heterogeneity are summarised in Table 4.8. Most estimated results on these continuous-time hazard models are similar across cases. For example, in Case 1 without predicted variables, the Age2 (aged 55 to 59), Age3 (aged 60 to 64), Age4 (aged 65 to 70), Gender, Race3 (Mainlander), and Health variables have positive and significant effects on retirement hazard, which can confirm that older workers, female workers, Mainlander workers, and workers with poor health have a higher hazard rate of retirement and are more likely to retire. However, the age variable is a continuous time parameterisation such that it might be not suitable to use as the basis for a grouped hazard. Therefore, the Age2 (aged 55 to 59), Age3 (aged 60 to 64), and Age4 (aged 65 to 70) variables have a significant effect on retirement hazard in the Exponential and Weibull models, but less significant in the Cox proportional hazard model. In contrast, the Edu3 (7 to 12 years of schooling), Edu4 (13 to 17 years of schooling), Married, and Rural variables have a negative and significant effect on labour force withdrawal, which implies that workers with better education, married workers, and rural workers have a lower hazard rate of retirement.

The result of Case 2 highlights that the variable P-Earnings has a significant negative effect on retirement hazard, and the P-Pension variable has a positive effect. This confirms that if workers expect to have higher predicted earnings, they are less likely to retire; in contrast, if they can have higher predicted pension income, they are more likely to retire. This result is consistent with the previous empirical studies reported by Mitchell and Fields (1981) and Diamond and Hausman (1984). That is, higher wages lead to delayed retirement, and higher pension benefits lead to earlier retirement.

Moreover, for unobserved heterogeneity, the empirical result of the frailty model without predicted variables can confirm that a model with unobserved heterogeneity may improve the model without unobserved heterogeneity in Table 4.6.1. The estimated coefficients of Age2, Age3, Age4, and Race3 are slightly larger in magnitude that the corresponding coefficients in the reference model. In particular, Weibull models without predicted variables have significant positive duration dependence across all years, the estimated values of the α parameter being significantly above one. Moreover, the α parameter changes from $\alpha = 1.537$ in Table 4.5.1 to $\alpha = 1.478$ in Table 4.6.1. Therefore, the unobserved heterogeneity may deduct a little effect of duration dependence.[19]

However, Table 4.6.2 shows the frailty model with predicted variables that the estimated coefficients on the covariates are almost exactly the same as those in the corresponding model without unobserved heterogeneity. In particular, Weibull models with predicted variables also have same estimated values, $\alpha = 1.537$ in Table 4.5.2 and Table 4.6.2. The possible reason for the failure to identify unobserved heterogeneity in the hazard model is that the predicted variables might not be suitable to be explanatory variables for frailty in an employment duration model. In particular, since the predicted variables used a smaller sub-sample to predict the whole sample, it might be difficult to find the effects of unobserved heterogeneity. Hence, ideally the study would use the data on real earnings and real pension benefits to find the effects of other unobserved heterogeneity in the employment duration model.

[19] See, Nolan (2000) uses a grouped hazard approach and concerns with estimating the tendency for duration dependence as well.

Table 4.8 Estimated Results of Three Models without Frailty: Summary

Models	Exponential Model		Weibull Model		Cox PH Model	
Duration	Case 1	Case 2	Case 1	Case 2	Case 1	Case 2
Age2	+**	+**	+**	+*	+	+
Age3	+***	+***	+***	+***	+**	+*
Age4	+***	+***	+***	+***	+***	+**
Gender	+***	+***	+***	+***	+***	+***
Race2	-	-	-	-	-**	-**
Race3	+***	(-)	+***	(-)	+***	(-)
Race4	-	-	-	-	-	-
Edu2	-	-	-	+	+	+
Edu3	-*	-	-*	-	-*	-*
Edu4	-**	(-)	-**	(-)	-**	(-)
Married	-*	-*	-*	-*	-*	-*
Health	+***	+***	+***	+***	+***	+***
Pension	+	+	+	+	+	+
Town	-	-*	-	-***	-	-***
Rural	-***	-***	-***	-***	-***	-***
P-Earnings		-***		-***		-***
P-Pension		+***		+***		+***
Constant	-***	-***	-***	-***		
$/\ln\_\alpha$			+***	+***		
α			+***	+***		
$1/\alpha$			+***	+***		

Note:

The estimated results of Case 1 are without predicted variables in the three models, but Case 2 with predicted variables. The variables of Race3 and Edu4 are dropped due to collinearity with predicted variables in the above three models.

4.7 Conclusions

This chapter investigated the factors that influence retirement behaviour among the middle-aged and elderly in Taiwan, using continuous-time hazard models in a sample of individuals aged 50 to 70 years. Both the parametric and semi-parametric approaches showed that personal, family, employment opportunity, and economic factors are important in individual retirement decisions. For instance, older workers, female workers, Mainlander workers, and workers with poor health all have a higher hazard rate of retirement. In contrast, workers with better education often have higher productivity and more employment opportunities, so have a lower hazard rate of retirement. For the family factors, married workers have a greater responsibility to earn money for their family, so need to retire later. For the employment opportunity, rural workers have a lower hazard rate of retirement because opportunities exist for them to be self-employed or to easily find low skilled jobs.

Further, the results highlight the importance of predicted earnings and predicted pension income for explaining retirement decisions. In particular, the information in this chapter shows that workers with higher predicted earnings have a lower hazard rate of retirement, but workers with higher predicted pension income have a higher hazard rate of retirement. In particular, the introduction of the 2005 portable pension system *is likely* to lead to an increase in expected pension incomes and workers might have a higher hazard rate of retirement. On the one hand, employees have more security for their jobs, particularly where there is a higher frequency of labour turnover. On the other hand, employers can also find better employees for their firms, particularly with higher productivity.

Finally, for unobserved heterogeneity, the frailty model without predicted variables supports that a model with unobserved heterogeneity may improve the model without unobserved heterogeneity. This has some implications for retirement decisions. In adopting a Weibull hazard specification, the estimated values of the α parameter being significantly above one, and the α parameter become smaller with unobserved heterogeneity. That is, the factors of unobserved heterogeneity may deduct a little from the effect of duration dependence. Further, for the frailty model with predicted variables, the estimated coefficients on the covariates are almost exactly the same as those in the corresponding model without unobserved heterogeneity. Hence, there is negligible unobserved heterogeneity. The study could be improved by incorporating new data on real earnings and pension income variables in employment duration models.

4.8 Appendix

The brief STATA commands for analysing the retirement decisions by the employment duration models are given as follows:

```
Use "C:\Documents and Settings\User\My Documents\Revised 2007
Summer\SHLS Data 2007\Chapter 4 Data Set 082007.dta",

Table 4.2.1
gen ey= earnings
reg ey age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3
reg ey age3-age4 gender race3 edu3-edu4 resid2 resid3
predict eyhat if history~=0
sum eyhat

Table 4.2.2
tab f17g3c1
gen peny= income if f17g3c1==3
tab peny
replace peny=. if peny==0
recode peny 1=5 2=20 3=45 4=80 5=150
tab peny
reg peny age3-age4 gender race2-race4 edu2-edu4 married poorh pension
resid2 resid3
reg peny age3-age4 gender race3 edu3-edu4 resid2 resid3
predict penyhat if history~=0
sum penyhat

Table 4.3
sum duration censor age1-age4 gender race1-race4 edu1-edu4 married poorh
pension resid1-resid3 ey eyhat peny penyhat if duration~=. & resid~=. &
race~=. & eyhat~=.

Table 4.4.1
streg age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=., distribution(exponential) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
if eyhat~=. & gender==0, distribution(exponential) nohr
```

```
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
if eyhat~=. & gender==1, distribution(exponential) nohr
```

Table 4.4.2

```
streg age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat, distribution(exponential) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
eyhat penyhat if gender==0, distribution(exponential) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
eyhat penyhat if gender==1, distribution(exponential) nohr
```

Table 4.5.1

```
streg age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=., distribution(weibull) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
if eyhat~=. & gender==0, distribution(weibull) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
if eyhat~=. & gender==1, distribution(weibull) nohr
```

Table 4.5.2

```
streg age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat, distribution(weibull) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
eyhat penyhat if gender==0, distribution(weibull) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
eyhat penyhat if gender==1, distribution(weibull) nohr
```

Table 4.6.1

```
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
if eyhat~=., distribution(weibull) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
if eyhat~=., distribution(weibull) nohr frailty(gamma) shared(gender)
```

Table 4.6.2

```
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
eyhat penyhat if eyhat~=., distribution(weibull) nohr
streg age2-age4 race2-race4 edu2-edu4 married poorh pension resid2 resid3
eyhat penyhat if eyhat~=., distribution(weibull) nohr frailty(gamma)
```

```
shared(gender)

Table 4.7.1a
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=., nohr
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=.& gender==0, nohr
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=.& gender==1, nohr

Table 4.7.1b
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=.
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=.& gender==0
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=.& gender==1

Table 4.7.2a
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat, nohr
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat if gender==0, nohr
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat if gender==1, nohr

Table 4.7.2b
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat if gender==0
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat if gender==1

Figure 4.1
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 if eyhat~=., basech (H0)
line H0 _t, c(J) sort
```

```
gen H1=3.065* H0
label variable H0 H0
line H1 H0 _t, c(J J) sort
```

Figure 4.2

```
stcox age2-age4 gender race2-race4 edu2-edu4 married poorh pension resid2
resid3 eyhat penyhat if eyhat~=., basech (H0)
line H0 _t, c(J) sort
gen H1=1.539* H0
label variable H0 H0
line H1 H0 _t, c(J J) sort
```

Chapter 5

Labour Force Transition

5.1 Introduction

Chapters 3 and 4 used cross-sectional data to analyse labour force participation and retirement decisions. The results show that personal, family, employment opportunity, and economic factors are important for determining individual labour force participation and retirement behaviour. However, cross-sectional data analyses are not satisfactory because they either assume the values of the explanatory variables to be constant with employment duration, or they do not examine retirement decisions with the same observations over the interval period (Slade, 1987). To address these deficiencies, this chapter uses panel data for analysis that can have time-varying covariates and capture individuals' labour force transition and retirement behaviour.

Previous studies examined the influence of a number of factors on labour force transition (LFT) and retirement, such as health, marital status, pension, and others. Regarding the impact of changes in health on LFT, Bound et al. (1999) and Disney et al. (2003) showed that "health shocks" greatly influenced the LFT of the elderly. Considering the impact of marital status on LFT, Hurd (1988), Blau (1998), and Blau and Riphahn (1999) showed that marital status is an important factor for joint retirement. Since most Taiwanese marry only once, any changes in their marital status may significantly affect their participation in the labour market. For instance, Table 5.1 shows that only a small percentage of the middle aged and elderly were divorced or separated. Most people were married and the percentage of those widowed increased with age. The third factor is pension status. Boskin (1977) and Slade (1987) noted that pension factors are important determinants of LFT. In Taiwan, most occupational pensions are received in a single payment. If the retirees deposit this pension sum in a designated financial institution, they can receive an interest rate as high as 18 percent, thus those elderly people are more likely to retire early to take advantage of this special law, and then re-enter the labour market. Finally, the impact of changes in residence status can also affect employment opportunities and living conditions for the elderly. Elderly workers who move from urban to rural areas may have a good living environment; those who move from rural to town may have more social benefits or employment opportunities. Therefore, this chapter will provide an important baseline for gauging changes in the employment pattern of older workers in the future as new policies continue to evolve. Which factors affect labour force transition behaviour? What is the role of health and "health shocks", marriage and "marriage shocks" in their transition decisions? What are the effects of unobserved heterogeneity, or do the characteristics of individuals affect their decisions?

To examine the above factors, probit models are first used to estimate the probability of LFT, including the probabilities of exit from and re-entry into employment. Secondly, the duration models in chapter 4 are going to be extended with panel data and add time-varying covariates to capture individuals' LFT and retirement behaviour. The exit employment probit model observes the sample of those who stop working, including unemployment and retirement, while the duration model focuses only on the retirement hazard.

Table 5.1 Marital Status of the Middle Aged and Elderly in Taiwan

Unit: %

Age Groups	Married	Spouse Deceased	Divorced or Separated	Single
50-54	85.3	5.7	6.3	2.7
55-59	82.9	9.0	6.1	2.0
60-64	77.9	15.9	5.1	1.1
65-69	69.5	26.6	3.1	0.8
70-74	62.3	33.4	3.0	1.3
75-79	50.0	45.9	1.6	2.5
80+	39.3	57.3	1.7	1.7

Source: Ministry of the Interior (2005), "*An Abstract Analysis for the 2005 Survey of Elderly Condition in Taiwan*", Department of Statistics, Taiwan (in Chinese). The website address is as follows: http://www.moi.gov.tw/stat/index.asp.

To illustrate, two waves of the Survey of Health and Living Status of the Middle Aged and Elderly in Taiwan (SHLS) between 1996 and 1999 are merged to create a multiple data set. This is a random sample of spells with right censoring but the censoring point varies. Estimation of continuous time parametric regression models incorporating time-varying covariates (hereafter, TVCs) requires episode splitting.[1] One has to split the survival time for each individual into sub-periods within which each TVC is constant. Multiple records are created for each individual, with one record per sub-period. Since the likelihood is only evaluated at the times at which failures occur in the data, the computation only depends on the risk pools at those failure times. Changes in covariates between failure times do not affect estimates for an ordinary regression model. Thus, to estimate a model with TVCs, all one has to do is define the values of these TVCs at all failure times at which a subject was at risk (Collett, 1994). Furthermore, considering the effect of unobserved heterogeneity in the duration models without or with TVC is also important for estimating the determinants of retirement hazard.

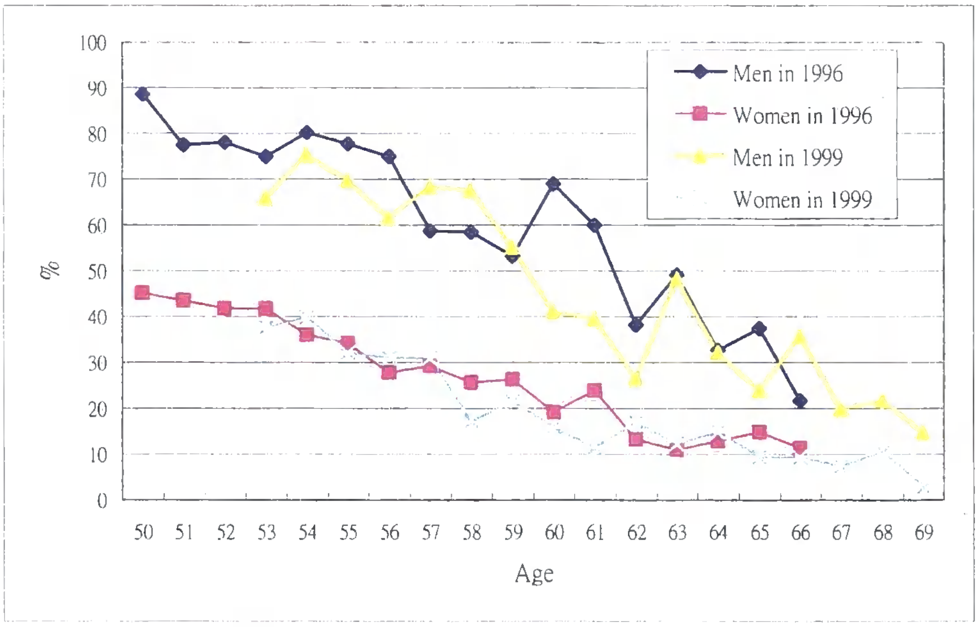
The rest of this chapter is organised as follows. Section 5.2 discusses the background of LFT, including some recent changes and specific covariates that vary with time. Section 5.3 presents the theoretical framework of LFT, including a review of related literature and major hypotheses. Section 5.4 devises the empirical specification of LFT, including probit and duration models. Section 5.5 describes the data sources and variables used in the analysis. The empirical results and discussion are presented in Section 5.6, followed by the conclusions in Section 5.7.

[1] See Jenkins (2003) for explanation of episode splitting.

5.2 Background to Labour Force Transition

5.2.1 Changes in Labour Force Participation

According to the SHLS data, Figure 5.1 shows the trends of labour force participation rates by men and women working full-time job between 1996 and 1999, respectively. The profiles for men are quite similar in showing a sharp drop between aged 54 and 59 in the 1996 wave, with a more fluctuating decline after these ages. After three years, the transitions from full-time work sharply declined between aged 57 and 62 in the 1999 wave. Further, the patterns of female labour force participation show a sharp decline between aged 50 and 56 in the 1996 wave and a more gradual decline after aged 61. However, this trend of female labour force transition is subject to more fluctuation after aged 57 in the 1999 wave. The trends in the two waves imply that men always have a higher participation rate in full-time work than women, and men also have a higher fluctuation rate than women.



Source: Author graphic using the 1996 and 1999 SHLS survey.

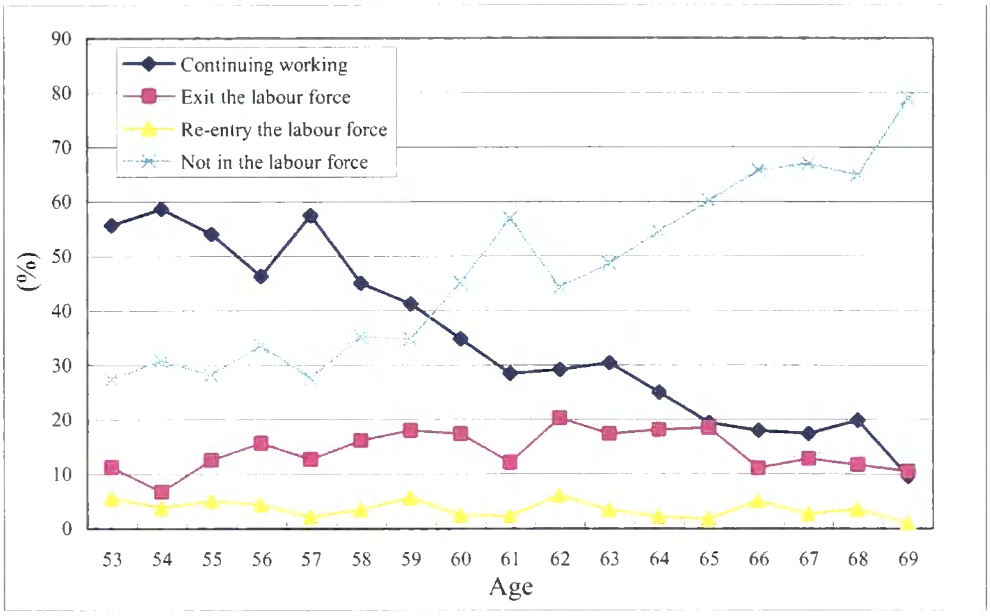
Figure 5.1 Proportion Rates of Full-Time Work by Age and Gender

Figure 5.2 shows the trends in labour force transitions (LFT) by age. Notice that there are two peaks for the LFT of continued working, which occur at aged 54 and 57 respectively. According to the Labour Standards Law (LSL) in Taiwan, a worker who is in one of the following categories may apply for voluntary retirement: (1) at the age of 55 after working 15 years; (2) after working more than 25 years. When the worker reaches aged 60 or when he/she is incapacitated owing to mental defect or physical handicap, retirement is mandatory. Hence, after aged 57, the proportion rates of transition from continued working gradually decrease.

Moving to the LFT for workers exiting the labour force shown in Figure 5.2, it shows four peaks, at aged 56, 59, 62 and 65 respectively. The first peak of retirement is due to the fact that many workers have completed 25 years of employment, a condition that presages retirement. The third peak, at aged 62, is the result of the mandatory retirement age for non-government workers. This peak is the largest shown in the SHLS survey, which contains approximately 70% non-governmental workers. The last peak, at aged 65, derives from the mandatory retirement age for government employees.

Regarding the LFT for workers re-entering the labour market, there is a pattern of roughly below 5%, especially for the aged 55, 59, 62 and 66. This could be because the three peaks of actual retirement age, at the aged 55, 60 and 65 (as previously mentioned and shown in Figure 2.5), have created a group of retirees who are still capable of continuing work but are retired. Due to their experience and valuable skills, they may re-enter the job market after a brief respite at aged 55 to 65 to further their career; some may even have switched to different types of work.

The trends of not in the labour force in Figure 5.2 show an increasing propensity from about 28 percent by aged 53 to 58 percent by aged 61, and to 79 percent by aged 69. In particular, more than 60 percent of people were not in the labour force after aged 65.



Source: Author graphic using the 1999 SHLS data.

