DE ECONOMIST 140, NR. 3, 1992

AN EMPIRICAL ANALYSIS OF UNIVERSITY CHOICE AND EARNINGS

BY

HESSEL OOSTERBEEK*, WIM GROOT** AND JOOP HARTOG*

1 INTRODUCTION

By introducing the concept of 'human capital,' economists express a view of education as an investment of current time and money in return for future pay (*cf.* Freeman (1986), p. 367). According to this human capital theory, there are two interdependent relationships between the amount of education an individual obtains and his or her future earnings. In the first, the level of earnings depends on the amount of education and other human capital variables such as experience and experience squared. This relationship has become extremely popular in empirical research due to work by Mincer (1974). Estimation of the Mincerian earnings equation *per se*, however, ignores the fact that the explanatory variable 'education' is not exogenous. On the contrary, the very notion of education as an investment implies a second relationship: namely, that individuals choose their optimal amount of education on the basis of their expectations with regard to future earnings (*cf.* Rosen (1977), Willis (1986)). As a result, the Mincerian earnings equation is only compatible with optimizing individual behaviour under restrictive assumptions.

Since Rosen's detection of this inconsistency in Mincer's schooling model, a number of contributions in the literature have focussed on this so-called self-selection problem in the context of educational choices.[1] Willis and Rosen (1979) analyze the choice of whether or not to attend college in the United States, and find that the expected earnings differential between college and high school has a significantly positive effect on the probability of choosing to do so. Similar results are reported by Micklewright (1989) for Great Britain and

* University of Amsterdam, Microeconomics Section, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands.

** University of Leiden, Department of Law.

1 In modern labour economics the problem of self-selection has been widely recognized; well-known examples include the choices of whether to join a union (Lee (1978)), and whether to work in the private or the public sector of the economy (Van der Gaag and Vijverberg (1987)).

by Oosterbeek (1990) for The Netherlands.[2] All these studies employ datasets containing observations from a cross-section of working individuals. Consequently, the information with respect to earnings does not really reflect expected earnings, but realized earnings predicted by the model. These models therefore operate on the strong assumption of *ex post* lack of bias in expectations, which is more stringent than the usual rational-expectations assumption of lack of bias *a priori* (*cf.* Manski and Wise (1983), p. 108).

The idea of education as an investment is not necessarily restricted to the choice of the level of education, but might also apply to an individual's decision to attend a particular university. In this paper we carry out an analysis of this particular decision. The questions addressed are therefore: (i) whether different earnings prospects are associated with attending different universities and, if so, (ii) whether the decision to choose a particular university is governed by these differing earnings prospects.

Both questions have an inherent policy interest. An assessment of earnings differentials among graduates from different universities might indicate quality differences between the various establishments. Earnings as an index of labour market performance are then viewed as a measure of post-academic performance. Information about the quality of the courses that universities offer is important from the point of view of university funding. At present, many western countries use input-financing schemes for the allocation of available public funds to their various universities. Discussions among public administrators, however, point to a shift towards output-financing schemes (*cf. Commissie Financieringsstructuur* (1985)). A necessary precondition for the implementation of such schemes is, however, the availability of information with respect to the quality of output.

The issue of whether decisions to attend a particular university are influenced by earnings prospects is also of interest from the point of view of a possible differentiation in tuition fees. To date, the level of tuition fees in The Netherlands has been decided by central government, and has been equal for all students (irrespective of field or phase of study). Recently, however policymakers have proposed giving establishments a measure of autonomy in this respect. Universities might then increase their tuition fees and use the extra revenues to improve the quality of their courses and thereby (presumably) the future earnings of their graduates. Such a policy would, however, only pay off if future earnings did in fact play a role in the choice of a particular university. If that were not the case, students would be reluctant to pay the higher tuition fees.

For the empirical investigation in this paper, we use a sample of economists who each graduated from one of the five economics departments in The Netherlands. This implies that the choice set of the respondents consists of five

2 Other contributions include Kenny *et al.* (1979), Garen (1984) and Hartog, Pfann and Ridder (1989).

options. To model this, we employ the multinomial logit model. As far as we know, this study is the first application of the multinomial logit model (which determines the probability of attendance at each specific university) to the analysis of university choice. In the study by Manski and Wise (1983) concerning college choice in the United States, students did not choose between specific establishments, but between quality levels (universities are ranked according to the average SAT score of freshmen entering).

The remainder of this paper is made up of four sections. In section 2 we describe the model, while in section 3 we offer a brief description of the data. In section 4 we present and discuss our empirical findings. Finally, section 5 summarizes our conclusions.

