Job performance and career prospects of auditors
Jonker, N.

Citation for published version (APA):
Chapter 8

Tenure

8.1 Introduction

In this chapter tenure of differently educated auditors in the Netherlands is analyzed. Since education is chosen by the individual himself and not given by nature it may be considered to be an endogenous variable (Willis and Rosen, 1979). The objective of this study is to investigate whether the more general educated auditors with a full-time education switch easier from employer than the part-time Nivra educated auditors. This latter group has relatively much firm-specific human capital whereas the full-time educated auditors possess relatively much general human capital. According to human capital theory, this may make it possible for the full-time educated auditors to transfer a higher amount of their human capital to other firms than for the Nivra educated auditors. Hence, it may be easier/less costly to find another job (better or comparable with the current job) for full-time educated auditors than the Nivra educated auditors. We investigate whether full-time educated auditors indeed switch jobs quicker than Nivra educated auditors by comparing tenures of the two groups of auditors. Tenure measures the number of years someone has worked/works with the firm where (s)he has worked at January 1990. Tenure is analyzed in a duration model framework. In the econometric model possible endogeneity of type of education is taken into account.

In medical research the type of treatment someone receives in a medical experiment is exogenous because of the random assignment of treatment type to the people in the experiment. In such a setting treatment is one of the exogenous covariates in the model. However, if one wants to evaluate the effect of a particular treatment in a non-experimental setting, the type of treatment individuals have received is not determined by a random process. If the decision of an individual to receive a particular treatment or not is based on self-selection,
treatment will not be exogenous. Under self-selection individuals who choose for a particular treatment may have a comparative advantage with that particular treatment, i.e. they will benefit more from that particular treatment than a randomly selected individual with the same observable characteristics would do. Then, unobserved heterogeneity of the individuals determines both the choice of treatment and the outcome under investigation. If one does not correct for endogeneity in treatment choice biased estimates of the treatment effect on the outcome under investigation will result.

In economic literature much attention has been paid to evaluating the effect of training on (social) welfare in linear models, see e.g. Maddala (1983) or Heckman and Smith (1996) for a survey. Little attention has been paid to evaluating the treatment effects in nonlinear models like duration models. Gritz (1993) is one of the first who incorporated endogeneity in a duration framework. He analyzed the effect of training on the frequency and duration of employment. Gritz allowed for endogeneity of training by introducing ‘training’ as a separate labor market state and by specifying joint probabilities of observed transition times between the different states in his model.

Abbring and Van den Berg (1999) present a method where treatment effects are non-parametrically identified in duration models. It deals with inference on treatment effects when both the assignment and the outcome are duration processes. With their method it is possible to evaluate the treatment effect if no proper instruments are available to correct for self-selection into treatments. Their model is not very suitable when individuals have full control over which treatment they get and when they get it since they assume that there is a purely random component in the duration until treatment.

Another paper which deals with the problem of self-selection in duration models is Holm (1996). He has estimated the effect of a training program on the duration of search for an apprentice vacancy. He has allowed for endogeneity by imposing a joint probability of training choice and the duration of search. In his statistical model he has allowed the search duration for no training and training to have a different distribution function. However, in the model he has estimated the effect on duration from choosing training through a single dummy variable.

In this chapter a method is introduced which deals with possible endogenous treatment effects in duration models. The model is based on a method discussed in Lee (1983). We assume that individuals made their educational choice before they started working at the 1990 employer. With Lee's method it is possible to correct for self-selection in a duration model for any distribution of the error term in the selection equation and for any distribution of the duration. A major advantage of Lee's method is that it is easy to use and yet quite flexible. The joint distribution function of the error term of the selection equation and the
duration of the event under investigation may be very complicated. However, if univariate distribution functions are transformed to standard normal variables the joint distribution of the transformed variables have a bivariate standard normal distribution, which is easy to deal with in estimations.

The outline of this chapter is as follows. Section 8.2 shows how Lee's method is used to derive a duration model which allows for endogenous treatment effects. In section 8.3 a description of the data is given and in section 8.4 the estimation results are shown. Section 8.5 concludes.