Figure 5.2 Trends in Labour Force Transition by Age

5.2.2 A Discussion of Time-Varying Covariates

According to the SHLS data, the effective sample of duration model only had 1732 observations, including 610 retired and 1122 continuing work in 1996. After three years, the effective sample were changed from 1122 people continuing work in 1996 to 253 people retired, 713 continuing work, 39 unemployed, and 117 missing for moved, deaths, or no answer in 1999. Possible reasons include individual and socioeconomic factors changed, such as changes in health, marital status, pension entitlement, and residence status factors. These factors can be defined as time-varying covariates, which may influence individuals' labour force transition and retirement behaviour. The details are described as shown in Table 5.2.

Comparing workers' health between 1996 and 1999, the SHLS data shows that, for the effective sample of the duration model in 1996, 20.6% of workers' health was excellent, 22.8% good, 32.6% average, 20.7% not so good, and 3.1% poor. Three years later, these workers' health had changed: 20.1% of workers' health was excellent, 31.1% good, 32.7% average, 14.5% not so good, and 1.6% poor. The proportion of workers with not so good and poor health declined rapidly, possibly health has been improved by the National Health Insurance (NHI) programmes from 1995. The other categories remain similar.

With regard to changes in the marital status of workers between 1996 and 1999, in 1996, 84.1% of workers were married, 2.7% not married, 2.3% divorced, 0.7% separated, and 10.2% widowed. Three years later, these workers' marital status changed slightly, 86.3% of workers were married, 2.3% not married, 2.3% divorced, 0.5% separated, and 8.6% widowed. The marital status of the workers in the study did not change as much as did their health status in the interim between the two surveys.

Workers' pension entitlements changed 7.8% between 1996 and 1999; 26.7% of individuals had pension entitlements in 1996, but after three years workers' pension entitlements fell to 18.9% in 1999. This means that a number of the interviewees had already received their pensions in the period between the two surveys, because most pensions are given in a single payment in Taiwan.

Finally, workers' residences changed between 1996 and 1999. In 1996, 38.0% of workers lived in urban areas, 23.7% in town areas, and 38.3% in rural areas. Three years later, in 1999, this had changed, and 35.6% of workers lived in urban areas, 32.4% in town areas, and 32.0% in rural areas. There is a significant influx of older people from rural and urban areas to towns. The main reason the elderly are moving from rural areas to towns is probably for convenience, such as access to medical treatment. For those moving from urban areas to towns, this could be because of the relatively cheaper cost of living there compared to that in urban areas.

Table 5.2 Relative Time-Varying Covariates between 1996 and 1999

Unit: %

Year	1996	1999
Effective sample*	1732	966
Health status		
Excellent	20.6	20.1
Good	22.8	31.1
Average	32.6	32.7
Not so good	20.9	14.5
Poor	3.1	1.6
Marital status		
Married	84.1	86.3
Not married	2.7	2.3
Divorced	2.3	2.3
Separated	0.7	0.5
Widowed	10.2	8.6
Pension		
Eligible	26.7	18.9
Otherwise	73.3	81.1
Residence status		
Urban	38.0	35.6
Town	23.7	32.4
Rural	38.3	32.0

Note:

1. The 1996 total sample of SHLS data has 2462 observations, including 1072 people working full-time, 124 working part-time, 246 unemployed, 610 retired, and 410 never worked. The effective sample of duration model only has 1732 observations, including 610 retired (event observations) and 1122 continuing work (right-censored observations).

2. Further, the effective sample of duration model were changed from 1732 people in 1996 to 966 observations in 1999 (including 253 people retired and 713 continuing work), 39 unemployed, and 117 missing for moved, dead, or no answer.

5.3 Theoretical Framework of Labour Force Transition

A large amount of literature is devoted to understanding labour force transition patterns in Western industrialised countries. The theoretical framework for these studies conceptualises labour force participation as a trade-off between work and leisure within the constraints of economic and non-economic factors. Table 5.3 shows that previous studies examined how these are influenced by health shocks (see Schoenbaum, 1995; Bound et al, 1999; Disney et al, 2003), by marital changes (Hurd, 1988; Blau, 1998; Blau and Riphahn, 1999), by pension status (Boskin, 1977; Slade, 1987), or by other factors (Nickell, 1979; Ham and Rea, 1987).

Probit analysis (Long and Jones, 1980; Slade, 1982) and duration models (Cox, 1972; Lancaster, 1990; Collett, 1994) are used here to analyse the LFT of the middle-aged and elderly in Taiwan. These models consider a number of variables, including time-constant and time-varying covariates that pertain to individual transitional behaviour. For example, the time-constant covariates include Gender(G_i), Race (R_i), and Education (E_i) variables. The time-varying covariates include Health$\left[X_i^H(t)\right]$, Marital Status$\left[X_i^M(t)\right]$, Pension$\left[X_i^P(t)\right]$, and Residence$\left[X_i^U(t)\right]$ variables. The Age(A_i) variable might belong to one of the time-varying covariates, but everyone has the same natural increasing process for age in the different periods. Hence, this study would assume the Age variable is one of the time-constant covariates. The general model takes the form

$$L_i^T = L\left[A_i, G_i, R_i, E_i, X_i^H(t), X_i^M(t), X_i^P(t), X_i^U(t)\right] \quad i = 1,2,\cdots,n. \tag{5.1}$$

where n is the total number of observations. Each of these factors is discussed in detail, particularly the time-varying covariates. Four different transition factors at varying times, including Health, Marital Status, eligibility for a Pension, and Residence Status are considered to examine the decisions involved in LFT behaviour.

First, the relationship between Health and LFT is a dynamic process that can best be examined longitudinally. Health declines with age and this may require adaptation or cessation of work activities. For example, Schoenbaum (1995) used the Taiwan SHLS data to test the effect of health on LFT among the elderly using four different measures.[2] He concluded that health is a major determinant of LFT, regardless of how it is measured. Individuals in poor health are significantly more likely to retire than people in good health. Bound et al. (1999) used the US Health and Retirement Survey (HRS) data to analyse the dynamic effect of health on LFT, including labour force exit, job change and application for disability insurance. Specifically, they examined how the timing of health shocks affects LFT. Thus, they concluded that not just poor health, but also declining health, can help explain retirement behaviour. Disney et al. (2003) applied the analysis from Bound et al. (1999) and used the British Household Panel Survey from 1991 to 1998 to examine the role of ill health in retirement decisions in the UK. They found that individual health shocks are an important predictor of individual retirement behaviour. In summary, health shocks are an important consideration in retirement behaviour.

Second, the relationship between Marital Status and retirement decision is attracting growing attention in the field of LFT studies. For example, Hurd (1988) examined joint retirement behaviour in the US New Beneficiary Survey (NBS) data

[2] The four types of health measures considered in his paper include: (1) a summary measure of limitations on activities of daily living (ADL), such as shopping and lifting; (2) a summary measure of health conditions, such as stroke and dizziness; (3) a summary measure of mood and depression using the Centre for Epidemiologic Studies Depression Scale (CES-D) that can measure how people have been feeling in the past week; and (4) health indices, such as crude birth rate, crude death rate, life expectancy at birth (years) constructed using an instrumental variables framework.

and found a relatively high incidence of such behaviour. Blau (1998) analysed the dynamics of joint labour force behaviour of older couples in the United States. He used data from an 11-year longitudinal sample of men and unmarried women who were aged 58-63 in 1969 by the Retirement History Survey (RHS), and applied Hurd's (1988) results on the incidence of joint retirement to examine the determinants of joint retirement. He suggested any policy that increased the incentive for one member of a married couple to retire would have an additional effect on the LFT of the other spouse. Furthermore, Blau and Riphahn (1999) extended this analysis and used the German Socio-Economic Panel (GSOEP) data to examine the labour force behaviour of older married couples in Germany. They found that the probability of one spouse's retirement was much larger if the other spouse was not employed than if the other spouse was employed, and an unemployed member of a couple was much more likely to enter employment if the spouse was employed.

Third, social security or Pension benefits are important determinants of retirement among the elderly (see Quinn, 1977; Gordon and Blinder, 1980; Gustman and Steinmeier, 1982; Lazear, 1986; Slade, 1987). Previous studies have found that eligibility for a social security benefit or pension is associated with earlier retirement. In particular, Slade (1987) examined the role of state dependence in explaining retirement status among older males in the US. He used data from a 2-year longitudinal sample of men aged 58 to 62 collected in 1969 by the US Retirement History Study (RHS). He found that the level of the present value of social security benefits had a negative and significant effect on retirement and private pension eligibility had a positive and significant effect on retirement. However, as noted previously, Taiwan's pension system is different from those in Western industrialised countries. The government does not provide public or state pensions for the elderly;

instead, they focus on compulsory occupational pensions for employees working in government or for large companies. Therefore, Chapter 3 noted that the pre-condition of being eligible for a pension provides a strong incentive for people to participate in work. Chapter 4 also stated that workers eligible for a pension have a higher hazard rate of retirement. This chapter further tests the effects of workers' pension status change from eligible to ineligible. For instance, Table 5.2 shows that the proportion of workers eligible for a pension decreased from 26.7% in 1996 to 18.9% in 1999. In particular, most occupational pension benefits in Taiwan are paid in a single payment on retirement, and the retirees can receive a higher interest rate from their retirement payments.

Finally, other aspects such as employment opportunity, the level of physical demands, and type of skills and training required can affect workers' Residence Status and ability to adapt to changing LFT. For example, Nickell (1979) estimated the probability of leaving unemployment and used the database of unemployed males from the 1972 General Household Survey (GHS) in Britain. He examined the impact of local labour demand and how this changes over the course of an unemployment spell. Ham and Rea (1987) analysed the hazard rate of unemployment using micro data from the Canadian Employment and Immigration Longitudinal Labour Force File for the period 1975–1980. They examined the effect of provincial unemployment rates on expected unemployment duration. This chapter follows these ideas and investigates how a change in workers' residence status affects LFT. Table 5.2 shows that the proportion of workers living in town areas increased from 23.7% in 1996 to 32.4% in 1999, and the proportions of workers living in urban and rural areas both decreased during the same period. So, residence status might also be able to influence individual retirement behaviour.

Table 5.3 Literature Review of Labour Force Transition

Authors	Data Source	Models	Key Findings
1. Health Status:			
Schoenbaum (1995)	Survey of Health and Living Status of the Middle Aged and Elderly in Taiwan (SHLS).	Used discrete model to examine the effect of health on LFT among the elderly by four different measures.	Health is a major determinant of LFT, regardless of how it is measured. Individuals in poor health are significantly more likely to retire than people in good health.
Bound, Schoenbaum, Stinebrickner, and Waidmann (1999)	US Longitudinal Health and Retirement Survey (HRS).	Used multinomial models to analyse the dynamic relationship between health status and labour force transitions.	Health is a very important determinant of labour force patterns for older men and women.
Disney, Emmerson, and Wakefield (2003)	First eight waves of the British Household Panel Survey (BHPS) from 1991 to 1998.	Used ordered probit model to examine the role of ill health in retirement decisions in Britain.	Found that adverse individual health problems are an important predictor of individual retirement decisions.
2. Marital Status:			
Hurd (1988)	US New Beneficiary Survey (NBS).	Used models of retirement age to examine the determinants of joint retirement.	The results supported the idea that retirement of husbands and wives is a joint process.
Blau (1998)	US Retirement History Survey (RHS).	Used discrete choice model to examine the determinants of joint retirement and the effect of one spouse's labour force status on the labour force transitions of the other spouse.	The results revealed strong associations between the labour force transition probabilities of one spouse and the labour force status of the other spouse.

Blau and Riphahn (1999)	German Socio-Economic Panel (GSOEP).	Used discrete time variant of competing risks hazard model to estimate labour force transition probabilities.	The probability of one spouse exiting employment was much larger if the other spouse was not employed than if the other spouse was employed.
3. Pension Status:			
Boskin (1977)	US 1968-1972 Panel Study of Income Dynamics	Used multi-nominal logistic model to estimate the probability of retirement behaviours.	Found that two key policy parameters of the social security system, including the income guarantee and the implicit tax on earnings, exert an enormous influence on retirement decisions.
Slade (1987)	US 1969-1971 Longitudinal Retirement History Study (RHS).	Used retirement transition model to examine the role of state dependence in explaining the retirement status of older men.	State dependence is empirically important when there is considerable variation in the work sequence, as is the case after normal working years and before permanent retirement.
4. Other Factors:			
Nickell (1979)	UK 1970 General Household Survey (GHS)	Used duration model with cross-section data to investigate a number of questions concerning unemployment duration.	The expected duration is significantly influenced by the replacement ratio with an elasticity of about unity. The impact of benefit levels on the conditional probability of obtaining work is significant for the first 20 weeks.
Ham and Rea (1987)	Canadian 1975-1980 Employment and Immigration Longitudinal Labour Force File (EILLFF)	Used unemployment duration model to examine the influencing factors on the probability of leaving unemployment.	Entitlement provisions in the unemployment insurance programmes and demand conditions are found to have a significant effect on the probability of leaving unemployment.

5.4 Empirical Specifications

The traditional approach to analysing labour force behaviour is based on the individual labour supply model. In this model behaviour is determined by maximisation of a single utility function subject to their budget constraint in which income is pooled (Killingsworth and Heckman, 1986). This approach has obvious limitations that have been widely noted: the individuals may have different preferences, making it difficult to justify their preferences; and they may have different outside opportunities, making it difficult to justify pooling income (Blau and Riphahn, 1999). This section specifies two econometric models, including probit and duration models, to examine the decisions to exit employment, re-enter employment, and retire. Due to the limitation of SHLS data, the specification only allows for the possibility that income is not pooled, and does not assume that the parameters of the models can be interpreted as preference parameters.

5.4.1 Probit Analysis

The first approach follows Long and Jones (1980) and Slade (1982, 1987) and uses Maximum-likelihood probit estimates to examine the labour force transition (LFT) equations, including exit and re-entry in the labour market. Each person-year is treated as a distinct observation and the probability of LFT can be defined for an individual i of labour force participation. For instance, the empirical specifications of exit by probit analysis are described as follows.

According to the description in Chapter 2, reasons for stopping working at last job include: (1) reached mandatory retirement age; (2) health problems, could not continue working; (3) work didn't suit, wanted to change work environment; (4) company layoffs or relocation, was let go; (5) business failed, poor economy, profits too low; (6) unhappy with income, wanted to earn more; (7) family reasons: got married or to take care of children. Terms (1), (2), and (7) can be defined as retirement, and terms (3) to (6) belong to unemployment. Hence, the dependent variable of the labour force exit equation equals one if the individual worked in 1996 but did not work in 1999 and can be defined as

$$Exit = \begin{cases} 1 & \textit{if individual worked in } 1996 \textit{ and did not work in } 1999, \\ 0 & \textit{if otherwise.} \end{cases} \tag{5.1}$$

Then, the dependent variables can be written as a latent variable model

$$y_{it}^{*} = \beta' X_{it} + v_{it}. \tag{5.2}$$

where $y_{it} = \begin{cases} 1 & \textit{if } y_{it}^{*} > 0, \textit{ for exit.} \\ 0 & \textit{if otherwise, for non-exit.} \end{cases} \qquad i = 1,\ldots,N, \quad t = 1,\ldots,T_i.$

i is the individual and t is the time subscript. The set of parameters, β, captures the effect of the vector of explanatory variables, X_{it} on the labour force exit decision. The error term, v_{it}, is an independent realisation of a random variable with cumulative distribution function $\Phi(\cdot)$. The probability of labour force exit can thus be written as

$$P(y_{it} = 1) = P(y_{it}^{*} > 0) = \Phi(\beta' X_{it}). \tag{5.3}$$

Maximum likelihood estimates of the parameters are obtained by maximising the following likelihood function. In its general form, the likelihood function can be written as

$$L(\beta' \mid X_{it}) = \prod_{i=1}^{N}\prod_{t=1}^{T_i}\Phi(\beta' X_{it})^{y_{it}}\left[1-\Phi(\beta' X_{it})\right]^{1-y_{it}}. \tag{5.4}$$

Taking logs to obtain the Log-Likelihood function as follows.

$$\ln L(\beta' \mid X_{it}) = \prod_{i=1}^{N}\prod_{t=1}^{T_i}\left\{y_{it}\cdot\ln\Phi(\beta' X_{it}) + (1-y_{it})\cdot\ln\left[1-\Phi(\beta' X_{it})\right]\right\}. \tag{5.5}$$

The solution for maximum likelihood is obtained when the parameter value of $\hat{\beta}$ is obtained by a sequence of iterative processes on the log-likelihood function (5.5). The values of $\hat{\beta}$ correspond to the maximum log-likelihood by finding parameters $\hat{\beta}$.[3]

5.4.2 Duration Analysis

The second approach follows Blau (1998) and Blau and Riphahn (1999) and uses duration models to examine the determinants of retirement behaviour. Recalling the exponential model in Chapter 4, the hazard function of employment duration is specified as

$$h(t \mid x_i) = \lambda = e^{(\beta_0+\beta_i x_i)}. \tag{4.7}$$

And from the Weibull model in Chapter 4, the hazard function is defined to be

$$h(t \mid x_i) = \alpha t^{\alpha-1}\cdot\lambda = \alpha t^{\alpha-1}\cdot e^{(\beta_0+\beta_i x_i)}. \tag{4.8}$$

When the model adds time-varying covariates (TVCs) to the set of factors

[3] To solve models of these forms via Maximum Likelihood requires solution of first-order condition. The parameters that maximise the general log likelihood (5.5) are required for the probit model.

$$\frac{\partial \ln L(\beta' \mid X_{it})}{\partial \beta'} = \sum_{i=1}^{N}\sum_{t=1}^{T_i}\frac{y_{it}-\Phi(\beta' X_{it})}{\Phi(\beta' X_{it})\cdot\left[1-\Phi(\beta' X_{it})\right]}\cdot\phi(\beta' X_{it})\cdot X_{it} = 0.$$

determining labour force transition behaviour, the exponential function $\exp(\beta' x_i)$ is the relative hazard function, and a non-negative function of covariates x_i. By generalising this model to situations in which some explanatory variables are time-dependent, x_{it}, a vector of explanatory variables can be written with unknown coefficients β. First, the exponential regression model becomes

$$h(t \mid x_i, x_{it}) = \lambda = e^{(\beta_0 + \beta_i x_i + \beta x_{it})}. \tag{5.6}$$

Second, the Weibull regression model becomes

$$h(t \mid x_i, x_{it}) = \alpha t^{\alpha - 1} \cdot \lambda = \alpha t^{\alpha - 1} \cdot e^{(\beta_0 + \beta_i x_i + \beta x_{it})}. \tag{5.7}$$

Furthermore, if considering the effect of unobserved heterogeneity, the Weibull model becomes

$$h(t \mid x_i, x_{it}, u) = \alpha t^{\alpha - 1} \cdot \lambda = \alpha t^{\alpha - 1} \cdot e^{(\beta_0 + \beta_i x_i + \beta x_{it} + u)}. \tag{5.8}$$

where u is an i individual's unobserved heterogeneity term.

Third, the Cox model is also used to examine the determinants of retirement behaviour. In particular, the Cox hazard model might not be an appropriate proportional hazard model now. For instance, recalling the Cox proportional hazard model from Chapter 4, the retirement hazard at time of t for the i^{th} individuals in a study can be specified as:

$$h_i(t \mid x_i) = h_0(t) \exp(\beta' x_i) \tag{4.12}$$

where $h_0(t)$ is the baseline hazard function, and depends on t. It summarises the pattern of "duration dependence" common to all people. The exponential function $\exp(\beta' x_i)$ is the relative hazard function, and a non-negative function of covariates

x_i. By generalising this model to situations in which some explanatory variables are time-dependent, x_{it}, a vector of explanatory variables can be written with unknown coefficients β and the Cox regression model becomes

$$h(t \mid x_i,\ x_{it}) = h_0(t)\exp(\beta' x_i + \beta'' x_{it}). \tag{5.9}$$

In this model, the baseline hazard function $h_0(t)$ is interpreted as the hazard function of retirement for an individual for whom all the variables are zero at the time of origin and remain at the same time value. It is important to note that the values of the variables x_{it} in (5.9) depend on the time t, which implies that the relative hazard $h(t \mid x_{it}) / h_0(t)$ is also time dependent. This means that the hazard of retirement at time t is no longer proportional to the baseline hazard. As a result, model (5.9) is no longer a proportional hazard model. If we take the natural logarithm, the hazard rate of retirement becomes:

$$\ln h_i(t) = h_0(t) + \sum_{j=1}^{p} b_j x_{ij} + \underbrace{\sum_{r=p+1}^{k} b_r x_{ir}(t)}_{\textit{Time-varying Covariates}} \tag{5.10}$$

where n is the total number of observations, and k is the total number of variables, including the number of time-constant covariates p and the number of time-varying covariates $(k-p)$. $\ln h_i(t)$ represents the natural logarithm of the hazard rate of retirement, $h_0(t)$ is the constant hazard baseline, $\sum_{j=1}^{p} \beta_j x_{ij}$ is the effect of the time-constant covariates (in the case, taken at the time of the 1999 survey), and $\sum_{r=p+1}^{k} \beta_r x_{ir}(t)$ is the effect of the time-varying covariates.