2 MODEL AND ESTIMATION METHOD

For the model in this paper we adopt the random utility framework devised by McFadden (1974). Assume that individuals aim to maximize expected utility. Let V_i represent the utility for the individual of choosing option i. We distinguish five different options: studying economics at the: (1) University of Amsterdam (UvA),[3] (2) Free University of Amsterdam (VU); (3) University of Groningen (RUG); (4) Erasmus University of Rotterdam (EUR), and (5) Catholic University of Brabant (KUB). Furthermore, assume that the indirect utility function V_i is linear in the log of the expected net present value of lifetime earnings ($logN_i$) and in the individual's characteristics X.

$$V_i = (logN_i)\gamma + X\boldsymbol{\beta}_i + u_i \tag{1}$$

where γ and $\boldsymbol{\beta}_i$ are (vectors of) parameters and u_i is a random term which reflects unobserved characteristics of the individual, unobserved taste variation and genuine indeterminacy of individual behaviour (*cf.* Cramer (1991), p. 49). University i is chosen if, and only if, studying economics at this university provides the highest utility level:

$$V_i > \max V_j \quad i, j = 1, 2, 3, 4, 5 \quad \text{and } i \neq j \tag{2}$$

The net present value of lifetime earning is, of course, not measured. For each option we calculate this variable as

$$N_i = \sum_t w_{it} h (1 + r)^{-t} \quad i = 1, 2, 3, 4, 5 \tag{3}$$

where t is a time index for years of working experience which ranges from 1 to 40, h is potential hours of work per year ($= 2000$), r is the discount rate which we set at 5 per cent, and w_{it} is the wage rate associated with option i in the t-th

3 We use the Dutch abbrevations.

year of experience. These university-specific expected wage rates are supposed to depend loglinearly upon a vector of individual characteristics Z, t and t^2:

$$log\ w_{it} = \boldsymbol{Z}\boldsymbol{\mu}_i + \alpha_{0i}t + \alpha_{1i}t^2 + \varepsilon_i, \quad i = 1,2,3,4,5 \tag{4}$$

where $\boldsymbol{\mu}_i$, α_{0i} and α_{1i} are (vectors of) parameters, and ε_i is a disturbance term. Substituting (4) into (3) and the result into (1) gives the reduced form of the indirect utility function[4]

$$V_i = \boldsymbol{W}\pi_i + \upsilon_i, \quad i = 1,2,3,4,5 \tag{5}$$

where $\boldsymbol{W} = [X, Z]$. We further assume that the disturbances $\upsilon_i (= \varepsilon_i\gamma + u_i)$ are independently and identically Gumbel distributed, where the distribution function is denoted by $F(\cdot)$. The probability that university i will be chosen, P_i, is then defined by[5]:

$$P_i = \{\textstyle\sum_j \exp(\boldsymbol{W}\boldsymbol{\pi}_j - \boldsymbol{W}\boldsymbol{\pi}_i)\}^{-1} \tag{6}$$

For each individual, w_i is only observed in one of the five possible wage regimes. Since individuals select themselves into one of the regimes by choosing their preferred university, the unobserved wages cannot be calculated by the OLS estimates of μ_i, α_{0i} and α_{1i} in equation (4) because of selection bias. Lee (1983) and Trost and Lee (1984) present a method for correcting for selection bias in wage equations in a model with more than two choices. We outline this method briefly in the paragraphs below.

First define $\eta_i = \max V_j - \upsilon_i$ $(j = 1,2,3,4,5;\ j \neq i)$. Let φ and Φ be the standard normal density function and the distribution function respectively. And let $J = \Phi^{-1}(F(\boldsymbol{W}\boldsymbol{\pi}_i))$. Lee (1983) defines the selectivity bias correction term by: $\varphi(J(\boldsymbol{W}\boldsymbol{\pi}_i))/F(\boldsymbol{W}\boldsymbol{\pi}_i)$. The selectivity-corrected wage rate equation of university i is then:

$$log\ w_i = \boldsymbol{Z}\boldsymbol{\mu}_i - \sigma_i\varrho_i\varphi(J(\boldsymbol{W}\boldsymbol{\pi}_i)) + \varepsilon_i \tag{7}$$

where the conditional expectation of ε_i is zero, σ_i is the standard deviation of ε_i, and ϱ_i the correlation coefficient of ε_i and $J(\eta_i)$.

The estimation procedure of equation (7) involves two stages. In the first stage, we estimate the reduced form of the decision model (5). Since vector $\boldsymbol{W}$ consists entirely of individual characteristics, we might refer to this model as the multinomial logit model. Estimation by the maximum likelihood method gives consistent estimates for $\boldsymbol{\pi}_i$. In the second stage, the estimated $\boldsymbol{\pi}_i$'s are

4 Substituting (4) into (3) gives $N_i = h\,exp(\boldsymbol{Z}\boldsymbol{\mu}_i + \varepsilon_i)\sum_t[exp(\alpha_{0i}t^2)(1+r)^{-t}]$, substitution of this expression into (1) gives $V_i = \gamma(logh + \boldsymbol{Z}\boldsymbol{\mu}_i + \varepsilon_i + log(\sum_t[exp(\alpha_{0i}t + \alpha_{1i}t^2)(1+r)^{-t}]) + \boldsymbol{X}\boldsymbol{\beta}_i + u_i$.

5 Cramer (1991, p. 50/1) gives the derivation of equation (6) from a random utility model.

substituted into equations (7) which in turn are estimated by ordinary least squares.