### 8.2 Econometric model

In this section an outline is given of the econometric model used in the estimations. A detailed derivation of the model can be found in appendix 8 B of this chapter. The basic of this econometric model is the joint distribution function of tenure and education. Suppose that there are two types of accountancy training, p (part-time Nivra) and f (full-time). Individuals who want to become auditor are not randomly assigned to these two types of education but they choose the education they want to have. Let $I^*_i$ be the latent educational choice variable and $Z_i$ be the vector of exogenous variables determining educational choice:

$$I^*_i = \gamma Z_i + \varepsilon_i$$  \hspace{1cm} (8.1)

with $\varepsilon_i$ identically and independently $F_{\varepsilon}$ distributed. Let $I_i$ represent the observed educational choice. $I_{1i} = 1$ if individual i has part-time Nivra education and $I_{1i} = 0$ if individual i has chosen full-time education:

$$I_{1i} = 1 \text{ iff } I^*_i > 0$$

$$I_{1i} = 0 \text{ iff } I^*_i \leq 0$$

We assume that each education has its own cumulative distribution function $F_j$ of $T_j$ with $j=p$ or $f$. If education affects tenure then this is reflected by differences between the distribution functions $F_p$ and $F_f$.

$$T_p - F_p(T_p) \text{ and } T_f - F_f(T_f)$$  \hspace{1cm} (8.2)

The joint distribution function of educational type and tenure can be obtained by using the method suggested by Lee (1983), which is also discussed in Maddala (1983). Lee proposes a method to transform two continuous random variables with known marginal distributions into a bivariate distribution in which the random variables are allowed to correlate. This transformation is very useful if one needs the joint distribution function of the marginal distribution functions
of two random variables from different families. By transforming these marginal
distribution functions to the standard normal distribution the joint distribution
function of the two transformed random variables is the standard bivariate dis­
tribution function. The standard bivariate distribution function is easy to deal
with in the estimations of the model in contrast to the joint distribution of the
original marginal distribution functions\(^1\).

Due to the possibility of right-censored tenures there are four groups of in­
dividuals to distinguish:

1. individuals with part-time Nivra education and observed tenures: \( T_p = t_i \)
and \( \varepsilon_i > -\gamma Z_i \)

2. individuals with full-time education and observed tenures: \( T_f = t_i \) and \( \varepsilon_i <= -\gamma Z_i \)

3. individuals with part-time Nivra education and right-censored tenures after
\( t_c \) years: \( T_p > t_c \) and \( \varepsilon_i > -\gamma Z_i \)

4. individual with full-time education and right-censored tenures after \( t_c \)
years: \( T_f > t_c \) and \( \varepsilon_i <= -\gamma Z_i \)

Define the dummy variable \( I_{2i} \) to indicate whether an individual's duration
is known (\( I_{2i} = 1 \)) or whether it is right-hand censored (\( I_{2i} = 0 \)) and let \( N \) be
the number of observations. Then the basic form of the log likelihood function
is as follows:

\[
\log L = \sum_{i=1}^{N} \left[ I_{1i}I_{2i} \log(\ell_{1i}) + (1 - I_{1i})I_{2i} \log(\ell_{2i}) \right] \\
+ I_{1i}(1 - I_{2i}) \log(\ell_{3i}) + (1 - I_{1i})(1 - I_{2i}) \log(\ell_{4i})
\] \hspace{1cm} (8.3)

In equation 8.3 the contribution to the likelihood function of individual \( i \) belong­
ing to group \( k, k=1..4 \) is denoted by \( l_{ki} \). The specification of the contributions
to the log likelihood function of these four groups differ.

\(^1\)Van Ophem and Jonker (1997) have used Lee's method when analyzing study dura­
tion in higher education. There, they used a semi-parametric dependent competing risks
model. Van Ophem (1999) shows that Lee's transformation method is not restricted to
deriving the joint distribution function of two continuous random variables but can be
extended to deriving the joint distribution function of any combination of continuous
and discrete random variables.
8.3 Data and variables

8.3.1 Data

The data used are from the auditors survey (see chapter 5 for a description of this data set). Individuals are selected with either a full-time or a part-time Nivra education. Auditors whose tenure in 1990 was unknown or with missing values on explanatory variables are excluded from the sample. Auditors (33) who were self-employed (zelfstandig gevestigde accountants) in 1990 are also excluded from the sample\(^2\). This resulted in 927 auditors whose records are used in the analysis. Tenure measures the number of years someone has worked/works with the firm where (s)he has worked at January 1990. It starts from the moment someone started working for the 1990 employer and it ends at the moment (s)he left the 1990 employer or at December 1998 depending on whether tenure was completed or not at December 1998. In this sample 255 auditors have a part-time Nivra education and a completed 1990 tenure, 219 auditors have a full-time training and a completed 1990 tenure, 284 auditors have a Nivra education and work at the same employer as at January 1990 and there are 169 auditors with a full-time training and who still work at the same employer as at January 1990.