For instance, assume that, for a person i, d represents the duration in employment before 1996. If $t < d$ and health status is 1, that represents health which is excellent before retirement. If $t \geq d$ and health status is 5, that represents workers being in poor health which may increase their hazard rate of retirement.

$$X_i^{H}(t) = \begin{cases} 1 & if \;\; t < d \\ 5 & if \;\; t \geq d. \end{cases} \tag{5.11}$$

Recall that the log-likelihood contribution for person i in the data structure is

$$\ln L_i = c_i \ln\left[h(T_i)\right] + \ln\left[S(T_i)\right]. \tag{5.12}$$

where $i's$ observed survival time is T_i and the censoring indicator $c_i = 1$ if $i's$ spell is complete (transition observed) and 0 if the spell is censored. But

$$\begin{aligned} \ln[S(T_i)] &= \ln\left[S(d)\frac{S(T_i)}{S(d)}\right] \\ &= \ln[S(d)] + \ln\left[\frac{S(T_i)}{S(d)}\right]. \end{aligned} \tag{5.13}$$

Thus the log of probability of survival until T = (log of the probability of survival to time d) + (log of the probability of survival to time T_i, conditional on entry at d). Table 5.4 presents the example of episode splitting. Here the multiple data record with $c_i = 0$, $t = d$ (a right censored episode), plus one new record with "delayed entry" at time d and censoring indicator c_i, has the same value as the original data. In the first episode and record, the time-varying covariate (TVC) takes on the value X_1 and the second record the TVC takes on the value X_2.

Table 5.4 Example of Episode Splitting

Data Set	Censoring Indicator	Entry Time	Survival Time	Time-Varying Covariates
Single Data	Record for i			
1996 Data Set	$c_i = 1$ *or* 0	0	T_i	-
Multiple Data	Records for i			
1996 Data Set	$c_i = 0$	0	d	X_1
1999 Data Set	$c_i = 1$ *or* 0	d	T_i	X_2

Note:
See Jenkins, S. P. (2003). "*Stata Programmes for Survival Analysis.*" Institute for Social and Economic Research, University of Essex.

5.5 Data Description

This chapter uses data from the second panel of the SHLS survey in 1996 and 1999. Before estimating the LFT data with the probit model and duration model, the different data sets are constructed for these two models. The sample utilised in the probit model contains workers who changed employment status: those who exited employment or re-entered the labour market between 1996 and 1999. By contrast, the sample utilised in the duration models examines the hazard rate of retirement.

5.5.1 Probit Analysis

In the probit analysis, two dependent variables are defined as in Section 5.4.1, including exit and re-entry variables. Explanatory variables are as follows: The Age groups include Age1 (aged 53 to 57), Age2 (aged 58 to 62), Age3 (aged 63 to 67), and Age4 (aged 68 to 73) in 1999. Next, the Gender variable is coded 1 for women and 0 for men. The Race variable can be separated into four groups, namely Race1 (Fujianese), Race2 (Hakka), Race3 (Mainlander), and Race4 (Aboriginal). The Education variable is years of schooling, namely Edu1 (Informal: 0 years of schooling), Edu2 (Primary School: 1 to 6 years of schooling), Edu3 (Junior and Senior High School: 7 to 12 years of schooling) and Edu4 (College and University: 13 to 17 years of schooling). On the other hand, some covariates indicate 1996 values (referred to here as "prior" variables). For instance, the Health variable is coded 1 for poor health, including "not so good" and "poor" health, and 0 for otherwise. The Marital Status variable is coded 1 for married and 0 for otherwise. The eligibility for a pension is coded 1 for those eligible and 0 otherwise. The Residence status includes workers living in urban, town and rural areas.

Furthermore, the other covariates prefixed by "Δ" denote the change in value between 1996 and 1999, such as $\Delta Health$, $\Delta Marital$, $\Delta Town$, and $\Delta Rural$. The $\Delta Health$ variable shows the worker's health becomes poorer. For example, the values of the health index include 1 (excellent), 2 (good), 3 (average), 4 (not so good), and 5 (poor). If the 1999 value of the health index minus the 1996 value is larger than zero, this means that worker's health becomes poorer. The $\Delta Marital$ variable represents marital status changes from married in 1996 to unmarried in 1999. However, most of the middle aged and elderly only married once. This might affect their LFT behaviour directly. The $\Delta Town$ variable shows the worker moves from town to non-town areas. The $\Delta Rural$ variable represents the worker moving from rural to non-rural areas. The full definitions of the variables by exit and re-entry cases and summary statistics are given in Tables 5.5 and 5.6 respectively.

Table 5.5 Descriptive Statistics of Variables by Exit Case

Variables	Description	Mean	Std. Err.
Exit	1 = Individual worked in 1996 and did not work in 1999, 0 = Otherwise.	.278	(.448)
AGE1	1 = Aged 53 to 57, 0 = Otherwise.	.387	(.487)
AGE2	1 = Aged 58 to 62, 0 = Otherwise.	.344	(.475)
AGE3	1 = Aged 63 to 67, 0 = Otherwise.	.210	(.407)
AGE4	1 = Aged 68 to 73, 0 = Otherwise.	.059	(.236)
GENDER	1 = Female, 0 = Male.	.310	(.463)
RACE1	1 = Fujianese, 0 = Otherwise.	.720	(.449)
RACE2	1 = Hakka, 0 = Otherwise.	.197	(.398)
RACE3	1 = Mainlander, 0 = Otherwise.	.067	(.251)
RACE4	1 = Aboriginal, 0 = Otherwise.	.016	(.126)
EDU1	1 = Informal education, 0 = Otherwise.	.213	(.409)
EDU2	1 = 1 to 6 years of schooling, 0 = Otherwise.	.481	(.499)
EDU3	1 = 7 to 12 years of schooling, 0 = Otherwise.	.220	(.415)
EDU4	1= 13 to 17 years of schooling, 0 = Otherwise.	.085	(.279)
PENSION	1 = Eligible for a pension, 0 = Otherwise.	.196	(.397)
HEALTH	1 = Poor health in 1996, 0 = Otherwise.	.147	(.354)
MARRIED	1 = Married in 1996, 0 = Otherwise.	.874	(.332)

TOWN	1 = Living in town in 1996, 0 = Otherwise.	.219	(.414)
RURAL	1 = Living in rural in 1996, 0 = Otherwise.	.412	(.492)
ΔHEALTH	1 = Health status become poorer from 1996 to 1999, 0 = Otherwise.	.326	(.469)
ΔMARRIED	1 = Marital status changed from married in 1996 to unmarried in 1999, 0 = Otherwise.	.022	(.146)
ΔTOWN	1 = Residence status changed from town in 1996 to non-town areas in 1999, 0 = Otherwise.	.062	(.241)
ΔRURAL	1 = Residence status changed from rural in 1996 to non-rural areas 1999, 0 = Otherwise.	.129	(.336)

Note:

According to the 1999 SHLS data, the overall exit sample has 1053 observations.

Table 5.6 Descriptive Statistics of Variables by Re-entry Case

Variables	Description	Mean	Std. Err.
RE-ENTRY	1 = Individual did not work in 1996 and worked in 1999, 0 = Otherwise.	.071	(.257)
AGE1	1 = Aged 53 to 57, 0 = Otherwise.	.170	(.376)
AGE2	1 = Aged 58 to 62, 0 = Otherwise.	.295	(.456)
AGE3	1 = Aged 63 to 67, 0 = Otherwise.	.340	(.474)
AGE4	1 = Aged 68 to 73, 0 = Otherwise.	.195	(.396)
GENDER	1 = Female, 0 = Male.	.558	(.497)
RACE1	1 = Fujianese, 0 = Otherwise.	.732	(.443)
RACE2	1 = Hakka, 0 = Otherwise.	.143	(.350)
RACE3	1 = Mainlander, 0 = Otherwise.	.109	(.312)
RACE4	1 = Aboriginal, 0 = Otherwise.	.016	(.124)
EDU1	1 = Informal education, 0 = Otherwise.	.332	(.471)
EDU2	1 = 1 to 6 years of schooling, 0 = Otherwise.	.449	(.498)
EDU3	1 = 7 to 12 years of schooling, 0 = Otherwise.	.171	(.377)
EDU4	1= 13 to 17 years of schooling, 0 = Otherwise.	.048	(.214)
PENSION	1 = Eligible in 1996, 0 = Otherwise.	.269	(.443)
HEALTH	1 = Poor health in 1996, 0 = Otherwise.	.339	(.474)
MARRIED	1 = Married in 1996, 0 = Otherwise.	.793	(.405)

TOWN	1 = Living in town in 1996, 0 = Otherwise.	.239	(.427)
RURAL	1 = Living in rural in 1996, 0 = Otherwise.	.356	(.479)
ΔHEALTH	1 = Health status become poorer from 1996 to 1999, 0 = Otherwise.	.293	(.456)
ΔMARRIED	1 = Marital status changed from married in 1996 to unmarried in 1999, 0 = Otherwise.	.045	(.208)
ΔTOWN	1 = Residence status changed from town in 1996 to non-town areas in 1999, 0 = Otherwise.	.071	(.257)
ΔRURAL	1 = Residence status changed from rural in 1996 to non-rural areas 1999, 0 = Otherwise.	.123	(.329)

Note:

According to the 1999 SHLS data, the overall re-entry sample has 706 observations.

5.5.2 Duration Analysis

The sample utilised in the duration analysis is different from the sample used in the previous analysis of exit from employment. The dependent variable is defined as the time of duration in employment. First, Table 5.7.1 shows the descriptive statistics of variables by cross-sectional data, the effective sample has 966 observations in 1999.

Second, Table 5.7.2 shows the descriptive statistics of variables by panel data. In particular, data from two waves of the second panel of SHLS survey data are merged to create a multiple data set by 1996 and 1999. The multiple data sets can be created in STATA.[4] The effective sample decreases from 966 (including 253 retirees, and 713 continuing work) to 915 observations (including 202 retirees, and 713 continuing work), and 51 retired observations missing for calculation by episode splitting. The average employment duration changes from 23.975 years by duration model with the 1999 cross-sectional data to 23.519 years by duration model with panel data between 1996 and 1999. The dependent variable is employment duration, and the explanatory variables are the same as those used in the previous models. In particular, for time-varying covariates such as health status, consider as an example the first observation records of all the information on the $ID = 40^{th}$ worker, who retired after 33 years employment, when health changed from 1 (excellent) to 5 (poor).

Observation Period	ID	Entry Time	Duration Employment	Retired	$X_{40}^{H}(t)$
1996-1999	40^{th}	30	33	1	5

[4] See 5.8 Appendix.

Now suppose the data on this particular worker looked like this,

Observation Period	ID	Entry Time	Duration Employment	Retired	$X_{40}^{H}(t)$
1996-1997	40^{th}	30	31	0	1
1997-1999	40^{th}	31	33	1	5

There are now two observations on this worker; these summarize the experiences of the worker over employment duration intervals $[30,31)$ and $[31,33]$. Note that for the first observation $Retired = 0$ as the worker did not retire at employment duration of 31 years. So assume that the covariates did not change at employment duration of 31 years. These two observations for $ID = 40^{th}$ worker record exactly the same information as the single observation did earlier. Continuing in this manner, when precisely health status changed, worker would make his/her decision to retire between 1998 and 1999. Therefore, workers being in poor health may increase their hazard rate of retirement.

Observation Period	ID	Entry Time	Duration Employment	Retired	$X_{40}^{H}(t)$
1996-1998	40^{th}	30	32	0	1
1998-1999	40^{th}	32	33	1	5

The descriptive statistics of the sample with panel data and summary statistics are given in Table 5.7.2. In this way, comment could have been included to the effect that the two sets of descriptive statistics in Table 5.7.1 and Table 5.7.2 are really very similar.

Table 5.7.1 Descriptive Statistics of Variables: Cross-Sectional Data

Variables	Description	Mean	Std. Err.
DURATION	1-55 years.	23.975	(14.540)
CENSOR	1 = Uncensored, 0 = Censored.	.262	(.440)
AGE1	1= Aged 53 to 57, 0 = Otherwise.	.394	(.489)
AGE2	1 = Aged 58 to 62, 0 = Otherwise.	.348	(.477)
AGE3	1 = Aged 63 to 67, 0 = Otherwise.	.201	(.401)
AGE4	1 = Aged 68 to 73, 0 = Otherwise.	.057	(.232)
GENDER	1 = Female, 0 = Male.	.295	(.456)
RACE1	1 = Fujianese, 0 = Otherwise.	.728	(.445)
RACE2	1 = Hakka, 0 = Otherwise.	.192	(.394)
RACE3	1 = Mainlander, 0 = Otherwise.	.066	(.249)
RACE4	1 = Aboriginal, 0 = Otherwise.	.014	(.120)
EDU1	1 = Informal education, 0 = Otherwise.	.208	(.406)
EDU2	1 = 1 to 6 years of schooling, 0 = Otherwise.	.482	(.499)
EDU3	1 = 7 to 12 years of schooling, 0 = Otherwise.	.223	(.416)
EDU4	1 = 13 to 17 years of schooling, 0 = Otherwise.	.087	(.282)
HEALTH	1 = Poor health, 0 = Otherwise.	.161	(.368)
MARRIED	1 = Married, 0 = Otherwise.	.863	(.344)

PENSION	1 = Eligible, 0 = Otherwise.	.189	(.392)
URBAN	1 = Living in urban areas, 0 = Otherwise.	.356	(.479)
TOWN	1 = Living in town areas, 0 = Otherwise.	.324	(.468)
RURAL	1 = Living in rural areas, 0 = Otherwise.	.320	(.467)

Note:

According to the 1999 SHLS data, the effective sample of duration model has 966 observations, including 253 retirees (event observations) and 713 continuing work (right-censored observations).

Table 5.7.2 Descriptive Statistics of Variables: Panel Data

Variables	Description	Mean	Std. Err.
DURATION	1-55 years.	23.519	(14.608)
CENSOR	1 = Uncensored, 0 = Censored.	.354	(.478)
AGE1	1= Aged 53 to 57, 0 = Otherwise.	.380	(.485)
AGE2	1 = Aged 58 to 62, 0 = Otherwise.	.351	(.478)
AGE3	1 = Aged 63 to 67, 0 = Otherwise.	.207	(.405)
AGE4	1 = Aged 68 to 73, 0 = Otherwise.	.063	(.242)
GENDER	1 = Female, 0 = Male.	.310	(.463)
RACE1	1 = Fujianese, 0 = Otherwise.	.727	(.446)
RACE2	1 = Hakka, 0 = Otherwise.	.184	(.388)
RACE3	1 = Mainlander, 0 = Otherwise.	.075	(.264)
RACE4	1 = Aboriginal, 0 = Otherwise.	.014	(.116)
EDU1	1 = Informal education, 0 = Otherwise.	.216	(.411)
EDU2	1 = 1 to 6 years of schooling, 0 = Otherwise.	.482	(.499)
EDU3	1 = 7 to 12 years of schooling, 0 = Otherwise.	.219	(.414)
EDU4	1 = 13 to 17 years of schooling, 0 = Otherwise.	.083	(.277)
HEALTH	1 = Poor health, 0 = Otherwise.	.165	(.371)
MARRIED	1 = Married, 0 = Otherwise.	.861	(.346)
PENSION	1 = Eligible, 0 = Otherwise.	.201	(.401)

URBAN	1 = Living in urban areas, 0 = Otherwise.	.358	(.479)
TOWN	1 = Living in town areas, 0 = Otherwise.	.316	(.465)
RURAL	1 = Living in rural areas, 0 = Otherwise.	.321	(.467)

Note:

As the model using panel data between 1996 and 1999, the effective sample becomes to 915 observations, including 202 retirees, and 713 continuing work. The details are discussed in page 251.

5.6 Empirical Results and Discussions

This section first uses probit analysis to estimate the probabilities of exit from employment and of re-entering the labour market between 1996 and 1999. The second subsection uses duration analysis to estimate the hazard rate of retirement with the 1999 cross-sectional data and the panel data between 1996 and 1999. These two analyses also show the empirical transition behaviours for men and women, respectively.

5.6.1 Probit Analysis

5.6.1.1 Transitions from Work to Non-work

The estimated result indicates that if $\beta_i > 0$, the probability of exiting employment will increase. If $\beta_i < 0$, then the probability of exiting employment will decrease. If $\beta_i = 0$, there is no effect on the probability of exiting employment.

Note that the estimates in Table 5.8.1 are quite precise. Age2 (aged 58 to 62), Age3 (aged 63 to 67), and Age4 (aged 68 to 73) variables have strong positive effects on the probability of exiting employment. Older workers have a higher probability of leaving the labour force. Furthermore, Table 5.8.2 indicates the marginal effect estimates that, holding other variables equal, people who are aged 58 to 62 (Age2) have a probability of leaving employment that is about 12.0 percentage points higher than those aged 53 to 57 (Age1), and those aged 63 to 67 (Age3) have a 17.7 percentage points higher probability, while those aged 68 to 73 (Age4) have a 10.8 percentage points higher probability than those aged 53 to 57 (Age1). Similarly, the Gender variable also has a strong positive effect on the probability of exiting employment. The marginal effect estimates indicate that female workers have a probability of leaving employment that is about 9.5 percentage points higher than males.

In contrast, the estimated coefficients of Edu3 (7 to 12 years of schooling) and Edu4 (13 to 17 years of schooling) variables are negative and statistically significant. Workers with better educational attainment have a lower probability of leaving the labour force. The marginal effect estimates show that workers with Edu3 (7 to 12 years of schooling) have a probability of leaving employment that is about 11.0 percentage points lower, and workers with Edu4 (13 to 17 years of schooling) about 19.4 percentage points lower than workers with informal education. These findings are similar to the results reported in Zimmer and Liu (1999), who suggested that people with a better education are more likely to work.

The demographic variable with by far the largest effect is poor health. Health96 has a positive and significant effect on the probability of entering early retirement. In particular, the $\Delta Health$ variable has a positive and significant effect on the probability of retirement transition. Hence, if workers' health declines, they have a higher probability of retirement. These findings are consistent with the results reported in Mete and Schultz (2002). To evaluate the other probabilities of exiting employment between 1996 and 1999, Table 5.8.1 shows that the coefficients for the $\Delta Married$ and $\Delta Rural$ variables are negative, which means that if a worker's marital status changed from married to unmarried, or his/her residence changed from rural to non-rural, they have a lower probability of leaving employment, but insignificantly so.

5.6.1.1.1 Gender Effects

Tables 5.8.1 and 5.8.2 also show the estimated results of male and female labour force transitions from work to non-work, respectively. The Age groups variables are strongly significant with a positive sign, implying that for both men and women all have a higher probability of leaving employment than the omitted category. For instance, the marginal effect estimates indicate that, holding other variables equal, men who are aged 58 to 62 (Age2) have a probability of exiting employment that is about 8.5 percentage points higher than for those aged 53 to 57 (Age1), and men aged 63 to 67 (Age3) have a 17.0 percentage points higher probability, and those aged 68 to 73 (Age4) have a 14.3 percentage points higher probability than men aged 53 to 57 (Age1). In addition, women who are aged 58 to 62 (Age2) have a probability of exiting employment that is about 18.8 percentage points higher than for those aged 53 to 57 (Age1), and those aged 63 to 67 (Age3) have a 17.8 percentage points higher probability, and women aged 68 to 73 (Age4) have a 2.8 percentage points higher probability than those aged 53 to 57 (Age1), but the coefficient of Age4 is insignificant. An explanation for this result may be that the employment opportunities for older workers are relatively limited and they therefore have a higher probability of leaving employment.

For the Race variables, male Hakka (Race2) and male Aboriginals (Race4) workers have a lower probability, and male Mainlanders (Race3) have a higher probability of exit from employment than male Fujianese (Race1), holding other variables equal. But the estimated coefficients are all insignificant. In contrast, female Hakka (Race2) and female Mainlanders (Race3) workers have a higher probability, and female Aboriginals (Race4) have a lower probability of exiting employment than female Fujianese (Race1), holding other variables equal. But the estimated coefficients are also all insignificant.