As a final step, we can compute for each individual the five university-specific lifetime earnings and include these in the vector of explanatory variables in the decision model. We can then estimate the structural decision function (1). Since the explanatory variables in this function are both choice-specific ($logN_i$) and individual-specific (X), we might refer to this model as the mixed logit model.[6]

A well-known property of the logit model is the independence of irrelevant alternatives (IIA). This property implies that the odds ratio of each two options in a choice set is determined exclusively by the characteristics of these two options. Adding or deleting an option will not affect the odds ratios (*cf.* Cramer (1991), p. 47). In the context of the present model this implies, among other things, that the choice between the five economic departments is independent of the possibility of choosing a department in some other field of study. Ignoring this possibility is equivalent to the assumption that the decision tree of our respondents has the sequential structure of a nested logit model. The decision as to which discipline to pursue (economics, history, science, *etc.*) is taken at the first node of the tree. Given that prior decision, the second node of the decision tree only involves the choice between departments offering courses in that preferred discipline. In this paper we restrict ourselves to the second node of the decision tree. Empirical evidence supporting the assumption of a sequential decision framework in this context is provided by Kodde and Theunissen (1985).

3 DESCRIPTION OF THE DATA AND CHOICE OF VARIABLES

In 1987 a questionnaire was distributed to all Dutch economists who received the monthly review "Economenblad". The data were collected by researchers from the Department of Economics and the Centre for Educational Research of the University of Amsterdam. The questionnaire included items on education, earnings, occupation, and various background variables, 2,060 (23 per cent) of the forms were returned. No systematic information is available with respect to the representativeness of the sample, although it is obvious from inspection of the sample means that people with doctoral degrees and university employees were over-represented, while the unemployed were under-represented.

The dependent variables in the empirical analysis are the wage rate and the university from which the respondent received his or her master's degree. The five Dutch universities with economics departments are listed at the beginning of the previous section. In addition, a very small number (less than 3 per cent)

6 A logit model in which the vector of explanatory variables consists entirely of choice-specific variables is called a conditional logit model.

of the respondents in the sample had graduated from a university other than the five with 'official' economics departments.[7] These respondents had degrees either from the Agricultural University of Wageningen or from a foreign university. We left this small group out of the analysis, because their inclusion often caused singularity in the dataset during an estimation procedure.

For the wage rate we used gross hourly wages. The questionnaire contains information on both net and gross earnings. With an individual choice framework, a net earnings concept might perhaps be preferable. But since we are also interested in quality assessment, and because the link between productivity and gross earnings is stronger than that between productivity and net earnings, we preferred a gross earnings concept. An additional reason for doing so was the large non-response on net earnings. Whereas 1,840 respondents reported their gross earnings, only 769 reported their net wage rate. This would only cause a minor problem if non-response were random. This is not the case, however, since non-response on net earnings is found to correlate positively with the value of gross earnings.

The variables thought to explain wages (vector Z) are:

- a dummy equal to 1 if the respondent is male;
- self-reported average grade at university: this variable is measured on an ordinal 1–6 scale, where 1 is the lowest score and 6 the highest[8];
- self-reported average grade during secondary education, also measured on an ordinal 1–6 scale. Along with the previous variable, this is supposed to proxy ability;
- experience and experience squared: inclusion of these variables is based on the human capital notion of on-the-job training investment (*cf.* Mincer (1974));
- a dummy equal to 1 if the respondent is a public sector employee. We include this variable because we know from other research that wages among economists differ by sector (see Oosterbeek, Hartog and Odink (1990));
- a dummy variable indicating that the respondent specialized in business economics. We include this variable out of curiosity. Teachers of business courses often claim that their courses have higher returns than the courses taught in the general economics curriculum. In our regressions we examine this claim.

Variables included in vector X, which are believed to influence the choice of the university into which respondents select themselves, are:

7 There are now six economics departments. The sixth is the economics department of the recently founded University of Limburg in Maastricht. At the time of our questionnaire, however, no graduates from this department had yet entered the labour market.

8 A scale from 1 to 10 is commonly used to grade Dutch students. Our six-point scale relates to this as follows: $1 = 6$; $2 = 6.5$; $3 = 7$; $4 = 7.5$; $5 = 8$; $6 \geq 8.5$. Grades below 6 are not considered since it is impossible to graduate with such a low average score.

- average grade at secondary school. This allows ability to affect university choice;
- gender (male = 1);
- educational levels of the father and mother, both measured on an 8-point ordinal scale (1 lowest, 8 highest). We include these variables to capture the possibility that social background may affect university choice.

Besides the individual-specific variables listed above, the choice of a particular economics department is likely to be influenced by the different characteristics of the various departments available. In the structural choice equation (1), we included one such state-specific variable, namely the net present value of lifetime earnings. But other characteristics of the departments – such as distance from home, attendance by friends and academic reputation – might be equally important. The questionnaire included an item on this issue, but the framing of the question was inappropriate for straightforward application in a logit analysis. Respondents were asked whether 'your decision to study at the preferred department was particularly determined by[9]:

- distance from home (46.5%),
- friends being enrolled there (12%),
- the right religious denomination (11.2%),
- the attractiveness of the city (15.9%), and
- the good academic reputation of the department (36.7%).'