Figure 8.1 shows the empirical distribution of tenure. The two upper graphs refer to completed tenures and the two lower graphs refer to uncompleted tenures. In the two upper graphs it is shown that only a small fraction of the people have a tenure of three years or less or a tenure larger than 10 (full-time education) or 15 (Nivra education) years. In the two lower graphs it is shown that most censored tenures lie between 9 and 15, but still a substantial proportion of the censored tenures lies beyond 15 years. This is especially true for the Nivra educated auditors\(^3\). The average uncensored tenures of Nivra educated auditors is 10.1 years (standard deviation 6.1 years) which is much higher than the average uncensored tenures of full-time educated auditors which is 7.0 years (standard deviation 4.5 years). A possible explanation is that Nivra educated auditors start working at a younger age than full-time educated auditors.

Figure 8.2 shows the Kaplan-Meier estimates of the survivor function of tenure for both educational types. The survivor functions become steeper during

---

\(^2\)This group of employees works in public auditing but they work on their own and are not employed by an audit firm. It is expected that these self-employed auditors behave differently at the labor market than the other auditors (other factors influencing the decision to start working elsewhere than for auditors working in an audit firm). Therefore they have been excluded in the analysis.

\(^3\)The lower bound of 9 years refers to people who have just started working at their 1990 employer in January 1990 and who are still employed in December 1998, the month the survey was taken.
Figure 8.1: Distribution of observed tenures
the first years of tenure, then they become a straight diagonal line for a while and after 25-30 years they become less steep. In terms of hazard rates, the profile of the hazard rate seems to be increasing and reaches its top somewhere between 5-8 years. After that point the hazard rate decreases.

A distribution which allows for both monotonous and non-monotonous hazards and which is analytically tractable is the Burr distribution (Lancaster, 1990, p. 68). Another attractive feature of the Burr distribution is that it can deal with unobserved heterogeneity in the sample. Lancaster shows that if one assumes that the hazard rate and the integrated hazard function contain a multiplicative individual specific random term \( v \) which follows a \( \Gamma(1, \sigma^{-2}) \) distribution such that \( \lambda(t, v) = \tilde{\lambda}(t)v \) and \( z(t, v) = \tilde{z}(t)v \) and if one assumes that the integrated hazard has the specification \( \tilde{z}(t) = \gamma t^\alpha \ (\alpha > 0) \), then one gets a Gamma mixture of Weibull distributions which is called the Burr(\( \gamma, \alpha, \eta \)) distribution with \( \eta = \sigma^{-2} \).

The Burr distribution is a quite general distribution and it has the following cumulative distribution, density and hazard function:
\[ F(T) = 1 - (1 + \sigma^2 \gamma T^\alpha)^{-\eta} \]  
(8.4)

\[ f(T) = \frac{\gamma \alpha T^{\alpha-1}}{(1 + \sigma^2 \gamma T^\alpha)^{\eta+1}} \]  
(8.5)

\[ \lambda(T) = \frac{\gamma \alpha T^{\alpha-1}}{(1 + \sigma^2 \gamma T^\alpha)} \]  
(8.6)

with \( \eta = \sigma^{-2} \). For \( \alpha > 1 \) the hazard function first increases with duration and then decreases. For \( 0 \leq \alpha \leq 1 \) the hazard function decreases with duration. The Burr distribution has the Log-Logistic distribution (non-monotonous hazard, \( \sigma^2 = 1 \)), the Weibull distribution (monotonous hazard function, \( \sigma^2 = 0 \)) and the exponential distribution (constant hazard function, \( \sigma^2 = 0 \) and \( \alpha = 1 \)) as special cases. In order to relate the effect of the individual specific explanatory variables on the duration of tenure we adopt the following often used assumption \( \gamma = \exp(X_i \beta) \), where \( X_i \) is a \( k \times 1 \) vector storing the values of the explanatory variables of individual \( i \).

### 8.3.2 Variables

The dependent duration variable in the model is tenure and the treatment variable is type of accountancy training. Tenure refers to the amount of time one has worked/works with the employer where one has worked at January 1990. Tenure is measured in years. The explanatory variables of tenure all date before the individuals started working at the 1990 employer/firm. So the explanatory variables are time-invariant. The education variable is a dummy variable called Nivra education; it equals 1 if one has part-time Nivra accountancy education and it equals 0 if one has full-time accountancy training.