For the Education variable, the estimated results of men's and women's labour force transitions from work to non-work are shown in Tables 5.8.1 and 5.8.2, respectively. In general, workers with better education have a lower probability of exiting employment. For instance, male workers with primary education (Edu2) have a probability of exiting employment that is about 7.4 percentage points lower, Edu3 is 11.3 percentage points lower, and Edu4 17.3 percentage points lower than workers with informal education (Edu1), respectively. Furthermore, female workers with primary education have a probability of exiting employment that is about 3.5 percentage points lower, high school education is 16.3 percentage points lower, and university is 31.8 percentage points lower than workers with informal education, respectively. But only the variables Edu3 and Edu4 for men and Edu4 for women have a significant and negative effect on the probability of exiting employment.

Further, the sign of the Health variable is as expected: it has a positive effect on exit from employment. From the marginal effect estimates, holding other variables constant, male workers with poor health in 1996 have a probability of exit from employment that is about 15.5 percentage points higher than male workers with good health; and female workers with poor health in 1996 have a probability of exit from employment that is about 9.0 percentage points higher than female workers with good health. Moreover, as their health became poorer, they had a higher probability of leaving the labour market. In particular, male workers significantly have a probability of exit from employment that is about 6.6 percentage points higher than the omitted category. However, poor health has significant effects on the probability of exiting employment for males but insignificant effects for females. An explanation is that the male sample size has 727 observations and female only 326 observations in the probit model. Hence, it might partially attribute the statistical insignificance of poor health effects for women to the smaller sample size.

For the Married variable, Table 5.8.2 shows that married male workers have a probability of exit from employment that is about 7.4 percentage points lower than unmarried male workers, and if their marital status changes from married to unmarried they also have a probability of exit from employment that is about 19.3 percentage points lower than the omitted category. In contrast, married female workers have a probability of exiting employment that is about 4.8 percentage points higher than unmarried female workers, and if their marital status changes from married to unmarried they also have a probability of exit from employment that is about 4.0 percentage points lower than the omitted category. These results imply that men are likely to have greater responsibility for their family and are likely to be the sole economic support in the family, and although their marital status changes to unmarried, they still have a lower probability of exit from employment. This finding is consistent with the results reported in Chan and Stevens (2001). However, the coefficients are all insignificant.

Finally, Residence status can represent some employment opportunity for people. In general, workers living in town and rural areas have a lower probability of exit from employment. For instance, Table 5.8.2 shows that male workers living in town and rural areas respectively have a probability of exiting employment that is about 3.2 and 1.7 percentage points lower than the omitted category, holding other variables constant. But these coefficients are all insignificant, only if the residence areas changed from town to non-town is there a significant probability of exiting employment that is about 11.6 percentage points higher than the omitted category.

Table 5.8.1 Probit Coefficient Estimates of Exit Cases

Sample	Overall		Male		Female	
Exit	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.373***	(.107)	.284**	(.136)	.522***	(.180)
Age3	.548***	(.120)	.568***	(.149)	.494***	(.208)
Age4	.335***	(.200)	.476**	(.235)	.076	(.430)
Gender	.295***	(.103)	-		-	
Race2	-.086	(.115)	-.235	(.148)	.147	(.190)
Race3	-.024	(.195)	-.353	(.233)	.729	(.447)
Race4	-.047	(.353)	-.127	(.485)	-.184	(.531)
Edu2	-.155	(.116)	-.248	(.168)	-.097	(.170)
Edu3	-.341**	(.150)	-.376*	(.195)	-.452	(.301)
Edu4	-.601***	(.206)	-.578**	(.245)	-.883*	(.504)
Pension	.727***	(.113)	.798***	(.136)	.682***	(.221)
Health96	.402***	(.122)	.518***	(.155)	.250	(.206)
Married96	-.009	(.129)	-.246	(.188)	.134	(.184)
Town96	-.123	(.134)	-.108	(.167)	-.110	(.235)
Rural96	-.086	(.117)	-.055	(.144)	-.161	(.208)
Δ Health	.207**	(.095)	.219*	(.118)	.222	(.168)
Δ Married	-.219	(.301)	-.642	(.553)	-.110	(.387)
Δ Town	.309	(.199)	.386*	(.231)	-.059	(.437)
Δ Rural	-.223	(.150)	-.304	(.197)	-.150	(.242)
Constant	-.971***	(.203)	-.678**	(.271)	-.837***	(.270)
No. of subjects	1053		727		326	
Log likelihood	-562.828		-363.552		-191.164	
LR chi2 (19)	119.61***		91.08***		35.84***	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (18), respectively.

Table 5.8.2 Probit Marginal Effect Estimates of Exit Cases

Sample	Overall		Male		Female	
Exit	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.120***	(.034)	.085**	(.041)	.188***	(.064)
Age3	.177***	(.038)	.170***	(.045)	.178**	(.075)
Age4	.108*	(.064)	.143**	(.070)	.028	(.155)
Gender	.095***	(.033)	-		-	
Race2	-.028	(.037)	-.070	(.044)	.053	(.068)
Race3	-.008	(.063)	-.106	(.070)	.263	(.161)
Race4	-.015	(.114)	-.038	(.145)	-.066	(.191)
Edu2	-.050	(.037)	-.074	(.050)	-.035	(.061)
Edu3	-.110**	(.048)	-.113*	(.058)	-.163	(.108)
Edu4	-.194***	(.066)	-.173**	(.073)	-.318*	(.181)
Pension	.235***	(.036)	.239***	(.041)	.245***	(.080)
Health96	.130***	(.039)	.155***	(.046)	.090	(.074)
Married96	-.003	(.042)	-.074	(.056)	.048	(.066)
Town96	-.040	(.043)	-.032	(.050)	-.039	(.085)
Rural96	-.028	(.038)	-.017	(.043)	-.058	(.075)
Δ Health	.067**	(.031)	.066*	(.035)	.080	(.060)
Δ Married	-.071	(.097)	-.193	(.166)	-.040	(.139)
Δ Town	.100	(.064)	.116*	(.069)	-.021	(.157)
Δ Rural	-.072	(.049)	-.091	(.059)	-.054	(.087)
Predicted Probability	.257		.225		.325	

Note:

Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

5.6.1.2 Transitions from Non-work to Work

One of the important contributions of this thesis is the use of the SHLS data set to estimate the probability of re-entry into employment. The estimation results are shown in Table 5.9.1. At the same time, the marginal effect estimates of re-entry for the sub-sample by gender are also presented in Table 5.9.2.

First, the probability of re-entering the labour market for the benchmark individual and other individuals with different demographic circumstances can be calculated. For instance, assume all explanatory variables take a value of zero, and then the benchmark individual in all cases is a Fujianese man aged 53 to 57, who has informal education, and we assume his health, marital status, and residence are not changed. This benchmark value reflects the constant variable. Table 5.9.1 shows the benchmark estimates lead to a probability estimate in the probit model of

$$P(d_{it} = 1) = \Phi(-0.367) = 0.357.$$

The effects on the probability of moving from non-work to work for different demographic circumstances can be calculated as follows. For example, holding other variables constant, the probability of re-entry into employment for workers aged 58 to 62 (Age2) is

$$P(d_{it} = 1) = \Phi(-0.367 - 0.631) = 0.159.$$

This means that a retired worker aged 58 to 62 has a lower probability of re-entering the labour market, holding other variables equal. Furthermore, the Edu3 (7 to 12 years of schooling) case can be calculated as

$$P(d_{it} = 1) = \Phi(-0.367 - 1.077) = 0.074.$$

This means that workers with middle levels of education have a lower probability of re-entry into employment. Therefore, if the estimated coefficients are negative, the probability of re-entering employment decreases, but if the estimated coefficients are positive, the probability of re-entering employment increases.

In general, the empirical results in Table 5.9.1 confirm the theoretical expectations in Section 5.3, with about half of the regressors being statistically significant. The estimated coefficients of Age2 (aged 58 to 62), Age3 (aged 63 to 67), Age4 (aged 68 to 73), and Gender variables are significantly negative. Elderly workers and female workers are less likely to re-enter the labour market. Furthermore, the estimated coefficients of Edu2 (1 to 6 years of schooling), Edu3 (7 to 12 years of schooling), and Edu4 (13 to 17 years of schooling) variables are also negative and statistically significant. Workers with better education have a lower probability of re-entering employment. Actually, the elderly have relatively fewer opportunities to re-enter the labour market again, although they have better education. The estimated coefficients of Race2 (Hakka) and Race3 (Mainlander) variables are positive. That means Hakka and Mainlander workers are more likely to re-enter the labour market, but all insignificantly so. In particular, there are no observations of Race4 variable in the re-entry sample and this is omitted from these regressions.

Furthermore, Table 5.9.1 shows that the coefficients for $\Delta Health$, $\Delta Married$, and $\Delta Rural$ variables have a negative effect. This means that as workers' health becomes poorer, workers' marital status changes from married to unmarried, particularly divorced and widowed, or residence changes from rural to non-rural areas, they have a lower probability of re-entering employment, but only the coefficient of $\Delta Rural$ is significant. In contrast, the estimated coefficient of $\Delta Town$ variable has a positive effect on re-entry into the labour market. That means as residence changes from town to non-town areas, workers are more likely to re-enter the labour market, but the coefficient is insignificant.

5.6.1.2.1 Gender Effects

Table 5.9.2 also shows marginal effect estimates of re-entry for men and women respectively. The Age group variables are strongly significant with a negative sign, implying that Age3 and Age4 for men, and Age2 and Age3 for women have a lower probability of re-entering employment than the omitted category. For instance, the marginal effect estimates indicate that, holding other variables equal, men who are aged 63 to 67 (Age3) have a probability of re-entering employment that is about 15.4 percentage points lower than for those aged 53 to 57 (Age1), and men aged 68 to 73 (Age4) have a probability 18.6 percentage points lower than those aged 53 to 57 (Age1). In addition, women who are aged 58 to 62 (Age2) have a probability of re-entering employment that is about 8.2 percentage points lower than for those aged 53 to 57 (Age1), and those aged 63 to 67 (Age3) have a 10.2 percentage points lower probability than those aged 53 to 57 (Age1). Further, there are no responses for the Age4, Race3, and Race4 variables in the re-entry cases by female sample. Hence, these variables are omitted in these regressions for the probability of re-entry into employment.

For the Education variables, only female workers with primary education (Edu2: 1 to 6 years of schooling) have a significantly lower probability of re-entering employment, about 5.2 percentage points lower than the omitted category, but the other coefficients of the education variables are insignificant for men.

For the Pension variable, Table 5.9.2 shows that male workers eligible for a pension have a lower probability of re-entering employment, about 3.6 percentage points lower than the omitted category. In contrast, female workers eligible for a pension have a higher probability of re-entering employment, about 3.8 percentage points higher than the omitted category. However, the above estimates of the Pension variables are all insignificant for men and women.

Further, the sign of the Health variable is as expected: it has a negative effect on the probability of re-entering employment. From the marginal effect estimates, holding other variables constant, male workers with poor health in 1996 have a probability of re-entering employment that is about 1.8 percentage points lower than male workers with good health; and female workers with poor health in 1996 have a probability of re-entering employment that is about 3.2 percentage points lower than female workers with good health. Moreover, as their health declines, male workers have a higher probability of re-entering employment, about 1.6 percentage points higher than the omitted category; female workers have a lower probability of re-entering employment, about 0.4 percentage points lower than the omitted category. But the estimates of the Health variables are all insignificant for men and women.

Finally, the Married, Town, Rural, $\Delta Married$, $\Delta Town$, and $\Delta Rural$ variables are all insignificant for men and women (Table 5.9.2). An explanation is that some variables in the 1999 SHLS survey missing for "don't know" or "not answer", such as the Race4, Town96, and Rural96 variables. The male samples were decreased from 310 to 263 observations, and female samples were decreased from 385 to 285 observations, so the above variables might be insufficient to identify a statistically significant effect on the re-entry cases.

Table 5.9.1 Probit Coefficient Estimates of Re-entry Cases

Sample	Overall		Male		Female	
Exit	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	-.631***	(.209)	-.479	(.312)	-.938***	(.325)
Age3	-1.048***	(.235)	-1.007***	(.330)	-1.165***	(.364)
Age4	-1.447***	(.341)	-1.216***	(.410)	-	
Gender	-.850***	(.193)	-		-	
Race2	.013	(.264)	-.503	(.578)	.269	(.333)
Race3	.014	(.351)	.091	(.395)	-	
Race4	-		-		-	
Edu2	-.416**	(.192)	-.203	(.290)	-.597**	(.298)
Edu3	-1.077***	(.322)	-.691*	(.392)	-	
Edu4	-.851*	(.451)	-.557	(.518)	-	
Pension	-.032	(.234)	-.235	(.304)	.439	(.433)
Health96	-.225	(.193)	-.118	(.269)	-.364	(.310)
Married96	-.190	(.207)	.001	(.309)	-.473	(.313)
Town96	-.286	(.244)	-.324	(.313)	-.355	(.446)
Rural96	-.007	(.205)	-.098	(.297)	.143	(.313)
Δ Health	-.002	(.188)	.104	(.264)	-.050	(.302)
Δ Married	-.065	(.481)	-		.086	(.527)
Δ Town	.176	(.341)	.262	(.418)	.266	(.655)
Δ Rural	-1.068**	(.436)	-		-.783	(.543)
Constant	-.367	(.362)	-.010	(.515)	-.103	(.428)
No. of subjects	695		263		285	
Log likelihood	-147.850		-81.435		-60.045	
LR chi2 (18)	63.80***		27.89**		19.52	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Due to some variables in the 1999 SHLS survey missing for "don't know" or "not answer", such as the Race4, Town96, and Rural96 variables, so the male sample was decreased from 310 to 263 observations and the LR chi2 of male is LR chi2 (15). The female sample was decreased from 385 to 285 observations and the LR chi2 of female is LR chi2 (13).

Table 5.9.2 Probit Marginal Effect Estimates of Re-entry Cases

Sample	Overall		Male		Female	
Exit	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	-.053***	(.019)	-.073	(.049)	-.082***	(.029)
Age3	-.089***	(.021)	-.154***	(.052)	-.102***	(.031)
Age4	-.123***	(.029)	-.186**	(.063)	-	
Gender	-.072***	(.017)	-		-	
Race2	.001	(.022)	-.077	(.087)	.023	(.029)
Race3	.001	(.030)	.014	(.060)	-	
Race4	-		-		-	
Edu2	-.035**	(.016)	-.031	(.045)	-.052**	(.025)
Edu3	-.091***	(.026)	-.106	(.058)	-	
Edu4	-.072*	(.038)	-.085	(.079)	-	
Pension	-.003	(.020)	-.036	(.046)	.038	(.037)
Health96	-.019	(.016)	-.018	(.041)	-.032	(.027)
Married96	-.016	(.017)	.001	(.047)	-.041	(.027)
Town96	-.024	(.021)	-.050	(.048)	-.031	(.039)
Rural96	.001	(.017)	-.015	(.046)	.013	(.027)
Δ Health	-.001	(.016)	.016	(.040)	-.004	(.026)
Δ Married	-.006	(.041)	-		-.008	(.046)
Δ Town	.015	(.029)	.040	(.064)	.023	(.057)
Δ Rural	-.091***	(.035)	-		-.068	(.045)
Predicted Probability	.039		.083		.041	

Note:

Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.

5.6.2 Duration Analysis

This section presents a detailed empirical analysis of labour force transition behaviour. First, using the 1999 cross-sectional data, the hazard rate of retirement can be estimated by the exponential model, Weibull model, and Cox hazard model. Second, the duration models can be extended with panel data and add time-varying covariates to capture individuals' retirement behaviour. The above analyses also consider the empirical transition behaviours for men and women, and the effects of unobserved heterogeneity.

5.6.2.1 Exponential and Weibull Models: Cross-Sectional Data

Recall from the analysis in Chapter 4 that the hazard functions of employment duration in the exponential and Weibull models are specified respectively as

$$h(t \mid x_i) = \lambda = e^{(\beta_0 + \beta_i x_i)}. \tag{4.7}$$

$$h(t \mid x_i) = \alpha t^{\alpha-1} \cdot \lambda = \alpha t^{\alpha-1} \cdot e^{(\beta_0 + \beta_i x_i)}. \tag{4.8}$$

The estimated result indicates that if $\beta_i > 0$, the hazard rate of retirement will increase. If $\beta_i < 0$, then the hazard rate of retirement will decrease. If $\beta_i = 0$, there is no effect on the hazard rate of retirement. Furthermore, the hazard function in the Weibull model increases in duration if $\alpha > 1$, decreases if $\alpha < 1$, and remains constant if $\alpha = 1$. The last, equality, is exactly the same as the exponential case.

First, comparing the estimated results of two cross-sectional data analyses, for the Exponential model, most estimated results of the 1999 SHLS data in Table 5.10.1 are similar and consistent with the reports of the 1996 SHLS data in Table 4.4.1. For instance, the estimated coefficients of Age2 (aged 58 to 62), Age3 (aged 63 to 67), Gender, and Race3 (Mainlander) variables are significantly positive. Older workers, female workers, and Mainlander workers are more likely to retire than otherwise. In contrast, the estimated coefficients of Edu3 (7 to 12 years of schooling) and Edu4 (13 to 17 years of schooling) variables have a significant negative effect on retirement hazard. This means that workers with better educational attainment have a lower hazard rate for retirement. However, there is also a little change for a few variables between these two waves. For example, the Race2 and Married variables change from a negative effect on the retirement hazard to a positive significant effect, but insignificantly. This implies that Hakka workers and married workers might gradually change their retirement decisions after three years, but the coefficients are insignificant. The Health variable significantly has a higher retirement hazard in 1996, but insignificantly in 1999. This suggests that workers with poor health have a higher hazard rate for retirement initially. After three years, they might give more consideration to other factors, so the coefficient became insignificant. In contrast, the Pension variable insignificantly has a higher retirement hazard in 1996, but significantly in 1999. This means that workers gradually reached retirement age after three years, so the pension variable might have a significant positive effect on retirement.

Second, extending the Exponential model to the Weibull model, most estimated results of the 1999 SHLS data in Table 5.10.2 are similar and consistent with the reports of the 1996 SHLS data in Table 4.5.1. Especially, the hazard rates of the Weibull model all have positive duration dependence, in 1996, $\alpha = 1.537 > 1$; and in 1999, $\alpha = 1.350$. This seems to prove that the hazard rate is increasing over elapsed employment duration. As employment duration gets longer, the hazard rate increases and workers are more likely to retire.

Next, comparing the estimated results of gender effects with previous analyses in Table 4.4.1, some results are similar, but some results are different. For instance, Table 5.10.1 shows the sample of women that the estimated coefficients of Age2 (aged 58 to 62) and Age3 (aged 63 to 67) variables have a significantly positive effect, and the variables of Town and Rural have a negative effect on retirement hazard by the Exponential model. This means that older female workers were still more likely to retire, and female workers living in town and rural areas were also less likely to retire. However, some variables have been changed. For example, the estimated coefficient of Health variable by women had a significantly positive effect in Table 4.4.1 and changed to insignificantly negative effect in Table 5.10.1. This implies that female workers with poor health were more likely to retire in 1996, after three years, they became to be less likely to retire. In particular, the Taiwanese government provided the National Health Insurance (NHI) programmes for people from 1995. Women might expect more supports from the NHI for their health and delay their retirement behaviour. In contrast, the estimated coefficient of Pension variable by women had a significantly negative effect in Table 4.4.1 and changed to have a significantly positive effect in Table 5.10.1. This implies that female workers eligible for a pension were less likely to retire in 1996, after three years, they were more likely to retire. Perhaps, female workers looked forward to receiving a new National Pension Programmes from 2000.[5]

Other interesting results are shown in Tables 5.10.1 and 5.10.2, which use sub-samples for men and women to estimate the hazard rate of retirement. For women, Table 5.10.1 shows that the estimated coefficients of Age2 (aged 58 to 62), Age3

[5] New National Pension Programmes was promoted from 1990s. However, Taiwan had a biggest earthquake in 1999, so the government delayed this programmes until now.