The phrasing of the question means that we only know how the preferred option was evaluated with regard to these motives, whereas a conditional logit analysis also requires information about the individual's evaluation of the unchosen alternatives. However, rather than disregard the information included in the response to this question, which is definitely relevant to the subject of this paper, we have chosen to reinterpret the question. Where a respondent answered that a particular motive determined his/her choice, we interpret this as meaning that the respondent attached great importance to this motive. This changes a variable which is, in theory, state-specific into an individual-specific variable about preferences. The variables 'distance,' 'friends,' 'religion,' 'city' and 'reputation' are measured in dummy form.[10]

Table 1 contains information with respect to the mean values and standard deviations of all the variables used, separately for each university. Some results in Table 1 stand out. To put it bluntly, UvA and RUG are chosen for the attractiveness of the associated cities, VU and KUB for their religious connotations

9 Figures in parentheses show the percentage of the sample that indicated that the motive was of importance. Since more (or less) than one motive might be mentioned, the percentages do not total one hundred.

10 The absence of choice-specific variables also makes it impossible to apply the logit model to predict the market share of the new department in Maastricht.

TABLE 1 — DESCRIPTIVE STATISTICS PER UNIVERSITY; MEANS AND STANDARD DEVIATIONS*

variable	UvA	VU	RUG	EUR	KUB
log wage rate	3.70	3.55	3.35	3.67	3.66
	(.53)	(.55)	(1.04)	(.98)	(.70)
male (%)	94.0	96.9	90.3	96.8	98.8
education father (1-8)	4.81	4.03	4.45	3.82	3.60
	(2.50)	(2.40)	(2.44)	(2.39)	(2.25)
education mother (1-8)	3.52	2.83	3.26	2.75	2.72
	(1.98)	(1.70)	(2.06)	(1.83)	(1.70)
grade secondary education (1-6)	3.55	3.52	3.40	3.54	3.44
	(1.14)	(1.04)	(1.09)	(1.13)	(1.14)
grade university (1-6)	3.54	3.41	3.57	3.46	3.36
	(1.05)	(1.01)	(1.04)	(1.02)	(.85)
public sector (%)	35.0	34.0	38.6	34.2	35.8
experience (years)	15.0	11.5	9.80	15.3	17.2
	(10.8)	(8.6)	(7.9)	(10.0)	(9.8)
business economics (%)	31.4	56.2	54.0	51.9	45.0
motive important:					
distance (%)	59.2	45.1	48.9	39.6	51.3
friends (%)	12.6	14.8	17.6	7.4	16.7
religion (%)	7.2	34.6	0.0	2.4	29.2
city (%)	40.4	21.0	48.9	1.9	0.8
reputation (%)	17.0	19.8	19.3	64.3	25.4
# of cases	223	162	176	462	240

* standard deviations in parentheses

and EUR for its good academic reputation.[11] Note also that UvA attracts the 'best' students in terms of grades in secondary and higher education and parents' level of education. In the next section we investigate whether this broad picture stands up to a multivariate analysis.

4 EMPIRICAL FINDINGS

In this section we present the estimation results. Table 2 contains the results of the reduced form multinomial logit model. For no other reason than our own affiliation, we choose UvA as the reference category. Whether a variable contributes significantly to the choice of another university over UvA is indicated by the *t*-values. In a multinomial logit model, the sign of a coefficient,

11 VU was founded by the leaders of the Protestant movement, and the K in KUB stands for *Katholiek*, which is Dutch for Catholic.

however, does not indicate whether the effect of the associated variable is either positive or negative. The reason is that the probabilities might be non-monotonic (*cf.* Cramer (1991), p. 46). For that reason we also present the mean sample probability derivatives.

We will first discuss the most notable results of Table 2. Those wishing to specialize in business economics are less likely to enrol at UvA. This result is probably to some extent artificial. The standard curriculum at UvA is called 'economics,' but within this curriculum it is possible to specialize in business or general economics by choosing the appropriate mix of options (instead of the

TABLE 2 — MULTINOMIAL LOGIT ESTIMATION RESULTS, REDUCED FORM

	VU/UvA	RUG/UvA	EUR/UvA	KUB/UvA
constant	−0.46 (0.5)	0.17 (0.2)	0.77 (1.0)	0.00 (0.0)
	[−0.072]	[0.003]	[0.124]	[−0.033]
male = 1	0.49 (0.9)	−0.55 (1.4)	0.87 (1.8)	1.46 (2.1)
	[−0.014]	[−0.109]	[0.043]	[0.135]
grade secondary education	0.07 (0.6)	−0.08 (0.8)	0.05 (0.6)	−0.03 (0.3)
	[0.007]	[−0.010]	[0.010]	[−0.007]
grade university	−0.07 (0.6)	0.12 (1.1)	−0.09 (0.9)	−0.13 (1.2)
	[−0.002]	[0.018]	[−0.008]	[−0.011]
education mother	−0.11 (1.5)	−0.07 (1.1)	−0.11 (1.7)	−0.03 (0.4)
	[−0.006]	[−0.001]	[−0.010]	[0.008]
education father	−0.05 (0.8)	−0.05 (1.0)	−0.12 (2.4)	−0.13 (2.4)
	[0.004]	[0.001]	[−0.008]	[−0.008]
public sector = 1	0.20 (0.8)	0.45 (1.9)	0.50 (2.3)	0.23 (1.0)
	[−0.009]	[0.023]	[0.043]	[−0.013]
business = 1	−1.15 (4.8)	1.08 (4.7)	1.02 (4.7)	0.63 (2.8)
	[0.048]	[0.049]	[0.052]	[−0.031]
motive important				
distance = 1	−0.77 (3.2)	−0.45 (1.9)	−1.12 (5.0)	−0.70 (3.0)
	[−0.015]	[0.010]	[−0.092]	[0.006]
friends = 1	0.15 (0.5)	0.24 (0.8)	−0.58 (1.8)	0.34 (1.1)
	[0.017]	[0.032]	[−0.118]	[0.072]
religion = 1	1.77 (5.4)	--**--	−1.33 (3.0)	1.45 (4.3)
	[1.528]		[1.900]	[1.459]
city = 1	−1.02 (3.7)	0.19 (0.8)	−4.08 (10.1)	−4.56 (6.2)
	[0.159]	[0.218]	[−0.314]	[−0.330]
reputation = 1	0.45 (1.6)	0.10 (0.4)	2.35 (9.4)	0.98 (3.6)
	[−0.055]	[−0.078]	[0.278]	−0.028]
# of cases	1263			
log likelihood	−1478.5			