In the tenure equation the following explanatory variables are used: gender, a dummy variable called 'children' indicating whether someone has children or not, number of years of work experience since started working in auditing (tenure at 1990 firm not included) and its square, three sector dummies (public sector + research, internal auditing and non-auditing job in financial sec-

---

4Education and research in accountancy usually takes place at universities. Since people employed by the universities are civil servants we included them in the public sector. Furthermore, we have lumped together auditors working as auditor with auditors who work in non-auditing jobs in the public sector. This has been done because of the low number (only 9) of auditors in non-auditing jobs in the public sector. In other chapters this problem did not occur and there these two groups have been considered as two separate sectors.
tor, reference group: public auditing), four job level dummies with job complexity/management responsibility increasing with job level (reference group is job level 1 consisting of assistants) and a dummy variable called 'not graduated' if the year in which the respondent finished his/her accountancy training was higher than the year in which (s)he started working at the 1990 employer.

The following variables are used in the education equation: gender, a dummy variable called 'not living near university' indicating whether someone lived more than 10 km from a university city with an economics department during his/her final year at pre-university education, a dummy variable called 're-examination' indicating whether someone has done a re-examination at pre-university education (is meant to reflect someone's intelligence/efficiency in studying)\(^5\), two dummies indicating the respondent's father educational level (lower vocational education or less and higher vocational education or more (reference is secondary general education or intermediate vocational education) and a dummy variable called 'started late' indicating whether there was a gap of two years or more between passing pre-university education and starting with the accountancy education\(^6\).

### 8.4 Empirical results

In table 8.1 the estimation results are shown. The estimated correlations between education and tenure are significant at a 10% level. There is a strong and negative correlation between the error term in the education equation and tenure of Nivra educated auditors. The correlation between the error term of the education equation and tenure of the full-time educated auditors is positive and significant at the 10% but not at the 5% level. The correlations imply the following\(^7\): Nivra educated auditors have a shorter tenure than the average auditor would have had in case of Nivra education and the full-time educated auditors have a shorter tenure than the average auditor would have had in case of a full-time education. The estimates of the coefficients of the \(\alpha\)'s are significantly higher than 1 indicating that the hazard of changing employer is

---

\(^5\)Another variable indicating ability in the data set is based on the average grade in the examination year at secondary education. This variable was also used in estimations but was very insignificant and has not been used anymore in later estimations.

\(^6\)Gaps of two years or more between graduating at secondary education and starting accountancy training may refer to for example military service, having attended another education before studying accountancy, started working after secondary education before deciding which study in higher education one wanted to do, etc.

\(^7\)For the interpretation of the effects of the correlations on expected tenure, see Maddala (1983), p. 367.
non-monotonous in time (this was already shown in figure 8.1 and 8.2). Furthermore, the heterogeneity terms \( \eta_{ij}, j=p, f \) are highly significant indicating that there is unobserved heterogeneity in the sample.

In order to have interpretable estimates in the tenure equation the effect of the explanatory variables on the logarithm of the expected duration are calculated:

\[
\frac{\partial \ln(E(T_{ij}))}{\partial X_{ij}} = -\frac{\beta_j}{\alpha_j}, \ j = p, f \tag{8.7}
\]

The estimation results of the tenure equation are quite similar for the two types of education: 10 out of 12 explanatory variables have the same sign. Nivra educated auditors with children stay longer than auditors without children, so having children reduces job mobility. Number of years of previous work experience has an "u-shaped" effect on tenure indicating that the amount of previous experience first decreases tenure but that there is a turning point after which the amount of previous experience has a positive effect on tenure. In all sectors people have a longer tenure than in public auditing, the reference sector. This may be the result of the 'up-or-out' culture in this sector which reduces tenure. Job level (reference category: assistant) seems to affect tenure negatively but this effect is not significant. Only Nivra educated auditors who start in job level 3 stay significantly shorter than employees who started as assistant. Nivra educated auditors who were still studying when they started working at the 1990 employer do not differ in tenure from auditors who were auditor when they became employed.