(aged 63 to 67), Race2 (Hakka), Race3 (Mainlander), and Pension variables have a significantly positive effect on retirement hazard by the Exponential model. This means that older female workers, female Hakka workers, female Mainlander workers, and female workers eligible for a pension were more likely to retire. However, there are also some different effects between males and females. For example, the Health variable has a positive effect on retirement hazard for males and a negative effect for females, but all insignificantly. A possible explanation is that females might do more unpaid housework and other family-related work and they might not like to complain if their health declines with age. Moreover, the estimated coefficient of the Race2 (Hakka) variable has a negative effect for men, and a positive effect for women. That means female Hakka workers are more likely to retire than males, but only the coefficient of female Hakka is significant. Further, the estimated coefficients of Town and Rural variables have a negative effect for men and women, but only the coefficient of female rural workers is significant. It seems that female workers living in rural areas have more job opportunities and are less likely to retire than those living in non-rural areas.

Table 5.10.1 Exponential Model: Cross-Sectional Data

Sample	Overall		Male		Female	
Duration	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.288*	(.167)	.177	(.221)	.503*	(.260)
Age3	.421**	(.180)	.275	(.233)	.639**	(.301)
Age4	.334	(.274)	.415	(.328)	.166	(.569)
Gender	.727***	(.150)	-		-	
Race2	.012	(.166)	-.302	(.226)	.505*	(.258)
Race3	.478*	(.258)	.203	(.307)	.962*	(.508)
Race4	-.002	(.592)	-.015	(.734)	-.182	(1.030)
Edu2	-.225	(.162)	-.148	(.239)	-.463*	(.239)
Edu3	-.551**	(.219)	-.430	(.286)	-.619	(.407)
Edu4	-.725**	(.298)	-.463	(.351)	-1.971*	(1.019)
Health	.210	(.157)	.261	(.196)	-.027	(.275)
Married	.081	(.184)	-.201	(.274)	.229	(.252)
Pension	1.114***	(.148)	1.212***	(.182)	.899***	(.276)
Town	-.061	(.163)	-.036	(.207)	-.028	(.274)
Rural	-.189	(.165)	-.038	(.205)	-.475*	(.288)
Constant	-5.017***	(.296)	-4.794***	(.413)	-4.364***	(.354)
No. of subjects	966		681		285	
No. of retirees	253		160		93	
Log likelihood	-670.742		-414.547		-248.673	
LR chi2 (15)	111.86***		68.52***		36.26**	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

Table 5.10.2 Weibull Model: Cross-Sectional Data

Sample	Overall		Male		Female	
Duration	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.228	(.168)	.070	(.223)	.494*	(.260)
Age3	.306*	(.183)	.011	(.237)	.627**	(.302)
Age4	.271	(.273)	.291	(.323)	.152	(.571)
Gender	.813***	(.152)	-		-	
Race2	.009	(.166)	-.330	(.227)	.509**	(.259)
Race3	.606**	(.259)	.491	(.307)	.972*	(.509)
Race4	.044	(.593)	.035	(.739)	-.165	(1.031)
Edu2	-.229	(.163)	-.170	(.240)	-.467*	(.240)
Edu3	-.582***	(.220)	-.439	(.289)	-.631	(.407)
Edu4	-.753**	(.297)	-.420	(.351)	-2.007**	(1.021)
Health	.184	(.157)	.212	(.199)	-.032	(.276)
Married	.081	(.185)	-.271	(.275)	.233	(.253)
Pension	1.215***	(.150)	1.328***	(.184)	.925***	(.279)
Town	-.059	(.164)	-.013	(.209)	-.029	(.275)
Rural	-.241	(.166)	-.079	(.205)	-.492*	(.290)
Constant	-6.171***	(.382)	-6.953***	(.557)	-4.530***	(.451)
/ln_α	.300***	(.053)	.513***	(.067)	.054	(.088)
α	1.350***	(.072)	1.670***	(.113)	1.055	(.093)
1/α	.741***	(.039)	.599***	(.040)	.948	(.083)
No. of subjects	966		681		285	
No. of retirees	253		160		93	
Log likelihood	-656.753		-390.816		-248.490	
LR chi2 (15)	124.31***		79.14***		36.61***	

Notes:

1. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

5.6.2.2 Frailty Model: Cross-Sectional Data

This section uses the 1999 cross-sectional data for examining the effects of unobserved heterogeneity on retirement behaviour. First, without unobserved heterogeneity, most results have the expected effects on retirement as shown in Table 5.10.3. The estimated coefficients of those with Race3 and Pension variables are positive and statistically significant and have higher hazard rates *ceteris paribus*. In contrast, the estimated coefficients for Edu2, Edu3, Edu4, and Rural variables are significantly negative. The estimate for the shape parameter is $\alpha = 1.313$ suggesting an increasing hazard over time.

Second, the frailty model is assumed to follow a gamma distribution with mean 1 and variance equal to theta(θ). The estimate of theta is 0.144 and it is significant at the 10% significance level. The likelihood ratio test for the inclusion of theta is provided at the bottom of the output and yields a chi-square value of 19.21 with 1 degree of freedom yielding a highly significant p-value of 0.000. Further, the estimated coefficients on the regressors Race3, Edu3, Edu4, and Pension are a little bit larger in magnitude than the corresponding coefficients in the reference model. In particular, the estimated coefficients of Race3 and Pension have a significantly positive effect on retirement duration, other things being equal. This implies that Mainlander workers and workers with eligible pension have a higher hazard rate of retirement. In contrast, the coefficients of Edu3 and Edu4 variables have a significantly negative effect on retirement duration. This means that workers with better education have a lower hazard rate of retirement. Finally, the Weibull distribution shape parameter $\alpha = 1.347$ is also a little bit larger in the frailty model than $\alpha = 1.313$ in the reference model. Hence, the neglected heterogeneity may let the bias underestimates duration dependence.

Table 5.10.3 Frailty Model: Cross-Sectional Data

	Without Unobserved Heterogeneity		With Gamma-Heterogeneity	
Variables	Coefficient	Standard Error	Coefficient	Standard Error
Age2	.122	(.166)	.221	(.167)
Age3	.146	(.180)	.294	(.181)
Age4	.132	(.271)	.259	(.272)
Race2	.041	(.165)	.012	(.166)
Race3	.518**	(.258)	.600**	(.257)
Race4	.129	(.593)	.051	(.593)
Edu2	-.494***	(.153)	-.248	(.163)
Edu3	-.893***	(.209)	-.606***	(.220)
Edu4	-1.062***	(.289)	-.776***	(.297)
Married	-.174	(.179)	.062	(.185)
Health	.173	(.157)	.183	(.157)
Pension	1.153***	(.149)	1.209***	(.148)
Town	-.095	(.163)	-.062	(.163)
Rural	-.290*	(.164)	-.244	(.165)
Constant	-5.234***	(.332)	-5.650***	(.435)
/ln_α	.272***	(.053)	.297***	(.051)
/ln_the			-1.939*	(1.049)
α	1.313***	(.070)	1.347***	(.069)
1/α	.761***	(.041)	.742***	(.038)
theta			.144*	(.151)
No. of subjects	966		966	
No. of retirees	253		253	
Log likelihood	-670.345		-660.739	
LR chi2 (14)	97.13***		97.98	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. In particular, Likelihood-ratio test of theta=0: chibar 2 (01) = 19.21, Prob >= chibar2 = 0.000.

5.6.2.3 Exponential and Weibull Models: Panel Data

Bringing time-varying covariates (TVCs) into our duration models, the exponential regression model becomes

$$h(t \mid x_i, x_{it}) = \lambda = e^{(\beta_0+\beta_i x_i+\beta' x_{it})}. \tag{5.8}$$

And the Weibull regression model becomes

$$h(t \mid x_i, x_{it}) = \alpha t^{\alpha-1} \cdot \lambda = \alpha t^{\alpha-1} \cdot e^{(\beta_0+\beta_i x_i+\beta'' x_{it})}. \tag{5.9}$$

The estimation results of the exponential and Weibull hazard models are reported in Tables 5.10.4 and 5.10.5. A positive coefficient means that this particular variable has positive effects on retirement, while a negative coefficient implies that a worker works longer, postponing retirement.

First, for the exponential model, Table 5.10.4 shows that the estimated coefficients of Gender, and Race3 (Mainlander) variables are significantly positive. Female workers and Mainlander workers have a higher hazard rate of retirement than otherwise. In contrast, the estimated coefficients of Edu3 (7 to 12 years of schooling) and Edu4 (13 to 17 years of schooling) variables are negative; workers with better educational attainment have a lower hazard rate of retirement. These are similar to the results reported in the cross-section analysis in 1999.

Further, for the time-varying covariates, the estimated coefficient of the *Health*(t) variable has a positive significant effect on retirement. That implies that a worker being in poor health increases the hazard rate of retirement other things being equal. This is consistent with the results reported in Bound et al. (1999), who used ordered probit models to examine how the timing of "health shocks" affects retirement, particularly for elderly people retiring in the US. They found that declining health was an explanation for this retirement behaviour. In addition, Disney et al. (2003) applied a two-stage method to examine the role of ill health on retirement. They also found that poor health reduced the probability of continuing to work and a change in health was a major determinant of retirement in the UK.

The estimated coefficient of the *Pension*(t) variable also has a positive significant effect on retirement. That means workers whose pension status becomes realisable have a higher hazard rate of retirement. This is consistent with the results reported in Chapter 4 using the cross-section analysis. In particular, most workers have a strong incentive to receive their pension benefits earlier and invest this lump sum retirement payment at a high interest rate.

Moreover, the variable *Married*(t) also has a positive effect on retirement hazard. That means workers whose marital status becomes unmarried have a higher hazard rate of retirement, but the estimated coefficient is insignificant for retirement. Comparing the results reported in Blau (1998) and Blau and Riphahn (1999), Blau (1998) analysed the dynamics of joint labour force behaviour of older married couples in the US and found that married couples tended toward joint retirement as there would no longer be an incentive for one spouse to remain employed. Blau and Riphahn (1999) analysed the dynamic effects on retirement of older workers and found that one member of a couple is more likely to enter employment if their spouse is employed than if their spouse is not employed.

In contrast, the estimated coefficients of *Town*(t) and *Rural*(t) variables have a negative effect on retirement, meaning that workers living in town and rural areas have a lower hazard rate of retirement than otherwise, but again the coefficients are insignificant under an exponential hazard.

Second, for the Weibull model, Table 5.10.5 shows that most of the parameter values resemble the results reported in the exponential model. For example, the estimated coefficients of Gender and Race3 (Mainlander) variables show significantly positive effects on retirement. Female workers and Mainlander workers are more likely to retire than otherwise. For the time-varying covariates analysis, the estimated coefficients of $Health(t)$ and $Pension(t)$ variables have positive and significant effects on retirement. That is, workers being in poor health and whose pension becomes realisable have a higher hazard rate of retirement other things being equal. In addition, the estimated coefficients of $Married(t)$ and $Town(t)$ variables also have a positive but insignificant effect on retirement hazard. In particular, $\alpha = 1.336 > 1$ and $1 \le t \le 55$, which indicates hazard rates have positive duration dependence, $dh(t)/dt > 0$. This means that as employment duration gets longer, hazard rates may increase and workers are more likely to retire.

5.6.2.2.1 Gender Effects

Tables 5.10.4 and 5.10.5 report the results for males and females of the Exponential and Weibull models. In particular, some estimated coefficients for females have a significant effect on retirement hazard, but insignificant for males. For instance, Table 5.10.4 shows that the estimated coefficients of Age2 (aged 58 to 62), Age3 (aged 63 to 67), and Race3 (Mainlander) variables for females are significantly positive, but insignificant for males. This means that female workers prefer to retire earlier, significantly between ages 58 and 67, and male workers might retire later. In addition, the female Mainlander worker variable has a more significant positive effect on retirement than that of males. This is different from the result reported in Chapter 4.

Furthermore, the estimated results of *Health*(t) and *Pension*(t) variables for the male sample have a positive significant effect on retirement. In contrast, the *Health*(t) variable for female sample has a negative insignificant effect, and the *Pension*(t) variable for female sample has a positive significant effect. This might be because the numbers of males with poor health are less than the female sample, or who eligible for a pension are greater than the female sample. These two speculations can be proved by consideration of appropriate sample means. That is, the sample mean of males with poor health is (20.4%) less than females (27.3%), and the sample mean of males eligible for a pension is (29.9%) and greater than females (9.7%).

Finally, for the Weibull model, Table 5.10.5 shows that most of the parameter values resemble the results reported in the Exponential model in Table 5.10.4. The estimated coefficients for females have a more significant effect on retirement hazard than for males. For example, the estimated coefficients of Age2 (aged 58 to 62), Age3 (aged 63 to 67), Race3 (Mainlander), and Edu4 (13 to 17 years of schooling) variables for females have a significant effect on retirement, but insignificant for males. In contrast, the duration dependence is significantly positive for males, $\alpha = 1.693$; but insignificant for females, $\alpha = 1.029$. Therefore, as employment duration gets longer, hazard rates may increase, that males might have a larger and significant effect on the retirement hazard than females.

Table 5.10.4 Exponential Model: Panel Data

Sample	Overall		Male		Female	
Duration	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
TCC						
Age2	.283	(.186)	.105	(.250)	.570**	(.284)
Age3	.326	(.204)	.157	(.267)	.578*	(.329)
Age4	.348	(.307)	.481	(.380)	.096	(.607)
Gender	.885***	(.167)	-		-	
Race2	-.169	(.194)	-.425	(.272)	.241	(.295)
Race3	.579**	(.286)	.165	(.346)	1.347**	(.607)
Race4	-.278	(.724)	-.678	(1.026)	-.184	(1.034)
Edu2	-.186	(.182)	-.135	(.275)	-.299	(.263)
Edu3	-.529**	(.241)	-.460	(.324)	-.558	(.421)
Edu4	-.751**	(.333)	-.525	(.395)	-1.844*	(1.022)
TVC						
Health(t)	.338**	(.174)	.488**	(.215)	-.044	(.303)
Married(t)	.156	(.207)	-.130	(.317)	.221	(.281)
Pension(t)	1.339***	(.164)	1.562***	(.201)	.921***	(.325)
Town(t)	-.001	(.178)	-.037	(.229)	.109	(.293)
Rural(t)	-.253	(.189)	-.178	(.239)	-.453	(.316)
Constant	-5.359***	(.326)	-5.101***	(.462)	-4.527***	(.389)
No. of subjects	915		644		271	
No. of retirees	202		123		79	
Log likelihood	-559.698		-329.229		-222.896	
LR chi2 (15)	121.79***		87.39***		26.91**	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

Table 5.10.5 Weibull Model: Panel Data

Sample	Overall		Male		Female	
Duration	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
TCC						
Age2	.227	(.186)	-.006	(.252)	.567**	(.284)
Age3	.217	(.206)	-.124	(.273)	.572*	(.329)
Age4	.302	(.306)	.373	(.371)	.085	(.610)
Gender	.962***	(.169)	-		-	
Race2	-.184	(.194)	-.447	(.274)	.240	(.295)
Race3	.707**	(.287)	.468	(.345)	1.357**	(.610)
Race4	-.262	(.725)	-.687	(1.028)	-.178	(1.034)
Edu2	-.177	(.183)	-.144	(.276)	-.298	(.264)
Edu3	-.545**	(.241)	-.458	(.328)	-.562	(.421)
Edu4	-.767**	(.331)	-.478	(.395)	-1.859*	(1.023)
TVC						
Health(t)	.315*	(.175)	.424*	(.218)	-.047	(.303)
Married(t)	.148	(.208)	-.219	(.318)	.224	(.282)
Pension(t)	1.426***	(.166)	1.695***	(.203)	.930***	(.326)
Town(t)	.007	(.179)	-.002	(.231)	.108	(.293)
Rural(t)	-.299	(.190)	-.226	(.240)	-.461	(.317)
Constant	-6.472***	(.424)	-7.333***	(.631)	-4.616***	(.495)
/ln_α	.290***	(.060)	.526***	(.077)	.028	(.096)
α	1.336***	(.080)	1.693***	(.130)	1.029	(.099)
1/α	.748***	(.045)	.591***	(.045)	.972	(.093)
No. of subjects	915		644		271	
No. of retirees	202		123		79	
Log likelihood	-549.273		-309.990		-222.852	
LR chi2 (15)	132.60***		99.84***		26.98**	

Notes:

1. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

5.6.2.4 Frailty Model: Panel Data

This section considers a "frailty" component included in the model with the panel data between 1996 and 1999. The main estimated results are shown in Table 5.10.6a. First, without unobserved heterogeneity, the estimated coefficient of those with Race3 variable is positive and statistically significant and has higher hazard rates *ceteris paribus*. In contrast, the estimated coefficients of Edu2, Edu3, and Edu4 variables are significantly negative. In particular, for the TVCs, the estimated coefficients of Health(t) and Pension(t) variables are positive and statistically significant and have increasing hazard rates over time *ceteris paribus*. Furthermore, the estimate for the shape parameter is $\alpha = 1.297$ suggesting an increasing hazard over time.

Second, the frailty model with TVCs is assumed to follow a gamma distribution with mean 1 and variance equal to theta(θ). The estimate of theta is 0.199, but it is insignificant. A variance of zero (theta = 0) would indicate that the frailty component does not contribute to the model. A likelihood ratio test for the hypothesis theta = 0 is shown directly below the parameter estimates and indicates a chi-square value of 22.53 with 1 degree of freedom yielding a highly significant p-value of 0.000. Further, compared to the 1996 result in Table 4.6.1 that the estimate for theta fell from 0.262 to 0.199. This confirms that the model with TVCs can reduce the influence of unobserved heterogeneity.

Moreover, for the time-constant covariates, the estimated coefficients on the regressors Race3, Edu3, and Edu4 are a little bit larger in magnitude that the corresponding coefficients in the reference model. In particular, the estimated coefficient of Race3 has a significantly positive effect on employment duration, other things being equal. This implies that Mainlander workers have a higher hazard rate of retirement. In contrast, the coefficients of Edu3 and Edu4 variables have a significantly negative effect on employment duration. This means that workers with better education have a lower hazard rate of retirement. For the time-varying covariates, the estimated coefficient of Health(t) is a little bit smaller, but the Pension(t) variable is a little bit larger in magnitude than the corresponding coefficients in the reference model. These two variables all have a significantly positive effect on employment duration, other things being equal. This implies that workers with poorer health and people with a realised pension have a higher hazard rate of retirement. The Weibull distribution shape parameter $\alpha = 1.333$ is also a little bit larger in the frailty model than $\alpha = 1.297$ in the reference model. Therefore, the neglected heterogeneity may let the bias underestimates duration dependence.

Finally, the unobserved factors may contribute an extra layer of heterogeneity, leading to greater variability in duration of employment than might be expected under the model without the frailty component, and the effect of unobserved heterogeneity might gradually be reduced for the model with TVCs. For example, in Table 5.10.6b, we can see a significant frailty effect. The variance (theta) decreased from 0.206 in the frailty model without TVCs (i.e. poor health variable) to 0.199 in the model with TVCs, $\hat{\alpha}$ decreased from 1.337 to 1.333, and most estimated coefficients are also decreased from the frailty model without TVCs to with TVCs. Therefore, the effect of unobserved heterogeneity can be reduced for the model with TVCs.

Table 5.10.6a Frailty Model: Panel Data

	Without Unobserved Heterogeneity		With Gamma-Heterogeneity	
Variables	Coefficient	Standard Error	Coefficient	Standard Error
TCC				
Age2	.082	(.184)	.217	(.186)
Age3	.051	(.204)	.206	(.204)
Age4	.202	(.302)	.294	(.305)
Race2	-.146	(.193)	-.180	(.194)
Race3	.540*	(.283)	.696**	(.285)
Race4	-.183	(.726)	-.256	(.724)
Edu2	-.481***	(.172)	-.196	(.183)
Edu3	-.878***	(.230)	-.567**	(.241)
Edu4	-1.096***	(.323)	-.789**	(.331)
TVC				
Health(t)	.350**	(.173)	.317*	(.174)
Married(t)	-.165	(.199)	.127	(.208)
Pension(t)	1.321***	(.163)	1.418***	(.164)
Town(t)	-.033	(.178)	.003	(.179)
Rural(t)	-.349*	(.187)	-.302	(.189)
Constant	-5.368***	(.366)	-5.845***	(.492)
/ln_α	.260***	(.060)	.287***	(.058)
/ln_the			-1.610	(1.029)
α	1.297***	(.078)	1.333***	(.077)
1/α	.771***	(.046)	.749***	(.043)
theta			.199	(.205)
No. of subjects	915		915	
No. of retirees	202		202	
Log likelihood	-564.646		-553.384	
LR chi2 (14)	101.86***		106.32***	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. In particular, Likelihood-ratio test of theta=0: chibar 2 (01) = 22.53, Prob >= chibar2 = 0.000.