- absolute value of asymptotic *t*-values in parentheses
- mean sample probability derivatives in square brackets
- ** could not be estimated due to singularity; see Table 1

prescribed mix in each specialization). In the questionnaire, individuals could still respond 'business economics' if their curriculum was dominated by such courses. But the answers might not be equivalent to answers for the other universities. Men are more likely to attend KUB. Quality of student input in terms of average grades in secondary education does not differ among the departments. There seems, however, to be some segmentation in terms of the social background of the students: those who graduated from EUR or KUB tend to have less well-educated fathers than graduates from UvA, VU and RUG. The more pronounced differences can be found for the motive-dummies. Those who think that the distance between university and place of residence is important are less likely to attend VU and EUR. Whether an individual takes into account where friends enrol is a factor which does not influence the odds ratios. Those motivated by religious denomination are more likely to enrol at VU or KUB. Attractiveness of the city is a factor of considerable importance among UvA, VU and RUG graduates as compared with the others. Finally, students particularly concerned about academic reputation are likely to enrol at EUR.

In order to investigate whether the departments really represent different alternatives or whether some distinctions are irrelevant, we applied a test recently developed by Cramer and Ridder (1991).[12] We pooled each possible pair of departments and applied a likelihood ratio test. With 12 degrees of freedom the critical value at the 1%-level is 26.2. The test statistics for all pooled pairs were well above this critical value (UvA and RUG are most alike with a test statistic of 70.1), implying that the original choice set of five departments contains no irrelevant distinctions. This establishes the heterogeneity of the departments.

In Table 3 we present the results from the estimation of the university-specific wage equations. This table contains some interesting results. The experience-wage profiles of the different departments assume different shapes as can be seen from lines 2 and 3 of the table. The KUB and EUR profiles peak before 30 years of experience, while the UvA and VU profiles occupy an intermediate position with a peak between 30 and 35 years, and the RUG profile reaches its maximum a long time after retirement (57.7 years of experience).

From all universities, graduates working in the public sector earn some 31 to 45 per cent less than those employed in the market sector. The VU graduates are exceptional in experiencing only a modest disadvantage in this respect. This

12 Cramer and Ridder propose a likelihood ratio test, in which the test statistic is $2(logL^u - logL^r)$, where $logL^u$ is the maximum likelihood of the original ('unpooled') model and $logL^r$ is the maximum likelihood of the model in which two or more states are pooled. Pooling of states comes down to giving the pooled states the same coefficients. For $logL^r$, they derive: $logL^r = \sum_i (n_{si} log n_s i) - n_s log n_s + logL^p$, where n_s is the total number of observations in the pooled states, n_{si} is the number of observations in pooled state i, and $logL^p$ is the unconstrained maximum of the loglikelihood of the pooled model.

finding confirms the anecdotal information concerning a network of VU economists in the public sector. This network, known as the Abraham Kuyper Society (after the founder of the VU), is said to hold regular informal meetings during which its members discuss, among other things, the 'transfer' market for economists in the public sector. Our results suggest that the Society must be fairly successful as far as wages are concerned.

Within the groups of graduates from any of the different departments, women generally earn neither more nor less than men. The only exception is that KUB male graduates tend to earn higher wages than KUB female graduates.

Pre-university ability, as measured by the average grade in secondary education, has no impact on post-university wages. University grades only have a significantly positive coefficient in the wage equations for EUR and KUB. This suggests that only at EUR and KUB grades are reliable indicators of ultimate professional ability. Notice in this respect that EUR and KUB are also the departments chosen by students who particularly care about academic reputation.

With respect to the dummy for a specialization in business economics, we note that it is only in the KUB wage equation that this variable has a coefficient significantly differing from zero. Claims about higher returns on business courses are therefore not supported by our results.

Of the correction terms for self-selection only the coefficient for the 'lambda' of UvA graduates is significantly different from zero. This indicates that there is significant selectivity bias in the UvA wage equation. The positive sign of this coefficient implies that an average individual from the sample would earn a lower wage rate had he graduated from the UvA than an actual UvA graduate does.