There are few variables affecting tenure of full-time educated auditors; only gender (which did not play a role in explaining tenure of Nivra educated auditors), experience and not being graduated have a significant effect. Women stay shorter than men. Experience has a negative and significant effect on tenure and its square has a positive and significant effect. So there is an 'u-shaped' effect of experience on tenure just as with the Nivra educated auditors. Full-time educated auditors who were still studying accountancy when they started working at the 1990 employer stay significantly shorter than auditors who had already finished their accountancy training when they became employed. Job level and sector have no significant effect.
### Table 8.1
ML estimation results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Nivra education</th>
<th>full-time education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T: -\beta_p/\alpha_p$</td>
<td>$T: -\beta_f/\alpha_f$</td>
</tr>
<tr>
<td><strong>tenure equation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.814**</td>
<td>2.815**</td>
</tr>
<tr>
<td>Female</td>
<td>-0.004</td>
<td>-0.663*</td>
</tr>
<tr>
<td>Experience/10</td>
<td>0.431*</td>
<td>0.297</td>
</tr>
<tr>
<td>Experience$^2$/100</td>
<td>-0.681*</td>
<td>-1.660*</td>
</tr>
<tr>
<td>Public sector/research</td>
<td>0.368*</td>
<td>0.353</td>
</tr>
<tr>
<td>Internal auditing</td>
<td>0.529*</td>
<td>0.201</td>
</tr>
<tr>
<td>Financial sector</td>
<td>0.253</td>
<td>0.171</td>
</tr>
<tr>
<td>Job level 2</td>
<td>-0.464</td>
<td>-0.049</td>
</tr>
<tr>
<td>Job level 3</td>
<td>-0.674*</td>
<td>-0.198</td>
</tr>
<tr>
<td>Job level 4</td>
<td>-0.739</td>
<td>0.045</td>
</tr>
<tr>
<td>Job level 5</td>
<td>-0.148</td>
<td>0.609</td>
</tr>
<tr>
<td>Not graduated</td>
<td>0.230</td>
<td>-0.498*</td>
</tr>
<tr>
<td><strong>Education equation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.003</td>
<td>(0.099)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.228</td>
<td>(0.207)</td>
</tr>
<tr>
<td>Not living near university</td>
<td>0.196*</td>
<td>(0.096)</td>
</tr>
<tr>
<td>Education father low</td>
<td>0.244**</td>
<td>(0.090)</td>
</tr>
<tr>
<td>Education father high</td>
<td>-0.394**</td>
<td>(0.110)</td>
</tr>
<tr>
<td>Re-examination</td>
<td>0.173</td>
<td>(0.190)</td>
</tr>
<tr>
<td>Started late</td>
<td>0.183*</td>
<td>(0.102)</td>
</tr>
<tr>
<td><strong>Parameters Burr distribution</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_p, \alpha_f$</td>
<td>2.283**</td>
<td>2.309**</td>
</tr>
<tr>
<td>$\eta_p, \eta_f$</td>
<td>0.198**</td>
<td>0.232**</td>
</tr>
<tr>
<td>$\rho_p, \rho_f$</td>
<td>-0.591**</td>
<td>0.447*</td>
</tr>
</tbody>
</table>

Loglikelihood: -2528.727; 927 observations

* (** ) = significant at 10% (1%)

$\chi^2$ test have been performed to test the hypotheses of equal coefficients in the tenure equations. Testing the hypotheses of equal coefficients in the tenure equation (intercepts excluded) resulted in a $\chi^2_{12}=7.571$ which is not significant. However, testing the hypothesis of equal coefficients of the significant variables in the tenure equation resulted in a $\chi^2=102.541$ which is highly significant.
Equality of the coefficients of the significant variables in the tenure equation is rejected.

In the education equation the following variables affect educational choice significantly: education of the father, residence during final year at pre-university education and not immediately started studying accountancy after pre-university training. People who have lived near a university offering a training in accountancy have chosen more often for a full-time (university) training in accountancy than others. The higher the educational level of the respondent’s father the more likely (s)he has had university training. Furthermore, those with 'started later'=1 are more likely to have a Nivra training. In this group there are people who are married and have to earn a living, who had served in the military and people who had already finished another education with poor labor market perspectives and could not afford another full-time education. Gender and having done a re-examination do not play a role in educational choice.

In figure 8.3 the estimated hazard rates and survivor functions are shown for each educational type. The estimated survivor functions are very similar to the estimated Kaplan-Meier survivor functions of figure 8.2. As could already be seen in figures 8.1 and 8.2 the hazard rates are non-monotonous and have their peaks at about 9 years for the Nivra educated and 5 years for the full-time educated. The estimated hazard rate of the full-time educated auditors is higher than the estimated hazard rate of the Nivra educated auditors, especially for the first ten years. The survivor function of the full-time educated auditors lies below the survivor function of the Nivra educated auditors. It declines much sharper during the first ten years than the survivor function of the Nivra educated auditors. This means that full-time educated auditors stay shorter at a particular firm than Nivra educated auditors. In particular during the first ten years of tenure full-time educated auditors are more likely to leave the firm than the Nivra educated.