Table 5.10.6b Frailty Models with or without TVCs: Panel Data

	With TVCs		Without TVCs	
Variables	Coefficient	Standard Error	Coefficient	Standard Error
TCC				
Age2	.217	(.186)	.241	(.185)
Age3	.206	(.204)	.220	(.204)
Age4	.294	(.305)	.320	(.305)
Race2	-.180	(.194)	-.171	(.194)
Race3	.696**	(.285)	.695**	(.286)
Race4	-.256	(.724)	-.208	(.724)
Edu2	-.196	(.183)	-.176	(.183)
Edu3	-.567**	(.241)	-.563**	(.241)
Edu4	-.789**	(.331)	-.804**	(.331)
TVC				
Health(t)	.317*	(.174)	-	
Married(t)	.127	(.208)	.158	(.208)
Pension(t)	1.418***	(.164)	1.428***	(.164)
Town(t)	.003	(.179)	.031	(.178)
Rural(t)	-.302	(.189)	-.241	(.187)
Constant	-5.845***	(.492)	-5.873***	(.496)
/ln_α	.287***	(.058)	.291***	(.058)
/ln_the	-1.610	(1.029)	-1.581	(1.027)
α	1.333***	(.077)	1.337***	(.077)
1/α	.749***	(.043)	.748***	(.043)
theta	.199	(.205)	.206	(.211)
No. of subjects	915		915	
No. of retirees	202		202	
Log likelihood	-553.384		-554.961	
LR chi2 (14)	106.32***		103.17***	

Notes:
1. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.
2. The LR chi2 (13) is 103.17 in the frailty model without poor health variable.
3. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. In particular, Likelihood-ratio test of theta=0: chibar 2 (01) = 23.23, Prob >= chibar2 = 0.000.

5.6.2.5 Cox Hazard Model

The Cox hazard estimates for the retirement model are given in Tables 5.11.2 and 5.11.4. Specially, the results derive from the duration model using panel data with time-varying covariates (TVCs) affecting individual hazard rates of retirement. The Cox hazard regression (5.11) becomes

$$\ln h_i(t) = h_0(t) + \sum_{j=1}^{p} \beta_j x_{ij} + \sum_{r=p+1}^{k} \beta_r x_{ir}(t). \tag{5.11}$$

A positive coefficient means that this particular variable has positive effects on retirement, while a negative coefficient implies that a worker works longer, postponing retirement. At the same time, the hazard rates of retirement for the sub-sample by gender are also presented in Tables 5.11.2 and 5.11.4.

5.6.2.5.1 Cox Hazard Model: Cross-Sectional Data

Table 5.11.1 first shows test statistics and p-values for the Cox hazard model with cross-sectional data by the log-rank test and generalised Wilcoxon test in 1999. The variables of Gender, Race3 (Mainlander), Edu1 (informal education), Pension, Urban, and Rural are significant to estimate the employment survival function for retirement behaviour. This implies that the above variables may affect the retirement hazard. However, the other variables may also be important factors, but insignificant for influencing retirement behaviour.

Tables 5.11.2a and 5.11.2b show that for the Cox hazard model with the 1999 cross-sectional data. Most estimated coefficients are similar to their counterparts from the exponential and Weibull models. For instance, the estimated coefficients of

Gender, Race3 (Mainlander), and Pension variables have a significantly positive effect on retirement implying that female workers, Mainlander workers, and workers eligible for a pension have a higher hazard rate of retirement. In contrast, the estimated coefficients of Edu3 (7 to 12 years of schooling) and Edu4 (13 to 17 years of schooling) have a significant negative effect on retirement meaning that workers with a better education have a lower hazard rate of retirement. However, the Cox hazard model has some different effects from the exponential and Weibull models. For example, the Age4 (aged 68 to 73), Race2 (Hakka) and Race4 (Aboriginal) variables have a positive effect on retirement hazard for the exponential and Weibull models, and negative effect for the Cox hazard model, but all insignificantly. Furthermore, the coefficients of Age2 (aged 58 to 62) and Age3 (aged 63 to 67) variables have positive effects on retirement, but also insignificant. These results might relate to the Cox hazard model lacking a negative constant term.

Further, for gender effects, the estimated results of the Cox hazard model with the 1999 cross-sectional data for men and women are shown in Tables 5.11.2a and 5.11.2b. For men and women, the estimated coefficients of Race3 and Pension variables have a significantly positive effect on retirement implying that Mainlander workers and workers eligible for a pension have a higher hazard rate of retirement. However, due to their traditional culture, the estimated coefficient of the Race2 variable for men has a significantly negative effect and for women has a significantly positive effect on retirement. This implies that male Hakka workers are less likely to retire and female Hakka workers are more likely to retire.

Table 5.11.1 Test Statistics for the Cox Hazard Model: Cross-Sectional Data

Variables	Log-rank Test		Generalised Wilcoxon Test	
	Value	p-Value	Value	p-Value
Age1*	2.93	0.086	4.81	0.028
Age2	2.29	0.130	0.57	0.451
Age3	0.00	0.976	1.37	0.241
Age4	0.05	0.819	1.00	0.317
Gender***	32.25	0.000	39.12	0.000
Race1	0.84	0.360	2.43	0.118
Race2	1.80	0.179	0.40	0.525
Race3***	27.13	0.000	19.66	0.000
Race4	0.01	0.933	0.27	0.606
Edu1***	10.36	0.001	14.91	0.000
Edu2	2.68	0.101	1.50	0.220
Edu3	0.81	0.369	2.62	0.105
Edu4	0.24	0.621	0.79	0.375
Poor Health	2.26	0.133	2.56	0.109
Married	1.81	0.178	2.21	0.136
Pension***	89.11	0.000	55.24	0.000
Urban**	5.16	0.023	2.48	0.115
Town	0.13	0.718	1.45	0.228
Rural*	3.09	0.078	7.57	0.005

Notes:

1. According to the 1999 SHLS data, the effective sample of duration model has 966 observations, including 253 retirees and 713 continuing work.

2. Effects by Log-rank Test are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

Table 5.11.2a Cox Hazard Model: Cross-Sectional Data

Sample	Overall		Male		Female	
Duration	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Age2	.197	(.168)	.021	(.224)	.463*	(.260)
Age3	.039	(.192)	-.342	(.252)	.455	(.313)
Age4	-.181	(.284)	-.391	(.348)	.166	(.560)
Gender	.863***	(.155)	-		-	
Race2	-.061	(.168)	-.408*	(.229)	.515**	(.262)
Race3	.699***	(.258)	.633**	(.309)	.957*	(.505)
Race4	-.011	(.594)	-.036	(.741)	-.202	(1.031)
Edu2	-.229	(.165)	-.209	(.242)	-.438*	(.241)
Edu3	-.453**	(.221)	-.337	(.291)	-.519	(.408)
Edu4	-.606**	(.298)	-.323	(.355)	-1.886*	(1.023)
Health	.146	(.158)	.159	(.200)	-.024	(.276)
Married	.104	(.186)	-.299	(.278)	.275	(.255)
Pension	1.321***	(.153)	1.474***	(.188)	.914***	(.284)
Town	-.051	(.165)	.039	(.211)	-.052	(.275)
Rural	-.206	(.167)	-.044	(.206)	-.425	(.291)
No. of subjects	966		681		285	
No. of retirees	253		160		93	
Log likelihood	-1450.884		-845.157		-440.298	
LR chi2 (15)	129.16***		87.73***		32.32***	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

Table 5.11.2b Cox Hazard Model: Cross-Sectional Data

Sample	Overall		Male		Female	
Duration	Haz. Ratio	Std. Err.	Haz. Ratio	Std. Err.	Haz. Ratio	Std. Err.
Age2	1.218	(.205)	1.021	(.228)	1.590*	(.414)
Age3	1.040	(.199)	.710	(.179)	1.577	(.493)
Age4	.835	(.237)	.676	(.235)	1.181	(.661)
Gender	2.371***	(.367)	-		-	
Race2	.941	(.158)	.665*	(.152)	1.674**	(.439)
Race3	2.011***	(.519)	1.883**	(.582)	2.605*	(1.315)
Race4	.989	(.588)	.964	(.714)	.817	(.842)
Edu2	.796	(.131)	.812	(.197)	.645*	(.156)
Edu3	.636**	(.140)	.714	(.208)	.595	(.243)
Edu4	.545**	(.163)	.724	(.257)	.152*	(.155)
Health	1.157	(.183)	1.172	(.235)	.976	(.270)
Married	1.110	(.206)	.741	(.206)	1.317	(.336)
Pension	3.747***	(.572)	4.368***	(.820)	2.494***	(.709)
Town	.950	(.157)	1.040	(.220)	.950	(.262)
Rural	.814	(.136)	.957	(.198)	.654	(.191)
No. of subjects	966		681		285	
No. of retirees	253		160		93	
Log likelihood	-1450.884		-845.157		-440.298	
LR chi2(15)	129.16***		87.73***		32.32***	

Notes:

1. Effects are significant at * $p \le .10$, ** $p \le .05$, *** $p \le .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

5.6.2.5.2 Cox Hazard Model: Panel Data

Table 5.11.3 shows the log-rank test, generalised Wilcoxon test, and p-values for the Cox hazard model with panel data between 1996 and 1999. Comparing with the cross-sectional data analysis in Table 5.11.1, the tested results have a special change, including the variables of Health and Married. This implies that poor health and married variables may change to affect the retirement hazard during this period. Therefore, these two time-varying covariates may be more important for investigating the retirement behaviour.

Tables 5.11.4a and 5.11.4b show the Cox hazard model with panel data. First, the estimated coefficients of time constant covariates, including Gender, Race3, Edu3, and Edu4 variables have the same effects as the Cox model in Table 5.11.1a. Second, for the time-varying covariates, the estimated coefficients of $Health(t)$, $Married(t)$, $Pension(t)$, and $Town(t)$ variables have a positive effect on retirement, but only the coefficient of $Pension(t)$ variable is significant. This means that workers able to claim their pension have higher hazard rates of retirement other things being equal. Most occupational pensions in Taiwan are received as a lump sum payment. Compared to the 1996 result, workers eligible for a pension have a higher hazard rate of retirement. This is consistent with the results reported in Slade (1987), who found that workers with pensions had a higher hazard rate of retirement. Finally, the coefficient of $Rural(t)$ variable has a negative, but statistically insignificant, effect on retirement.

For the hazard ratio, the Gender variable is 2.861. This means that, other variables being constant, the estimated hazard ratio of female workers compared to that of male workers is 2.861 times greater. Other results can be described as follows:

the estimated hazard rate of Mainlander workers compared to that of Fujianese workers is 2.406 times higher. The estimated hazard ratio of workers with high education is only 0.511 times that of workers with informal education. The estimated hazard ratio of workers eligible for a pension is 4.777 times greater than for those ineligible.

Tables 5.11.4a and 5.11.4b also report the results for males and females. Some effects are different. For example, the estimated coefficient of the $Health(t)$ variable for males is positive on retirement and negative for females. That implies male workers in poor health increase the hazard rate of retirement, and female workers in poor health decrease the hazard rate of retirement, other things being equal, but insignificantly. Further, the estimated coefficient of the $Married(t)$ variable for males has a negative effect and for females has positive effect on retirement. This means that male married workers have a lower hazard rate of retirement and female married workers have a higher hazard rate of retirement, but this is statistically insignificant.

Table 5.11.3 Test Statistics for the Cox Hazard Model: Panel Data

Variables	Log-rank Test		Generalised Wilcoxon Test	
	Value	p-Value	Value	p-Value
Age1	2.36	0.124	4.70	0.030
Age2	1.00	0.316	0.05	0.816
Age3	0.01	0.941	2.91	0.088
Age4	0.53	0.467	1.37	0.241
Gender***	31.32	0.000	35.63	0.000
Race1	0.42	0.519	1.21	0.272
Race2*	3.10	0.078	1.29	0.256
Race3***	30.69	0.000	21.85	0.000
Race4	0.06	0.801	1.36	0.243
Edu1***	7.39	0.006	8.32	0.003
Edu2	1.86	0.172	0.46	0.496
Edu3	0.57	0.450	2.38	0.122
Edu4	0.19	0.660	0.31	0.576
Poor Health*	3.21	0.073	1.97	0.160
Married*	3.03	0.082	3.90	0.048
Pension***	103.03	0.000	60.75	0.000
Urban**	4.22	0.040	2.02	0.155
Town	0.03	0.868	1.45	0.227
Rural**	4.12	0.042	6.25	0.012

Notes:

1. As the model with panel data between 1996 and 1999, the effective sample becomes to 915 observations, including 202 retirees, and 713 continuing work.

2. Effects by log-rank test are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

Table 5.11.4a Cox Hazard Model: Panel Data

Sample	Overall		Male		Female	
Duration	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
TCC						
Age2	.193	(.187)	-.041	(.253)	.522*	(.285)
Age3	-.062	(.217)	-.472	(.291)	.371	(.343)
Age4	-.149	(.318)	-.340	(.403)	.089	(.589)
Gender	1.051***	(.173)	-		-	
Race2	-.226	(.196)	-.521*	(.275)	.257	(.299)
Race3	.878***	(.288)	.694**	(.343)	1.393**	(.589)
Race4	-.312	(.726)	-.714	(1.032)	-.186	(1.035)
Edu2	-.170	(.185)	-.178	(.279)	-.271	(.265)
Edu3	-.442*	(.243)	-.342	(.330)	-.491	(.421)
Edu4	-.671**	(.333)	-.416	(.402)	-1.806*	(1.025)
TVC						
Health(t)	.257	(.176)	.301	(.223)	-.026	(.305)
Married(t)	.181	(.209)	-.299	(.320)	.322	(.285)
Pension(t)	1.564***	(.169)	1.814***	(.207)	1.051***	(.328)
Town(t)	.012	(.180)	.099	(.235)	.044	(.292)
Rural(t)	-.272	(.191)	-.145	(.242)	-.450	(.317)
No. of subjects	915		644		271	
No. of retirees	202		123		79	
Log likelihood	-1134.687		-627.901		-369.608	
LR chi2(15)	140.52***		105.73***		25.70**	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis that all coefficients except the intercept are 0 at the 0.01 level. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

Table 5.11.4b Cox Hazard Model: Panel Data

Sample	Overall		Male		Female	
Duration	Haz. Ratio	Std. Err.	Haz. Ratio	Std. Err.	Haz. Ratio	Std. Err.
TCC						
Age2	1.212	(.227)	.960	(.242)	1.685*	(.480)
Age3	.940	(.204)	.624	(.182)	1.449	(.497)
Age4	.861	(.274)	.712	(.287)	1.092	(.644)
Gender	2.861***	(.495)	-		-	
Race2	.798	(.156)	.594*	(.163)	1.293	(.386)
Race3	2.406***	(.692)	2.002**	(.687)	4.028**	(2.373)
Race4	.732	(.532)	.490	(.505)	.830	(.860)
Edu2	.844	(.156)	.837	(.234)	.763	(.202)
Edu3	.643*	(.156)	.710	(.234)	.612	(.258)
Edu4	.511**	(.170)	.659	(.265)	.164*	(.168)
TVC						
Health(t)	1.293	(.228)	1.352	(.301)	.974	(.297)
Married(t)	1.198	(.250)	.741	(.237)	1.379	(.393)
Pension(t)	4.777***	(.806)	6.136***	(1.269)	2.861***	(.939)
Town(t)	1.012	(.182)	1.104	(.259)	1.045	(.306)
Rural(t)	.762	(.146)	.865	(.209)	.638	(.202)
No. of subjects	915		644		271	
No. of retirees	202		123		79	
Log likelihood	-1134.687		-627.901		-369.608	
LR chi2(15)	140.52***		105.73***		25.70**	

Notes:

1. Effects are significant at * $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$.

2. Goodness of fit: the result of Log-likelihood ratio test can reject the hypothesis. Considering the Gender variable, the LR chi2 of male and female samples is LR chi2 (14), respectively.

5.7 Conclusions

This chapter aims at contributing to the understanding the influences of labour force transition by the different specification models, including probit models to investigate the probabilities of exit from and re-entry into employment between 1996 and 1999, and duration models without or with time-varying covariates (TVCs) to estimate the hazard of retirement behaviour. In particular, this chapter has paid special attention to examining the frailty models without or with TVCs of labour force transition.

The first main results by probit models estimate the probability of labour force transition behaviour, including the probabilities of exit from and re-entry into employment between 1996 and 1999. Most results can confirm that older workers, female workers, workers eligible for a pension, workers with poor health in 1996, and those whose health becomes poorer have higher exit rates from employment. In contrast, workers with better education have lower exit rates from employment. Furthermore, older workers and female workers have lower re-entry rates into employment. Interestingly, workers with formal education also have lower re-entry rates into employment. Perhaps, they have reached retirement age and do not wish to re-enter the labour market.

The second main results by duration models are used to estimate the hazard rate of retirement. In the duration models without time-varying covariates, most results are consistent with the results reported in Chapter 4. That is, older workers, female workers, Mainlander workers, and workers with eligible pension have higher hazard rates of retirement, and workers with better education have a lower hazard rate of retirement. Further, bringing unobserved heterogeneity into the frailty model, most estimated coefficients on the regressors are larger in magnitude than the corresponding coefficients in the reference model, but only Race3 and Pension variables have a significant positive effect, and Edu3 and Edu4 variables have a significant negative effect on retirement hazard.

Moreover, for the duration models with TVCs, the empirical results indicate that workers with poor health have a higher hazard rate of retirement, in particular as workers being in poor health increase the hazard rates of retirement, other things being equal. For examining the effect of unobserved heterogeneity on retirement behaviour, most estimated coefficients of the frailty models are larger in magnitude that the corresponding coefficients in the reference model. Comparing to the 1996 results in Table 4.6.1, the estimated theta has reduced from 0.262 to 0.199 by the model with TVCs in Table 5.10.6. This implies that unobserved heterogeneity might be expected to be less serious once TVCs are included in the hazard model.

Lastly, because the SHLS survey has limited information about income and wealth, this chapter does not discuss changes in income and wealth which assuredly affect retirement behaviour. A possible later analysis could include more time-varying factors that deeply affect retirement behaviour, such as changes in financial health.

5.8 Appendix

The brief STATA commands for analysing the labour force transition are given as follows:

Table 5.2

```
use "C:\Documents and Settings\User\My Documents\Revised 2007 Summer\SHLS Data 2007\Chapter 4 Data Set 082007.dta"
sum heal1- heal5 marit1- marit5 pension resid1-resid3 if duration~=. & resid~=. & race~=. & eyhat~= .
use "C:\Documents and Settings\User\My Documents\Revised 2007 Summer\SHLS Data 2007\1999 (3) Data Set 082007.dta"
sum heal91- heal95 marit91- marit95 pen9 resid91- resid93 if cen~= . & resid~=.
```

Table 5.5

```
use "C:\Documents and Settings\User\My Documents\Revised 2007 Summer\SHLS Data 2007\1999 (3) Data Set 082007.dta"
sum exit age91- age94 gender race1- race4 edu1- edu4 pen9 poorh6 married6 resid62 resid63 healthcp maritc resid2cc resid3cc if history6==8 & resid6~=.
```

Table 5.6

```
sum reentry age91-age94 gender race1-race4 edu1-edu4 pen9 poorh6 married6 resid62 resid63 healthcp maritc resid2cc resid3cc if history6==1 & resid6~=.
```

Table 5.7.1 ***(without TVC)***

```
stset dur, failure( cen)
sum dur cen age91- age94 gender race1- race4 edu1- edu4 poorh9 married9 pen9 resid91-resid93 if dur~=. & resid~=.
```

Table 5.7.2 ***(with TVC)***

```
expand 2 if cen==1
sort id
by id: gen count=_n
gen const=1
by id: egen total = sum(const)
drop const
gen died=1 if cen==1
replace died=0 if count==1 & total==2
replace poorh9= poorh6 if count==1 & total==2
```

replace married9= married6 if count==1 & total==2
replace pen9= pen6 if count==1 & total==2
replace resid92= resid62 if count==1 & total==2
replace resid93= resid63 if count==1 & total==2
replace dur= dur619 if count==1 & total==2
stset dur, id(id) failure(died)
sum dur cen age91- age94 gender race1- race4 edu1- edu4 poorh9 married9 pen9 resid91-resid93 if _t~=. & resid~=.