TABLE 3 — UNIVERSITY-SPECIFIC WAGE EQUATIONS

	UvA	VU	RUG	EUR	KUB
constant	2.94 (14)	3.05 (8.3)	3.20 (6.5)	2.33 (7.2)	2.18 (4.7)
experience	0.05 (5.6)	0.06 (4.5)	0.05 (2.0)	0.09 (6.3	0.08 (5.4)
experience squared/100	−0.09 (3.5)	−0.09 (2.0)	−0.04 (0.5)	−0.16 (4.6)	−0.15 (4.1)
male = 1	−0.01 (0.1)	0.06 (0.3)	−0.33 (1.2)	0.18 (0.8)	0.68 (1.8)
public sector = 1	−0.31 (4.4)	−0.17 (1.9)	−0.41 (2.6)	−0.45 (4.9)	−0.40 (4.6)
grade secondary/10	0.18 (0.6)	0.17 (0.5)	−1.50 (1.9)	−0.10 (0.3)	−0.43 (1.2)
grade university/10	0.13 (0.4)	0.08 (0.2)	−0.37 (0.5)	1.21 (2.8)	1.19 (2.3)
business = 1	0.00 (0.0)	−0.04 (0.4)	0.24 (1.5)	0.15 (1.7)	0.19 (2.2)
λ	0.21 (2.3)	−0.05 (0.6)	0.30 (1.6)	0.07 (0.7)	−0.09 (0.8)
R^2	0.31	0.35	0.19	0.19	0.26
adj. R^2	0.29	0.32	0.15	0.18	0.24
# of cases	223	162	176	462	240

– absolute value of t-values in parentheses.

The explained parts of the variance of the wage equations differ considerably between departments, from 19% for RUG and EUR to 35% for VU. Note that the unexplained parts are almost proportional to the standard deviations of wages reported in Table 1.

Next we investigated whether the wage equations for different departments were really different. To do this, we pooled each possible pair of departments and estimated restricted and unrestricted wage equations. In the restricted wage equations intercepts and slopes are equal, while in the unrestricted equations intercepts and slopes are allowed to differ. Using an F-test for the residual sum of squares, we found that at the 5%-level the wage equations of UvA and RUG, of UvA and KUB, and of VU and RUG are different. At the 10%-level we also have to reject equality of the wage equations of VU and KUB, and of RUG and KUB.

Subsequently, we compared the wage structures for the five universities by calculating the expected wages in each of the five wage regimes. Using the methods often applied in discrimination analysis, we weighted the characteristics of those who graduated at university i with the coefficients of the wage equation of university j. Table 4 summarizes the results. The table reads as follows; the field ij in the table indicates the average log wage rate for individuals with characteristics in the i-th row weighted with the wage regime in the j-th column. For example, the average log wage rate for graduates from VU (i) had they fallen in the KUB-wage regime (j) is 3.70.

From the results in Table 4 we find that, for all endowment combinations, a VU-degree gives the highest returns. This result is caused by the high intercept of the VU-wage equation, the low penalty on public sector employment and an advantageous experience-wage profile. At the other extreme we find the RUG wage regime: for all endowment combinations, the RUG coefficients generate

TABLE 4 — LOGS OF EXPECTED WAGES

	UvA	VU	RUG	EUR	KUB	#
UvA	3.27	3.86	2.74	3.48	3.69	223
	(0.09)	(0.13)	(0.15)	(0.14)	(0.14)	
VU	3.16	3.77	2.60	3.40	3.70	162
	(0.09)	(0.10)	(0.27)	(0.14)	(0.13)	
RUG	3.12	3.68	2.61	3.32	3.59	176
	(0.10)	(0.11)	(0.25)	(0.14)	(0.15)	
EUR	3.38	3.84	2.94	3.58	3.75	462
	(0.10)	(0.13)	(0.35)	(0.12)	(0.15)	
KUB	3.33	3.94	2.84	3.60	3.83	240
	(0.09)	(0.11)	(0.34)	(0.13)	(0.12)	
average	3.29	3.83	2.80	3.51	3.73	1263
	(0.10)	(0.12)	(0.32)	(0.13)	(0.14)	

– standard errors in parentheses.

the lowest returns. We also investigated whether the sample averages of the university-specific wages in the last row of Table 4 differed significantly. The following inequalities hold at the 5%-level of significance[13]:

$$w_{VU} > w_{UvA};\ w_{KUB} > w_{UvA};\ w_{VU} > w_{RUG};\ w_{VU} > w_{EUR};\ w_{EUR} > w_{RUG};\ w_{KUB} > w_{RUG}.$$

After calculating for each individual in the sample the net present values of lifetime earnings for each option, we are able to estimate the conditional logit model and the mixed logit model. In the conditional logit model the explanatory variables are choice-specific. In our model the only choice-specific variable is lifetime earnings. The mixed logit includes both choice-specific and individual-specific explanatory variables. Results are presented in Tables 5 and 6.