It is interesting to know what causes the different shapes of the education specific hazard rates. According to table 8.1 the duration dependence parameters $\alpha_j$ ($j=p,f$) are not significantly different from each other and the same holds for the heterogeneity terms $\eta_j$ ($j=p,f$). Differences in hazard rates seem to stem from differences in the $\gamma_j$’s which are related to the personal characteristics of the auditors: $\gamma_j = \exp(X_j \beta_j)$. Both the values of the explanatory variables $X_j$ and the tenure coefficients $\beta_j$ may cause differences in $\gamma_j$. It has already been noted that the significant variables in the tenure equations are statistically different. In order to see whether differences in explanatory variables cause differences in $\gamma_j$ the average $\gamma_j$’s are calculated (for each educational type $j$ the average $\gamma_j$ is calculated with the characteristics of both Nivra and full-time educated auditors). They are listed in table 8.2. The average $\gamma_j$’s differ by
Figure 8.3: Estimated hazard rates and survivor functions
educational type (i.e. $\beta_j$) but they do not differ by the personal characteristics of the differently educated auditors. For both types of students the average value of $\gamma_p$ equals about 0.0015 and the average value of $\gamma_f$ equals about 0.004. The difference in hazard rates (and survivor functions) shown in figure 8.3 stem from differences in estimated coefficients in the tenure equation $\beta_j$ and not from differences in the personal characteristics of the differently educated auditors.

Table 8.2
Average $\gamma$'s for each education type

<table>
<thead>
<tr>
<th>Variable matrix</th>
<th>$\gamma_p = \exp(X_i\beta_p)$</th>
<th>$\gamma_f = \exp(X_i\beta_f)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_p$</td>
<td>0.001596</td>
<td>0.004030</td>
</tr>
<tr>
<td>$X_f$</td>
<td>0.001504</td>
<td>0.003961</td>
</tr>
</tbody>
</table>

8.5 Summary and concluding remarks

In this chapter tenures of differently educated auditors have been compared. The full-time accountancy training has a more general character than the Nivra training whereas students from the Nivra education have relatively much firm-specific human capital. This may make full-time educated auditors more job mobile than Nivra educated auditors since a larger amount of their human capital is transferable to other firms.

A duration model has been presented with education specific distribution functions of tenure and which allows for endogeneity of educational type. The distribution of the duration variable and the selection variable are assumed to be known univariate distributions. The univariate distributions are the marginals of the joint distribution function. Since this joint distribution will often not be estimable directly we propose to use Lee's method (1983) to transform the original joint distribution of duration and treatment to the bivariate normal distribution. With this transformation estimation of duration models with endogenous treatments becomes very simply.

The estimation results support the idea that job mobility among full-time educated auditors is higher than among Nivra educated auditors. Tenure of full-time educated auditors is shorter than tenure of Nivra educated auditors. The more general full-time accountancy training makes it for its graduates in accountancy easier to transfer their skills to other than the current firm than for Nivra educated auditors.
### Appendix 8 A

#### Table 8 A
Summary statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Nivra education</th>
<th>full-time education</th>
</tr>
</thead>
<tbody>
<tr>
<td>tenure</td>
<td>15.362 (8.796)</td>
<td>10.744 (6.846)</td>
</tr>
<tr>
<td>female</td>
<td>0.035 (0.185)</td>
<td>0.0515 (0.221)</td>
</tr>
<tr>
<td>children=1</td>
<td>0.349 (0.477)</td>
<td>0.273 (0.446)</td>
</tr>
<tr>
<td>experience/10</td>
<td>0.637 0.727</td>
<td>0.259 (0.442)</td>
</tr>
<tr>
<td>experience²/100</td>
<td>0.934 (1.456)</td>
<td>0.262 (0.689)</td>
</tr>
<tr>
<td>public auditing</td>
<td>0.568 (0.496)</td>
<td>0.593 (0.492)</td>
</tr>
<tr>
<td>public sector</td>
<td>0.163 (0.370)</td>
<td>0.111 (0.314)</td>
</tr>
<tr>
<td>internal auditing</td>
<td>0.093 (0.290)</td>
<td>0.067 (0.250)</td>
</tr>
<tr>
<td>financial sector</td>
<td>0.176 (0.381)</td>
<td>0.229 (0.421)</td>
</tr>
<tr>
<td>job level 1</td>
<td>0.466 (0.499)</td>
<td>0.523 (0.500)</td>
</tr>
<tr>
<td>job level 2</td>
<td>0.148 (0.356)</td>
<td>0.130 (0.336)</td>
</tr>
<tr>
<td>job level 3</td>
<td>0.221 (0.415)</td>
<td>0.222 (0.416)</td>
</tr>
<tr>
<td>job level 4</td>
<td>0.095 (0.293)</td>
<td>0.083 (0.275)</td>
</tr>
<tr>
<td>job level 5</td>
<td>0.071 (0.256)</td>
<td>0.044 (0.205)</td>
</tr>
<tr>
<td>not graduated</td>
<td>0.731 (0.444)</td>
<td>0.665 (0.473)</td>
</tr>
<tr>
<td>not living near university</td>
<td>0.787 (0.410)</td>
<td>0.729 (0.445)</td>
</tr>
<tr>
<td>education father low</td>
<td>0.477 (0.500)</td>
<td>0.325 (0.469)</td>
</tr>
<tr>
<td>education father intern.</td>
<td>0.362 (0.481)</td>
<td>0.376 (0.485)</td>
</tr>
<tr>
<td>education father high</td>
<td>0.161 (0.368)</td>
<td>0.299 (0.458)</td>
</tr>
<tr>
<td>re-examination</td>
<td>0.056 (0.230)</td>
<td>0.046 (0.211)</td>
</tr>
<tr>
<td>started late</td>
<td>0.245 (0.430)</td>
<td>0.219 (0.414)</td>
</tr>
</tbody>
</table>