Table 5.8.1

probit exit age92- age94 gender race2- race4 edu2- edu4 pen9 married6 poorh6 resid62 resid63 maritc healthcp resid2cc resid3cc if history6==8
probit exit age92- age94 race2- race4 edu2- edu4 pen9 married6 poorh6 resid62 resid63 maritc healthcp resid2cc resid3cc if history6==8 & gender==0
probit exit age92- age94 race2- race4 edu2- edu4 pen9 married6 poorh6 resid62 resid63 maritc healthcp resid2cc resid3cc if history6==8 & gender==1

Table 5.8.2

probit exit age92- age94 gender race2- race4 edu2- edu4 pen9 married6 poorh6 resid62 resid63 maritc healthcp resid2cc resid3cc if history6==8
mfx compute, nodiscrete
probit exit age92- age94 race2- race4 edu2- edu4 pen9 married6 poorh6 resid62 resid63 maritc healthcp resid2cc resid3cc if history6==8 & gender==0
mfx compute, nodiscrete
probit exit age92- age94 race2- race4 edu2- edu4 pen9 married6 poorh6 resid62 resid63 maritc healthcp resid2cc resid3cc if history6==8 & gender==1
mfx compute, nodiscrete

Table 5.9.1

probit reentry age92- age94 gender race2- race4 edu2- edu4 pen9 poorh6 married6 resid62 resid63 healthcp maritc resid2cc resid3cc if history6==1
probit reentry age92- age94 race2- race4 edu2- edu4 pen9 poorh6 married6 resid62 resid63 healthcp maritc resid2cc resid3cc if history6==1 & gender==0
probit reentry age92- age94 race2- race4 edu2- edu4 pen9 poorh6 married6 resid62 resid63 healthcp maritc resid2cc resid3cc if history6==1 & gender==1

Table 5.9.2

probit reentry age92- age94 gender race2- race4 edu2- edu4 pen9 poorh6 married6

resid62 resid63 healthcp maritc resid2cc resid3cc if history6==1
mfx compute, nodiscrete
probit reentry age92- age94 race2- race4 edu2- edu4 pen9 poorh6 married6 resid62 resid63 healthcp maritc resid2cc resid3cc if history6==1 & gender==0
mfx compute, nodiscrete
probit reentry age92- age94 race2- race4 edu2- edu4 pen9 poorh6 married6 resid62 resid63 healthcp maritc resid2cc resid3cc if history6==1 & gender==1
mfx compute, nodiscrete

Duration Models without TVC analysis
use "C:\Documents and Settings\User\My Documents\Revised 2007 Summer\SHLS Data 2007\1999 (3) Data Set 082007.dta"
stset dur, failure(cen)

Table 5.10.1
streg age92- age94 gender race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, distribution (exponential) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==0, distribution (exponential) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==1, distribution (exponential) nohr

Table 5.10.2
streg age92- age94 gender race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, distribution (weibull) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==0, distribution (weibull) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==1, distribution (weibull) nohr

Table 5.10.3
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, distribution (weibull) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, distribution (weibull) nohr frailty (gamma) shared (gender)

Table 5.11.1a
stcox age92- age94 gender race2- race4 edu2- edu4 poorh9 married9 pen9 resid92

resid93, nohr
stcox age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==0, nohr
stcox age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==1, nohr

Table 5.11.1b
stcox age92- age94 gender race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93
stcox age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==0
stcox age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==1

Duration Models with TVC analysis
stset dur, id(id) failure(died)

Table 5.10.4
streg age92- age94 gender race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, distribution (exponential) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==0, distribution (exponential) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==1, distribution (exponential) nohr

Table 5.10.5
streg age92- age94 gender race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, distribution (weibull) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==0, distribution (weibull) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==1, distribution (weibull) nohr

Table 5.10.6
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, distribution (weibull) nohr
streg age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, distribution (weibull) nohr frailty (gamma) shared (gender)

Table 5.11.2a

stcox age92- age94 gender race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93, nohr

stcox age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==0, nohr

stcox age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==1, nohr

Table 5.11.2b

stcox age92- age94 gender race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93

stcox age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==0

stcox age92- age94 race2- race4 edu2- edu4 poorh9 married9 pen9 resid92 resid93 if gender==1

Chapter 6

Conclusions and Recommendations

6.1 Summary of the Findings

This study has investigated the determinants of retirement behaviour, using data from the second panel of the SHLS survey of the middle-aged and elderly in Taiwan. In doing so, we examined three broad fields of labour economics: labour force participation, retirement, and labour force transition. Although the SHLS survey has been used for analysing the living arrangements of the elderly (Chen, 1994; Chang, 1999), the economic well being of the elderly (Hermalin et al, 1999), and health status and health-care utilisation of the elderly (Chen, 1999), this is the first time SHLS data has been used for the economic analysis of retirement issues. In particular, the empirical work focuses on duration models in analysing retirement behaviour. First, how people make their decisions to enter the labour market or work was discussed in Chapter 3. Second, how people make their decisions to exit the labour market or retire in Chapter 4. Third, how people make their transition decisions to exit from or re-enter into the labour market in Chapter 5.

In general, the empirical results in this study confirm theoretical expectations. First, for the Age variables, older workers have a lower probability of labour force participation as shown in Chapter 3 and have a higher hazard rate of retirement as discussed in Chapter 4. In particular, workers retire at around ages 55, 60 and 65. This is consistent with the results reported in Shih (1999). The hazard rate of retirement increases rapidly after age 50. This is because of the unique "duration" retirement concept. As the workers are relatively young when they first retire, many may find a second job before they retire again. However, as workers age, their health declines and they retire. This is the key result in the time-varying covariate analysis in Chapter 5.

Second, with respect to the Gender variable, female workers tend to retire earlier. This is evident from the nonparametric estimation in Chapter 2. In particular, female workers have a lower probability of labour force participation according to the probit analysis in Chapter 3, have a higher hazard rate of retirement from the duration model in Chapter 4, and also have a higher hazard rate of retirement from the duration models in Chapter 5. These confirm the results reported in Blau and Riphahn (1999), who found that wives have a lower probability of labour force participation than husbands between aged 50 and 70. Single women, however, are more likely to continue work for longer.

Third, for the Race variables, Mainlander workers have a lower probability of employment as discussed in Chapter 3 and have a higher hazard rate of retirement as shown in Chapter 4. In contrast, Hakka workers have a higher probability of labour force participation as seen in Chapter 3 and have a lower hazard rate of retirement as discussed in Chapter 4. Comparing the 1999 SHLS data, these results have not changed as demonstrated in Chapter 5. These are consistent with the results reported in Shih (1999), who found that Mainlander workers had a higher hazard rate of retirement.

Fourth, with respect to the effects of Education, workers with better educational attainments have a higher probability of employment as demonstrated in Chapter 3 and have a lower hazard rate of retirement as seen in Chapter 4. These results are similar to those reported in Chang (1999), who found that the education status of the elderly is an important factor in determining their economic independence and health status in Taiwan. Comparing the 1999 SHLS data, this factor has remained the same, so elderly workers have a lower hazard rate of retirement, as shown in Chapter 5, but the coefficients are insignificant.

Fifth, for the Health variables, workers with poor health have a lower probability of participation in work as demonstrated in Chapter 3 and have a higher hazard rate of retirement as discussed in Chapter 4, consistent with the results reported in Diamond and Hausman (1984), who found that the demographic variable with by far the largest effect is bad health and workers in poor health have a higher hazard rate of retirement. Comparing the 1999 SHLS data, workers being in poor health increased their hazard rate of retirement, other things being equal, as discussed in Chapter 5.

Sixth, with respect to the effects of Marital Status, married workers have a higher probability of employment as shown in Chapter 3 and have a lower hazard rate of retirement as discussed in Chapter 4. In particular, married male workers have a higher probability of participation in work than unmarried males, but married female workers have a lower probability of employment than unmarried females. These results are consistent with those reported in Chan and Stevens (2001), who noted that married female workers have a lower probability of entry-to-work and a higher probability of exit-from-work. Comparing to the 1999 SHLS data, changes in marital status have a higher but insignificant hazard rate of retirement as seen in Chapter 5.

Seventh, this study indicates that the Pension variable has a conditional effect on retirement behaviour. That is, a pre-condition for being eligible for a pension can provide a strong incentive for people to participate in work before working 35 years as shown in Chapter 3. Workers eligible for a pension also have a lower hazard rate of retirement after working 35 years as described in Chapter 4. However, male workers eligible for a pension generally have a low participation rate and a higher hazard rate of retirement after working 17 years. Furthermore, if workers expect to have a higher pension income then they have a higher hazard rate of retirement. These results are somewhat different from those reported in Diamond and Hausman (1984), who noted that both pension and social security have an expected strong positive effect on retirement behaviour. Comparing the 1999 SHLS data, pension reliability shows a higher and significant effect on the retirement.

Eighth, with respect to the Residence effect, town and rural workers have a higher probability of employment, as discussed in Chapter 3, and have a lower hazard rate of retirement as shown in Chapter 4. These results differ from those reported in Gunderson (1977), who argued that rural workers have a lower participation rate, assuming urban residence could provide more employment opportunities for workers than rural areas. However, when town and rural workers change their residence status, they show a lower, but also insignificantly different hazard rate of retirement, as seen in Chapter 5.

Finally, examining the effect of unobserved heterogeneity on retirement behaviour, most estimated coefficients of the frailty models are larger in magnitude than the corresponding coefficients in the reference model. In particular, the model with time-varying covariates might be expected to reduce the effect of unobserved heterogeneity on the retirement hazard.

6.2 Policy Implications

This study documents the work-to-retirement processes of the middle-aged and elderly in Taiwan. Despite low levels of pension for retirees, the determinants of retirement decisions are similar to those found in Western industrial countries. For example, older workers, female workers, and workers with poor health significantly have higher hazard rates of retirement. In contrast, workers with better educational attainments and workers living in rural areas have significantly lower hazard rates of retirement.

Pension is one of the key factors in the labour force analysis (Lazear, 1986). The empirical results discussed in Chapter 3 showed that workers eligible for a pension have a higher incentive to continue working, but the duration analysis in Chapter 4 showed that workers with higher predicted pension incomes are more likely to retire. These are consistent with the preliminary results reported in Chapter 2 whereby, before 35 years employment duration, workers eligible for a pension have a higher incentive to participate in work; and after 35 years employment duration, workers eligible for a pension have a higher incentive to retirement. This seems to be contradictory for individual decisions. Actually, this presents a good case for re-thinking Taiwan's pension policies: for example, the current Labour Standards Law (LSL) in Taiwan, state that no employee is eligible to receive a retirement pension until the employee has worked 25 years for the same employer or is over age 55. Hence, this study suggests that the current occupational pension system may extend the employment duration to 35 years or until the worker is over age 65. On the one hand, workers can pay longer contributions (for 35 years) and then they would receive more benefits for their later life than the current pension system (for 25 years). On the other hand, the current pension system easily induces those workers eligible for a pension to retire early and receive their pension benefits. In particular, for interest payments on the pension benefits, most workers eligible for a pension are governmental employees. When they retire and receive their lump sum pension benefits they can deposit part of this amount with financial institutions, which are legally required to pay the market interest rate.[1] Further, if the government employees retired after working only 25 years, or aged 55, they might easily return to work again. Younger potential workers would be crowded out lessening their opportunity and increasing their difficulty in finding a job or entering into the labour market. This is really not good for the development of the labour force.

The empirical results presented in Chapter 4 demonstrate that workers with higher predicted earnings have a lower hazard rate of retirement, and workers with higher predicted pension incomes have a higher hazard rate of retirement. In particular, under the new portable pension policy operating from 2005, workers have an individual retirement account, employers and employees will pay defined contributions during the employee's working life, and they will have higher pension incomes than before. Hence, the portable pension system is *likely* to lead to an increase in expected pension incomes and workers might have a higher hazard rate of retirement in the future.

Chapter 5, considering time-varying covariates, highlights that workers with poor health have a higher hazard rate of retirement, in particular as workers being in poor health increases the hazard rate of retirement other things being equal. Hence, the result suggests that the Taiwanese government should invest more in the health care industry and National Health Insurance. The first is for improving people's health,

[1] It is estimated that the interest rate is 18 percentage points.

and the second to help workers extend their time in the labour market.

6.3 Limitations of the Study

Retirement decisions are dependent on many factors, including individual decisions from labour supply aspects, employers' decisions from labour demand aspects, and government pension policy from the social security system. However, the main limitation of this study is the survey data. The SHLS questionnaires mainly focused on individual employment histories and collected limited data on employer responses and government social welfare programmes. Further, the response rates on wages, income, and assets were low, reflecting the reluctance of participants to divulge their true income. More detailed micro-data for retirement behaviour similar to the Retirement History Survey and the Health and Retirement Study in the US and the Retirement Survey in the UK simply do not exist for Taiwan. One more hidden danger of the SHLS data is the unknown accuracy of the responses given by the participants. No obvious means of verifying these responses exists.

The second problem is the lack of information about Taiwan's social security system or programmes from the SHLS data sets. There is a lack of analysis of pension policies. Moreover, until now the definitions of defined benefit and defined contribution have not been properly stated. Furthermore, the important issue of the types of state pensions, occupational pensions or personal pensions for the elderly are not fully covered in the survey. Lastly, the patterns of pension benefits paid and whether pensions should be taxed are matters that have not been settled. Issues relating to the social security system in Taiwan are currently under debate, and suitable occupational pension systems are expected to be developed in the near future.

The third problem is the limited information about the employers' labour demand effects from the SHLS data sets. For example, the productivity of middle-aged and elderly workers declines as they grow older, unless they acquire new skills to improve their abilities. In particular, they might take more time for caring for their family or for finding a second job to earn more money; their loyalty would be relatively lower than before. On the other hand, their wages and pension benefits will increase along with their years of employment. Both situations are detrimental to employers' businesses. The former will decrease productivity and revenues, and the latter will increase costs and expenditures. Therefore, senior employees are more likely to lose their job than their younger colleagues. None of this is reflected in the SHLS data sets. This issue is important in Taiwan because most private employers are small and medium sized enterprises.

6.4 Prospects for Future Research

The study has made important inroads into examining the determinants of labour force participation, retirement, and labour force transition of the middle-aged and elderly in Taiwan, but there are many issues left for future research. First, it would be useful to analyse joint retirement decisions by spouses. This is because most families depend on double salaries for their modern lives. The partners usually have similar ages and retirement situations. An et al. (1999) revealed a strong association between the retirement probabilities of spouses. They found the effects of wages were significant and asymmetric by gender. If we can collect more information about this, joint retirement decisions in Taiwan can be studied.

Second, it would be interesting to analyse retirement phases in Taiwan. Many workers retire after 25 years of working, but often wish to re-enter the labour market either full-time or part-time. This situation provides a good opportunity for examining determinants of retirement. The government has recently attempted to change the retirement rules for compulsory occupational pension policies from 75 to 85,[2] to take into account the longer life span of Taiwanese in recent years, as documented in chapters 1 and 2. For example, comparing the sample of SHLS data aged 50 and older in 1996, those people are same generation as shown in Table 1.2 in 2011. That is, the life expectancy of male is 76.2 years and female 82.8 years. If workers only work and pay 25 years of pension contributions but can receive more than 25 years of pension benefits, a pension fund deficit is almost certainly going to emerge. Therefore, if people can work longer, pay more pension contributions, or retire later that may solve the problem of finance crisis in the future.

Third, it would be useful to analyse financial management for retirees' later life in Taiwan, as most research focuses on living arrangements for the elderly (Chen, 1994; Chang, 1999) and health-care utilisation (Chen, 1999). Further research should focus on the financial management of elderly retirement plans. The financial management of bequests can be investigated.

6.5 Concluding Remarks

Mencius, "If you believe everything in a book, throw the book away."

Different societies confront different situations and have enacted different polices. Taiwan could learn from the social security systems of advanced countries and provide a better social welfare system for its citizens. These advanced countries have their own problems, including population ageing and financial crises, and they

[2] That is, the retirement condition of workers aged 50 and having worked for 25 years (rule 75) will change to workers aged 60 who must have worked for 25 years (rule 85).

have started considering how to solve these problems by possibly delaying retirement age, and changing pension systems from defined benefits to defined contributions. The Taiwanese government should note whether these countries tackle the above mentioned problems successfully, as this information will be helpful.

This thesis has produced some valuable results on retirement behaviour. The implications are useful for policymakers, especially since the government needs to construct a social security system for the elderly. In particular, the social security system not only provides an occupational pension for employees, but also takes into account public pensions and personal pensions for individuals. Further, the social security system needs to manage pension funds for defined contributions or taxing of defined benefits, and decide how to pay the benefits. Whether a pension benefit is paid one-off, per month or per year, has yet to be ironed out. Moreover, portable pensions might be suitable for workers because of frequent changes in jobs. People should be able to open individual accounts, pay contributions from their earnings, and get pension benefits after their retirement.

Finally, another specific recommendation is for policymakers. This concerns the survey data. The government could construct detailed panel surveys for employment and retirement. This is very important for investigating human resource management, including the decisions of labour force participation, retirement, and labour force transition.

Bibliography

1. Aalen, O. O. (1978). "Nonparametric Inference for a Family of Counting Processes." *Annals of Statistics*, 6, 701-726.
2. Aitchison, J. and Silvey, S. D. (1957). "The Generalization of Probit Analysis to the Case of Multiple Responses." *Biometrika*, 44(1), 131-140.
3. An, M. Y., Christensen, B. J., and Gupta, N. D. (1999). "A Bivariate Duration Model of the Joint Retirement Decisions of Married Couples." Centre for Labour Market and Social Research, *Working Paper*, 99(10), Denmark.
4. Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (1993). *Statistical models based on counting processes.* Springer Verlag, New York.
5. Antolin, P. and Scarpetta, S. (1998). "Micro-econometric Analysis of the Retirement Decision: Germany." *Maintaining Prosperity in an Ageing Society*, OECD, Working Paper AWP 1.6.
6. Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*, New York: Springer-Verlag.
7. Arulampalam, W., and Stewart, M. B. (1995). "The Determinants of Individual Unemployment Durations in an era of High Unemployment." *The Economic Journal*, 105, 321-332.
8. Baltagi, B. H. (2001). *Econometric Analysis of Panel Data*, Second edition, John Wiley & Sons, Ltd.
9. Barmby, T. A., Orme, C. D., and Treble, J. G. (1991). "Worker Absenteeism: An Analysis Using Microdata." *The Economic Journal*, 101(405), 214-229.
10. Berkovec, J., and Stern, S. (1991). "Job Exit Behaviour of Older Men." *Econometrica*, 59(1), 189-210.

11. Blau, D. M. (1994). "Labour Force Dynamics of Older Men." *Econometrica*, 62(1), 117-156.

12. Blau, D. M. (1998). "Labour Force Dynamics of Older Married Couples." *Journal of Labor Economics*, 16(3), 595-629.

13. Blau, D. M. and Riphahn, R. T. (1999). "Labour Force Transitions of Older Married Couples in Germany." *Labour Economics*, 6, 229-251.

14. Blundell, R. and Macurdy, T. (1999). "Labour Supply: A Review of Alternative Approaches." in *Handbook of Labour Economics*, 3A, 1559-1695.

15. Boheim, R., and Taylor, M. P. (2000). "Options or Obligation: The Determinants of Labour Supply Preferences in Britain." *Institute for Social and Economic Research, University of Essex.*

16. Borjas, G. J. (2008). *Labor Economics*, McGraw-Hill Irwin. New York.

17. Boskin, M. J. (1977). "Social Security and Retirement Decisions." *Economic Inquiry*, 15, 1-25.

18. Boskin, M. J., and Hurd, M. D. (1978). "The Effect of Social Security on Early Retirement." *Journal of Public Economics*, 10, 361-377.

19. Bound, J., Schoenbaum, M., Stinebrickner, T. R. and Waidmann, T. (1999). "The Dynamic Effects of Health on the Labour Force Transitions of Older Workers." *Labour Economics*, 6, 179-202.

20. Bridges, S. (2001). "*Organisational Performance and Human Resource Management.*" PhD thesis, Department of Economics, University of Newcastle upon Tyne.

21. Buckley, N. J., Denton, F. T., Robb, A. L., and Spencer, B. G. (2004). "The transition from good to poor health: an econometric study of the older population." *Journal of Health Economics*, 23, 1013-1034.

22. Burdett, K., Kiefer, N. M., Motensen, D. T., and Neumann, G. R. (1984). "Earnings, Unemployment and the Allocation of Time Over Time." *Review of Economic Studies*, 51, 559-578.

23. Burkhauser, R. (1980). "The Early Acceptance of Social Security: An Asset Maximization Approach." *Industrial and Labour Relations Review*, 33, 484-492.

24. Burtless, G. (1986). "Social Security, Unanticipated Benefit Increases, and the Timing of Retirement." *The Review of Economic Studies*, 53(5), 781-805.