From the conditional logit model we can infer that, if it is only lifetime earnings that motivate the choice of university, the revealed choices are certainly in accordance with earnings maximization. Expressed as a quasi-elasticity ($\partial P_i / \partial log N_i$), a one per cent increase in lifetime earnings for option i, increases *ceteris paribus*, the probability that i will be attended by between .067 and .082 per cent. The effect is thus quite moderate. If we also include individual-specific variables (extend the model from an earnings to a utility maximizing framework), the results are not straightforward. If individuals' choices differ according to gender, ability, social background, field of specialization and preferences with respect to the characteristics of an option as well, then choices are not significantly affected by lifetime earnings. The other results in Table 6 confirm the results already reported in our discussion of the reduced form equation.

Besides information about the university actually chosen, the dataset also contains information regarding the university that would now be chosen if the respondent were able to choose again. The question was: 'Where would you study economics now, if you had to start again?' It was, of course, extremely interesting to find out whether there was any relation between the probability of choosing a particular university as predicted by our mixed logit model and the currently stated preferences. To do this, we estimated a mixed logit model

TABLE 5 — CONDITIONAL LOGIT MODEL

$log N_i$	0.48 (3.9)
number of cases	1263
log likelihood	−2024.97

– asymptotic t-value in parentheses.

13 The standard errors of the wage differentials are calculated according to the method proposed by Stewart (1987, p. 146).

with the currently preferred university as the dependent variable, and the predicted probabilities as the choice-specific explanatory variables. Since many respondents are likely to choose their 'own' university, and since the calculations of the predicted probabilities are based on the information about the actual university, we also included the university actually chosen as an explanatory variable. If our model has anything to say about preferences, we expect a positive coefficient for the predicted probabilities. The results in Table 7 reveal that currently stated preferences do indeed correlate positively with predicted probabilities. If, therefore, the currently stated preference implies a switch from the actual choice to another university, this switch is not purely random but implies adaptation in the direction predicted by our model, *i.e.* in the direction of the choice which (*ex ante*) yields the highest predicted utility.

TABLE 6 — MIXED LOGIT ESTIMATION RESULTS, STRUCTURAL FORM

N_i	−0.23 (0.6)	[between −0.023 and −0.034]		
	VU/UvA	RUG/UvA	EUR/UvA	KUB/UvA
constant	0.21 (0.3)	1.47 (2.2)	1.31 (2.0)	−0.02 (0.0)
	[−0.041]	[0.107]	[0.151]	[−0.106]
male = 1	0.51 (0.9)	−0.64 (1.6)	0.87 (1.8)	1.62 (2.2)
	[−0.015]	[−0.122]	[0.036]	[0.157]
score secondary education	−0.03 (0.3)	−0.15 (1.3)	−0.02 (0.3)	−0.11 (1.1)
	[0.004]	[−0.012]	[0.008]	[−0.009]
education mother	−0.10 (1.4)	−0.06 (0.9)	−0.11 (1.6)	−0.02 (0.3)
	[−0.005]	[−0.001]	[−0.010]	[0.007]
education father	−0.05 (1.0)	−0.06 (1.1)	−0.13 (2.5)	−0.14 (2.5)
	[0.004]	[0.001]	[−0.008]	[−0.008]
motive important				
distance = 1	−0.74 (3.1)	−0.43 (1.8)	−1.07 (4.9)	−0.67 (2.9)
	[−0.016]	[0.009]	[−0.089]	[0.007]
friends = 1	0.27 (0.9)	0.37 (1.3)	−0.46 (1.5)	0.44 (1.5)
	[0.020]	[0.038]	[−0.114]	[0.072]
religion = 1	1.74 (5.4)	--**--	−1.31 (3.0)	1.43 (4.3)
	[1.903]		[2.567]	[1.884]
city = 1	−1.00 (3.7)	0.22 (0.9)	−4.04 (10.1)	−4.54 (6.2)
	[0.158]	[0.219]	[−0.314]	[0.332]
reputation = 1	0.39 (1.4)	0.02 (0.1)	2.28 (9.3)	0.93 (3.4)
	[−0.055]	[−0.081]	[0.277]	[−0.027]
# of cases	1263			
log likelihood	−1501.87			

- absolute value of asymptotic *t*-values in parentheses.
- mean sample probability derivatives in square brackets.
- ** could not be estimated due to singularity; see Table 1.

TABLE 7 — LOGIT ESTIMATES FOR CURRENTLY PREFERRED UNIVERSITY

P_i		1.54 (8.01)		
	VU/UvA	RUG/UvA	EUR/UvA	KUB/UvA
constant	−1.57 (5.4)	−2.94 (9.2)	−3.07 (10.1)	−18.3 (13.8)
actual choice	0.87 (5.5)	1.50 (9.7)	1.84 (12.1)	5.07 (16.1)
log likelihood	−995.36			
# of cases	1091			

- absolute value of asymptotic t-values in parentheses.

5 CONCLUSIONS

The aim of this paper was to analyze the interdependent relationships between university choice and earnings. To this end, we estimated a switching regression model. In this model the choice function determines the decision to apply to a particular university and thereby the switch to a particular wage regime. The wage structures of the options determine the lifetime earnings prospects associated with those options, and these prospective earnings influence university choice.