n 539 388
Appendix 8 B

In this appendix the technical details of the log likelihood are given. There are two types of accountancy training, p (part-time Nivra education) and f (full-time university education). Remind from section 8.2 that:

\[ I_{it}^* = \gamma Z_i + \varepsilon_i \]  

(a 8.1)

with \( \varepsilon_i \) identically and independently \( F_\varepsilon \) distributed. \( I_{it} \) represents the observed educational choice with \( I_{1i} = 1 \) if individual \( i \) has Nivra education and \( I_{1i} = 0 \) if individual \( i \) has full-time education:

\[ I_{1i} = 1 \text{ iff } I_{1i}^* > 0 \]
\[ I_{1i} = 0 \text{ iff } I_{1i}^* \leq 0 \]

In section 8.2 it has been assumed that each education has its own cumulative distribution function \( F_j \) of \( T_j \) with \( j=p \) or \( f \).

\[ T_p - F_p(T_p) \text{ and } T_f - F_f(T_f) \]  

(a 8.2)

The joint distribution function of educational type and tenure can be obtained by using the method suggested by Lee (1983), which is also discussed in Maddala (1983). Starting from assumptions a 8.1 and a 8.2 and denoting the correlation between \( T_p \) and \( \varepsilon \) by \( \rho_{1p} \) and the correlation between \( T_f \) and \( \varepsilon \) by \( \rho_{1f} \), the following variables are defined:

\[ \tau_p = J_p(T_p) = \Phi^{-1}(F_p(T_p)) \]  

(a 8.3)
\[ \tau_f = J_f(T_f) = \Phi^{-1}(F_f(T_f)) \]
\[ \tau_\varepsilon = J_\varepsilon(\varepsilon) = \Phi^{-1}(F_\varepsilon(\varepsilon)) \]

where \( \Phi^{-1}(.) \) is the inverse of the standard cumulative normal distribution function. The transformed variables \( \tau_p \), \( \tau_f \) and \( \tau_\varepsilon \) are standard normal random variables irrespective of the distributions of the original durations. The bivariate distributions having marginal distributions \( F_j \) and \( F_\varepsilon \) and correlation \( \rho_{1j} \) with \( j=p, f \) are given by:

\[ H_p(T_p, \varepsilon; \rho_{1p}) = B(\tau_p, \tau_\varepsilon; \rho_{2p}) = B(J_p(T_p), J_\varepsilon(\varepsilon); \rho_{2p}) \]  

(a 8.4)
\[ H_f(T_f, \varepsilon; \rho_{1f}) = B(\tau_f, \tau_\varepsilon; \rho_{2f}) = B(J_f(T_f), J_\varepsilon(\varepsilon); \rho_{2f}) \]

where \( B(.,.; \rho_{2j}) \) is the bivariate normal distribution with zero means, unit variances and correlation \( \rho_{2j} \). Note that through the transformation to normality of
of the original marginal distribution functions the correlation of the original distributions $H_p$ and $H_f$ is not the same as the correlation of the standard bivariate normal distribution of the transformed durations. The corresponding bivariate density functions equal:

\[
\begin{align*}
    h_p(T_p, \varepsilon; \rho_{1p}) &= \frac{f_p(T_p)}{\phi(J_p(T_p))} \frac{f_\varepsilon(\varepsilon)}{\phi(J_\varepsilon(\varepsilon))} b(J_p(T_p), J_\varepsilon(\varepsilon); \rho_{2p}) \\
    h_f(T_f, \varepsilon; \rho_{1f}) &= \frac{f_f(T_f)}{\phi(J_f(T_f))} \frac{f_\varepsilon(\varepsilon)}{\phi(J_\varepsilon(\varepsilon))} b(J_f(T_f), J_\varepsilon(\varepsilon); \rho_{2f})
\end{align*}
\]

where $f_p$, $f_f$ and $f_\varepsilon$ are the marginal density functions of $T_p$, $T_f$ and $\varepsilon$ and $b(\cdot; \cdot; \rho_{2j})$ is the bivariate standard normal density function of $(\tau_j, \varepsilon)$ with $j=p, f$.

The contribution to the likelihood function of person $i$ with Nivra education and a completed tenure $t_i$ is:

\[
\ell_{1i} = \int_{-\infty}^{\infty} h_p(t_i, \varepsilon_i; \rho_{1p}) d\varepsilon_i
\]

This contribution can be written as (cf. Maddala, 1983, p. 272):

\[
\ell_{1i} = \int_{-\infty}^{\infty} h_p(t_i, \varepsilon_i; \rho_{1p}) d\varepsilon_i - \int_{-\infty}^{\infty} h_p(t_i, \varepsilon_i; \rho_{1p}) d\varepsilon_i
\]

\[
= f_p(t_i) - \frac{\partial}{\partial T_p} H_p(T_p, \varepsilon_i; \rho_{1p}) \bigg|_{T_p=t_i}
\]

\[
= f_p(t_i) \left( 1 - \Phi \left( \frac{J_\varepsilon(-\varepsilon Z_i) - \rho_{2p} J_p(t_i)}{\sqrt{1 - \rho_{2p}^2}} \right) \right)
\]

The contribution of the likelihood function of an individual $i$ with full-time education and a completed tenure $t_i$ equals:

\[
\ell_{2i} = \int_{-\infty}^{\infty} h_f(t_i, \varepsilon_i; \rho_{1f}) d\varepsilon_i
\]

\[
= f_f(t_i) \Phi \left( \frac{J_\varepsilon(-\varepsilon Z_i) - \rho_{2f} J_f(t_i)}{\sqrt{1 - \rho_{2f}^2}} \right)
\]

Individuals with Nivra education and right-hand censored tenures after $t_c$ years have the following contribution to the likelihood function:
\[ \ell_{3i} = \int_{-\infty}^{\infty} \int_{t_c}^{\infty} h_p(T_{pi}, \varepsilon_i; \rho_{1p})dT_{pi}d\varepsilon_i \tag{a 8.9} \]
\[ = \int_{-\infty}^{\infty} \int_{t_c}^{\infty} b(\tau_{pi}, \tau_{\varepsilon_i}; \rho_{2p})d\tau_{pi}d\varepsilon_i \]
\[ = B(-J_p(t_c), -J_c(-\gamma Z_i); \rho_{2p}) \]

Analogously, the log likelihood contribution of individuals with full-time education and right-hand censored tenures after \( t_c \) years can be derived:

\[ \ell_{4i} = \int_{-\infty}^{\infty} \int_{t_c}^{\infty} h_f(T_{f_i}, \varepsilon_i; \rho_{1f})dT_{f_i}d\varepsilon_i \tag{a 8.10} \]
\[ = \int_{-\infty}^{\infty} \int_{t_c}^{\infty} b(\tau_{f_i}, \tau_{\varepsilon_i}; \rho_{2f})d\tau_{f_i}d\varepsilon_i \]
\[ = B(-J_f(t_c), J_c(-\gamma Z_i); -\rho_{2f}) \]

Define the dummy variable \( I_{2i} \) to indicate whether an individual’s duration is known \((I_{2i} = 1)\) or whether it is right-hand censored \((I_{2i} = 0)\) and let \( N \) be the number of observations, the log likelihood function will be as follows:

\[ \log L = \sum_{i=1}^{N} I_{1i}I_{2i}\log(\ell_{1i}) + (1 - I_{1i})I_{2i}\log(\ell_{2i}) + I_{1i}(1 - I_{2i})\log(\ell_{3i}) + (1 - I_{1i})(1 - I_{2i})\log(\ell_{4i}) \tag{a 8.11} \]