25. Burtless, G. (1999). "An Economic View of Retirement." In Aaron, H. J., editor, *Behavioural Dimensions of Retirement Economics*, Brookings Institution Press, 7-42.

26. Cahuc, P. and Zylberberg, A. (2004). *Labour Economics*, Massachusetts Institute of Technology.

27. CEPD. (2000). *Taiwan Statistical Data Book.* Council for Economic Planning and Development, Executive Yuan, Republic of China.

28. CEPD. (2002). "Labour Force Participation Rates by Age and Sex." *Manpower Indicators*, Council for Economic Planning and Development, Executive Yuan, Republic of China.

29. Chan, S., and Stevens, A. H. (2001). "Job Loss and Employment Patterns of Older Workers." *Journal of Labour Economics*, 19(2), 484-521.

30. Chang, C. (1980). "The Impact of Marriage, Fertility, and Number of Children on Married Women's Labour Supply." Economic Essays, (in Chinese).

31. Chang, M. (1996). "Structural Changes of the Elderly and Its Effects on Economic Independency in Taiwan Area." *Book Series of Sun Yat-Sen Institute for Social Science and Philosophy*, Academia Sinica, 37, 1-30.

32. Chang, M. (1999). "A Longitudinal Study on Living Arrangements of the Elderly in Taiwan." In Chen C., Hermalin, A. I., Hu, S., and Smith, J. P., editors, *Emerging Social Economic Welfare Programs for Aging in Taiwan in a World Content*, The Institute of Economics, Academia Sinica, Taipei, Taiwan, 43-64.

33. Chen, C. (1994). "The Determinants of Satisfaction with Living Arrangements for the Elderly in Taiwan." *Journal of Population Studies*, 16, 29-52 (in Chinese).

34. Chen, C. (1999). "Health Status and Health-care Utilisation of the Elderly in Taiwan." In Chen C., Hermalin, A. I., Hu, S., and Smith, J. P., editors, *Emerging Social Economic Welfare Programs for Aging in Taiwan in a World Content*, The Institute of Economics, Academia Sinica, Taipei, Taiwan, 279-303.

35. Chiu, S.B. (2004). "Taiwan: Compulsory Occupational Pensions Still Dominate." International Conference on the Pensions in Asia: Incentives, Compliance and Their Role in Retirement. *Hitotsubashi Collaboration Centre*, Tokyo, Japan.

36. Chou, Y.J., and Staiger, D. (2001). "Health Insurance and Female Labour Supply in Taiwan." *Journal of Health Economics*, 20, 187-211.

37. Chuang, Y. (2006). "Introduction for the Survey of Health and Living Status of the Middle Aged and Elderly in Taiwan." at the Workshop for Panel Data Analysis, Central for Survey Research, Academia Sinica, Taipei, Taiwan, 12 (in Chinese).

38. Clark, R., Johnson, T., and McDermed, A.A. (1980). "Allocation of Time and Resources by Married Couples Approaching Retirement." *Social Security Bulletin*, 43, 3-16.

39. Cleves, M.A., Gould, W.W., and Gutierrez, R.G. (2004). *An Introduction to Survival Analysis Using Stata*, Revised Edition, Stata Corporation, USA.

40. Cochrane, J. H. (2001). *Asset Pricing*. Princeton University Press.

41. Coile, C. and Gruber, J. (2000). "Social Security and Retirement." National Bureau of Economics Research, Working Paper, 7830.

42. Collett, D. (1994). *Modelling Survival Data in Medical Research.* Chapman and Hall.

43. Cox, D. R. (1972). "Regression Models and Life-Tables (with discussion)." *Journal of Royal Statistical Society*, Series: B34, 187-220.

44. Cox, D. R., and Oakes, D. (1984) *Analysis of Survival Data*, Chapman and Hall, London.

45. Devine, T., and Kiefer, N. (1991). *Empirical Labour Economics: The Search Approach*, Oxford University Press.

46. Diamond, P., and Hausman, G. A. (1984). "Individual Retirement and Saving Behaviour." *Journal of Public Economics*, 23, 81-114.

47. Disney, R., Emmerson, C., and Wakefield, M. (2003). "Ill-health and Retirement in Britain: A Panel Data-Based Analysis." *The Institute for Fiscal Studies*, WP03/02.

48. Duffie, D. and Singleton, K. (1993). "Simulated Moments Estimation of Markov Models of Asset Prices." *Econometrica*, 61(4), 929-952.

49. Duncan, A. (2000). "Cross-Section and Panel Data Econometrics." *Lectures and Seminar Programme Notes*, University of Nottingham.

50. Dwyer, D.S., and Mitchell, O.S. (1998). "Health Problems as Determinants of Retirement: Are Self-Rated Measures Endogenous?" *Population Ageing Research Centre*, University of Pennsylvania, Working paper Series: 98 (02).

51. Ehrenberg, R. G., and Smith, R. S. (2000). *Modern Labour Economics: Theory and Public Policy*, Addison Wesley Longman, Inc.

52. Fields, G., and Mitchell, O. (1982). "Economic Determinants of the Optimal Retirement Age: An Empirical Investigation." *NBER Working Paper*, 876.

53. FIES. (2000). *The Family Income and Expenditure Survey*, Directorate-General of Budget, Accounting and Statistics, Taiwan.

54. Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*, John Wiley & Sons, Inc.

55. Goodman, R., and Peng, I. (1996). "The East Asia Welfare States: Peripatetic Learning, Adaptive Change, and Nation-Building." In Gosta, E. editor, *Welfare States in Transition: National Adaptations in Global Economies*, United Nations Research Institute for Social Development, SAGE Publications, 192-224.

56. Gordon, R., and Blinder, A. (1980). "Market Wages, Reservation Wages and Retirement Decisions." *Journal of Public Economics*, 14, 277-308.

57. Greene, W. H. (2000). *Econometric Analysis*, 4th edition, Prentice Hall International, Inc.

58. Gunderson, M. (1974). "Probit and Logit Estimates of Labour Force Participation." *Industrial Relations*, 19, 216-220.

59. Gunderson, M. (1977). "Logit Estimates of Labour Force Participation Based on Census Cross-Tabulations." *Canadian Journal of Economics*, 3, 453-462.

60. Gustman, A. L., and Steinmeier, T. L. (1982). "Partial retirement and the analysis of retirement behaviour." *NBER Working Paper*, 1000.

61. Gustman, A. L., and Steinmeier, T. L. (1986). "A Structural Retirement Model." *Econometrica*, 54(3), 555-584.

62. Gutierrez, R. G. (2002). "Parametric Frailty and Shared Frailty Survival Models." *The Stata Journal*, 1, 22-44.

63. Hall, A., and Johnson, T.R. (1980). "The Determinants of Planned Retirement Age." *Industrial and Labour Relations Review*, 33(2), 241-254.

64. Ham, J. C., and Rea, S.A. (1987). "Unemployment Insurance and Male Unemployment Duration in Canada." *Journal of Labor Economics*, 5, 325-353.

65. Hanoch, G., and Honig, M. (1983). "Retirement, Wages and Labour Supply of the Elderly." *Journal of Labour Economics*, 1(2), 131-151.

66. Hayashi, F. (2000). *Econometrics*, Princeton University Press.

67. Hayward, M. D., Hardy, M. A., and Liu, M. C. (1994). "Work after Retirement: The Experiences of Older Men in the United States," *Social Science Research*, 23, 82-107.

68. Heckman, J.J. (1981). "Heterogeneity and State Dependence," in *Studies in Labour Market*, ed. Rosen, S., University of Chicago Press, 91-139.

69. Heckman, J.J. (2000). "Microdata, Heterogeneity and the Evaluation of Public Policy." *Lecture to the memory of Alfred Nobel*, Aula Magna, Stockholm University, http://www.nobel.se/economics/laureates/2000/heckman-lecture.pdf.

70. Hermalin, A.I., Roan, C., and Chang, M. (1999). "The Economic Status of the Elderly in Taiwan." In Chen, C., Hermalin, A. I., Hu, S., and Smith, J. P., editors, *Emerging Social Economic Welfare Programs for Aging in Taiwan in a World Content*, The Institute of Economics, Academia Sinica, Taipei, Taiwan, 157-203.

71. Honig, M., and Reimers, C. (1987). "Retirement, Re-entry, and Part-time Work." *Eastern Economic Journal*, 13, 361-371.

72. Hosmer, D. W., and Lemeshow, S. (1999). *Applied Survival Analysis: Regression Modelling of Time to Event Data.* John Wiley & Sons, Inc.

73. Hosmer, D. W., and Lemeshow, S. (2000). *Applied Logistic Regression.* 2nd Edition, John Wiley & Sons, Inc.

74. Hsiao, C. (2003). *Analysis of Panel Data.* 2nd Edition. Cambridge University Press.

75. Hung, W. (2003a). "Probit Estimates of Labour Force Participation for the Middle Aged and Elderly in Taiwan." *Web Journal of Chinese Management Review*, 6(2), 1-16.

76. Hung, W. (2003b). "Semi-parametric Duration Analysis of Retirement Decisions for the Middle Aged and Elderly in Taiwan." *Web Journal of Chinese Management Review*, 6(6), 1-14.

77. Hurd, M. D. (1988). "The Joint Retirement Decisions of Husbands and Wives." *NBER Working Paper*, 2803.

78. Hurd, M. D. (1990). "Research on the Elderly: Economic Status, Retirement, and Consumption and Saving." *Journal of Economic Literature*, 28(2), 565-637.

79. Jenkins, S. P. (1992). "Lone Mothers' Employment and Full-Time Work Probabilities." *The Economic Journal*, 102, 310-320.

80. Jenkins, S. P. (2002). "Estimation of Discrete Time Proportional Hazards Model: pgmhaz." ESRC Research Centre on Micro-Social Change, University of Essex.

81. Jenkins, S. P. (2005). "*Stata Programs for Survival Analysis.*" Institute for Social and Economic Research, University of Essex. The website address is: http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/.

82. Jovanovic, B. (1984). "Matching, Turnover, and Unemployment." *The Journal of Political Economy*, 92(1), 108-122.

83. Kan, K., Park, A., and Chang, M. (2001). "A Dynamic Model of Elderly Living Arrangements in Taiwan." *Population Association of America*, Los Angeles, C.A.

84. Kiefer, N. M. (1988). "Economic Duration Data and Hazard Functions." *Journal of Economic Literature*, 26(2), 646-679.

85. Killingsworth, M. R., and Heckman, J. J. (1986). "Female Labour Supply: A Survey", *Handbook of Labour Economics*, Vol.1, Edited by O. Ashenfelter and R. Layard, 103-204.

86. Klein, J. P., and Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data.* Springer-Verlag. New York, Inc.

87. Kleinbaum, D. G., and Klein, M. (2005). *Survival Analysis: A Self-Learning Text.* Springer Science + Business Media. New York, Inc.

88. Kuan, C. (2003). "The Stylised Facts and Development of Retirement System in Taiwan." *Journal of Pension Fund*, 1(2), 3-8. (In Chinese)

89. Kuo, S. (1997). *Economic Policies: The Taiwan Experience: 1945-1995.* Hwa-Tai Publishing Company. Taiwan.

90. Lancaster, T. (1979). "Econometric Methods for the Duration of Unemployment." *Econometrica*, 47, 939-956.

91. Lancaster, T. (1990). *The Econometric Analysis of Transition Data.* Cambridge University Press.

92. Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data.* A John Wiley & Sons, Inc., Publication.

93. Lazear, E. P. (1986). "Retirement from the Labour Force." In Ashenfelter, O., and Card, D., editors, *Handbook of Labour Economics*, 1. Elsevier Publisher.

94. Long, J. E., and Jones, E. B. (1980). "Labour Force Entry and Exit by Married Women: A Longitudinal Analysis." *The Review of Economics and Statistics*, 62(1), 1-6.

95. Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables.* SAGE Publications.

96. Lumsdaine, R. L., and Mitchell, O. S. (1999). "New Developments in the Economic Analysis of Retirement." *Handbook of Labour Economics*, edited by Ashenfelter, O. and Card, D., Elsevier Publisher, 3C, 3261-3303.

97. Maddala, G.S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

98. McCall, B. P. (1990). "Occupational Matching: A test of Sorts." *The Journal of Political Economy*, 98(1), 45-69.

99. Mete, C., and Schultz, T. P. (2002). "Health and Labour Force Participation of the Elderly in Taiwan." *Centre Discussion Paper* No. 846, Economic Growth Centre, Yale University.

100. Meyer, B.D. (1990). "Unemployment Insurance and Unemployment Spells." *Econometrica*, 58, 757-782.

101. Miller, R. A. (1984). "Job Matching and Occupational Choice." The *Journal of Political Economy*, 92(6), 1086-1120.

102. Mitchell, O. S. and Fields, G. S. (1981). "The Effects of Pensions and Earnings on Retirement: A Review Essay." NBER Working Paper, 772.

103. Mitchell, O. S. and Fields, G. S. (1982). "Economic Determinants of the Optimal Retirement Age: An Empirical Investigation." NBER Working Paper, 876.

104. Nakamura, A., and Nakamura, M. (1981). "A Comparison of the Labour Force Behaviour of Married Women in the United States and Canada, with Special Attention to the Impact of Income Taxes." *Econometrica*, 49, 451-489.

105. Narendranathan, W. and Stewart, M. B. (1993). "How does the benefit effect vary as unemployment spells lengthen?" *Journal of Applied Econometrics*, 8(4), 361-381.

106. Nelson, W. (1972). "Theory and Applications of Hazard Plotting for Censored Failure Data." *Technometrics* 14, 945-965.

107. Nickell, S. J. (1979). "Estimating the Probability of Leaving Unemployment." *Econometrica*, 47, 1249-1266.

108. Nolan, M. A. (2000). "Spell durations and the impact of censoring." *Empirical Economics*, 25(4), 699-714.

109. Oi, W. Y. and Idson, T. L. (1999). "Firm Size and Wages." In Ashenfelter, O., and Card, D., editors, *Handbook of Labour Economics*, Elsevier Publisher, 3B, 2165-2214.

110. Pagan, A., (1984). "Econometric Issues in the Analysis of Regressions with Generated Regressors." *International Economic Review*, 25(1), 221-247.

111. Pellechio, A. (1978). "Social Security and the Decision to Retire." *NBER Working Paper*, 260.

112. Pencavel, J., (1986). "Labour Supply of Men: A Survey." In Ashenfelter, O., and Layard, R. editors, *Handbook of Labour Economics*, Elsevier Science Publishers BV, 1, 3-103.

113. Prentice, R., and Gloeckler, L. (1978). "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data." *Biometrics*, 34, 57-67.

114. Quinn, J. F. (1977). "Microeconomic Determinants of Early Retirement: A Cross-sectional View of White Married Men." *Journal of Human Resources*, 12, 329-346.

115. Raymo, J. M., and Cornman, J. C. (1999) "Trends in Labour Force Status Across the Life Course in Taiwan: 1970-1990." in Chen, C., Hermalin, A. I., Hu, S., and Smith, J. P., editors, *Emerging Social Economic Welfare Programs for Aging in Taiwan in a World Content*, The Institute of Economics, Academia Sinica, Taipei, Taiwan, 359-383.

116. RMUS. (2002). "*Report on the Manpower Utilization Survey*." Directorate-General of Budget, Accounting and Statistics, Taiwan, 8.

117. Ross, C. and Wu, C. (1995). "The Links between Education and Health." *American Sociological Review*, 60(5), 719-745.

118. Rust, J. (1989). "A Dynamic Programming Model of Retirement Behaviour." *NBER Working Paper*, 2470.

119. San, G. (1988). "The Effect of Income Taxes on Female Labour Supply in Taiwan." in Liu, P. K. C., Lee, Y., and Wu, C., editors, *Economic Development and Social Welfare in Taiwan*, The Institute of Economics, Academia Sinica, Taipei, Taiwan, 613-629.

120. Schmidt, P. and Strauss, R. P. (1975) "The Prediction of Occupation Using Multiple Logit Models", *International Economic Review*, 16(2), 471-486.

121. Schoenbaum, M. (1995). "The Health Status and Labour Force Behaviour of the Elderly in Taiwan." Research Report, Ann Arbour, University of Michigan, 95-32.

122. Shih, S. R. (1999). "Pension and Retirement Decisions of Older Men in Taiwan." in Chen, C., Hermalin, A. I., Hu, S., and Smith, J. P., editors, *Emerging Social Economic Welfare Programs for Aging in Taiwan in a World Content*, The Institute of Economics, Academia Sinica, Taipei, Taiwan, 385-414.

123. SHLS. (1996). "Survey of Health and Living Status of the Middle Aged and Elderly in Taiwan." *Comparative Study of the Elderly in Four Asian Countries Taiwan Ageing Studies Series.* 7-1.

124. SHLS. (1999). "Survey of Health and Living Status of the Middle Aged and Elderly in Taiwan." *Comparative Study of the Elderly in Four Asian Countries Taiwan Ageing Studies Series.*

125. Singer, J. D. and Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modelling Change and Event Occurrence.* Oxford University Press, Inc.

126. Slade, F. P. (1982). "Labour Force Entry and Exit of Older Men: A Longitudinal Study" *NBER Working Paper Series*, 1029.

127. Slade, F. P. (1987). "Retirement Status and State Dependence: A Longitudinal Study of Older Men." *Journal of Labor Economics*, 5(1), 90-105.

128. Social Indicators. (2000). *Social Indicators: The Republic of China, 2000*, Directorate-General of Budget, Accounting and Statistics, Taiwan.

129. Spataro, L. (2002). "New Tools in Micro-modelling Retirement Decisions: Overview and Applications to the Italian Case." *Centre for Research on pensions and Welfare Policies*, working paper 28/02, Department of Economics, University of Pisa, Italy.

130. Stata. (2001). *Stata Statistical Software: Release 7.0.* College Station, TX: Stata Corporation.

131. Stern, S. (1998). *Lecture Notes in Labour Economics*, Chapter 9, Human Capital. http://www.people.virginia.edu/~sns5r/classes/undergrad/econ415stf/econ415.html.

132. Steuerle, E., Spiro, C., and Carasso, A. (2000). "Measuring Replacement Rates at Retirement." *Straight Talk on Social Security and Retirement Policy*, Urban Institute, 24.

133. Stock, J. H., and Wise, D. A. (1990). "Pensions, the Option Value of Work and Retirement." *Econometrica.* 58(5), 1151-1180.

134. Tiebout, C. M. (1956). "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64, 416-24.

135. Uccello, C.E. (1998). "Factors Influencing Retirement: Their Implications for Raising Retirement Age." *American Association of Retired Persons*, USA.

136. Vere, J. P., and Wong, G. (2002). "Women's Labour Force Participation and Occupational Choice in Taiwan." *Hong Kong Institute of Economics and Business Strategy*, working paper, 1043.

137. Wang, L. (2000). "Annuity and Financial Schemes of Retirement." *Journal of Pension Fund*, 1(2), 9-28 (in Chinese).

138. Willmore, L. (2000). "Three Pillars of Pensions? A Proposal to End Mandatory Contributions." DESA, Discussion Paper, 13, United Nations.

139. Wooldridge, J. M., (1997). Introductory Econometrics: A Modern Approach. Michigan State University.

140. World Bank. (1993). "The East Asian Miracle: Economic Growth and Public Policy", in *A World Bank Policy Research Report*, Oxford University Press, 162.

141. World Bank. (1994). "*Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth*", in *A World Bank Policy Research Report*, Oxford University Press.

142. YMSS. (2002). "*Yearbook of Manpower Survey Statistics, Taiwan Area, Republic of China, 2001.*" Directorate-General of Budget, Accounting and Statistics, Taiwan.

143. Yu, R., and Chang, Y. (2004) "The Retirement Behaviour of Public Sector Employees in Taiwan." *Working Paper*, The Institute of Economics, Academia Sinica, Taipei, Taiwan (in Chinese).

144. Zimmer, Z., and Liu, X. (1999). "A Comparison of the Impact of Education on Functional Limitations Across Three Countries: Taiwan, Thailand and the Philippines." in Chen, C., Hermalin, A. I., Hu, S., and Smith, J. P., editors, *Emerging Social Economic Welfare Programs for Aging in Taiwan in a World Content*, The Institute of Economics, Academia Sinica, Taipei, Taiwan, 305-335.

145. Zimmer, Z., and Lin, H.S. (2000). "Changes in Prevalence and Transition Rates of Functioning Difficulties, and Limitation Severity, Among Older Adults in Taiwan: 1989 to 1996." *Population Studies Centre*, at the Institute for Social Research, University of Michigan.