The results reveal some notable differences between the wage structures of graduates from different universities: (i) higher grades are only rewarded for KUB and EUR graduates; (ii) public sector economists earn up to 45% less than their market sector equivalents; (iii) only in the KUB wage equation does a specialization in business economics bring a higher return than a specialization in general economics. Paired comparison of the wage equations indicates that not all wage structures are really different. Equality of the wage structures of UvA and RUG, UvA and KUB, VU and RUG, VU and KUB, and RUG and KUB, could all be rejected.

Comparison of the university-specific wage rates reveals that the VU wage regime generates the highest returns, and the RUG regime the lowest. Comparing the calculated wage differentials and their associated standard errors reveals that the following inequalities are significant at the 5%-level: $w_{VU} > w_{UvA}$; $w_{KUB} > w_{UvA}$; $w_{VU} > w_{RUG}$; $w_{VU} > w_{EUR}$; $w_{EUR} > w_{RUG}$; $w_{KUB} > w_{RUG}$. If we take labour market performance as the ultimate measure of quality, the results indicate that VU offers the highest quality and RUG the poorest.

In the choice function, we find that graduates from different universities differ with respect to social background (measured by the level of education of the parents), gender and the motives that they consider important in choosing a department. A test on the possibility of pooling alternatives establishes heterogeneity of departments. In the structural choice function, the choice-specific lifetime earnings have a wrongly signed, but insignificant coefficient. This result indicates that earnings prospects are not a particularly important factor in the choice of a specific university. It is therefore doubtful whether it

would be worthwhile for a university to increase tuition fees in order to improve the quality of its courses.

Finally, we compared the predicted probabilities resulting from our structural choice model with the currently stated preferences. The results indicate that currently stated preferences adapt in the direction predicted by our choice model.

REFERENCES

Commissie Financieringsstructuur Onderzoek en Ontwikkeling, 1985, *Rekenschap van creativiteit*, The Hague.

Cramer, J.S., 1991, *The LOGIT Model: An Introduction for Economists*, London.

Cramer, J.S. and Geert Ridder, 1991, 'Pooling States in the Multinominal Logit Model,' *Journal of Econometrics*, 47, pp. 267–272.

Freeman, Richard, 1986, 'Demand for Education,' in: Orley Ashenfelter and Richard Layard (eds.), *Handbook of Labor Economics*, Volume 1, Amsterdam.

Gaag, Jacques van der and Wim Vijverberg, 1988, 'A Switching Regression Model for Wage Determinants in the Public and Private Sectors of a Developing Country,' *Review of Economics and Statistics*, 70, pp. 244–252.

Garen, John, 1984, 'The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable,' *Econometrica*, 52, pp. 1199–1218.

Hartog, Joop, Gerard Pfann and Geert Ridder, 1989, '(Non-)graduation and the Earnings Function: An Inquiry on Self-selection,' *European Economic Review*, 33, pp. 1373–1395.

Kenny, Lawrence, Lung-Fei Lee, G.S. Maddala and Robert Trost, 1979, 'Return to College Education: An Investigation of Self-selection Bias Based on the Project Talent Data,' *International Economic Review*, 20, pp. 775–789.

Kodde, David and M.A.M. Theunissen, 1985, *De schoolloopbaan van snelle vwo-leerlingen, sexe-ongelijkheid en de keuze voor hoger onderwijs*, Ministerie van Onderwijs en Wetenschappen, The Hague.

Lee, Lung-Fei, 1978, 'Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables,' *International Economic Review*, 19, pp. 415–433.

Lee, Lung-Fei, 1983, 'Generalized Econometric Models with Selectivity,' *Econometrica*, 51, pp. 507–512.

Manski, Charles and David Wise, 1983, *College Choice in America*, Cambridge Mass.

McFadden, Daniel, 1974, 'Conditional Logit Analysis of Qualitative Choice Behavior,' in: P. Zarembka (ed.), *Frontiers in Econometrics*, New York.

Mincer, Jacob, 1974, *Schooling, Experience, and Earning*, NBER, New York.

Micklewright, John, 1989, 'Choice at Sixteen,' *Economica*, 56, pp. 25–39.

Oosterbeek, Hessel, 1990, 'Education and Earnings in the Netherlands; An Empirical Analysis,' *European Economic Review*, 34, pp. 1353–1375.

Oosterbeek, Hessel, Joop Hartog and Joop Odink, 1990, 'Wage Differentials between the Public and Private Sectors: An Application to Dutch Economists,' Mimeo, University of Amsterdam, Faculty of Economics and Econometrics.

Rosen, Sherwin, 1977, 'Human Capital: Relations between Education and Earnings,' in: M. Intriligator (ed.), *Frontiers in Quantitative Economics*, Amsterdam.

Stewart, Mark, 1987, 'Collective Bargaining Arrangements, Closed Shops and Relative Pay,' *Economic Journal*, 97, pp. 140–156.

Trost, Robert and Lung-Fei Lee, 1984, 'Technical Training and Earnings: A Polychotomous Choice Model with Selectivity,' *Review of Economics and Statistics*, 66, pp. 151–156.

Willis, Robert, 1986, 'Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Function,' in: Orley Ashenfelter and Richard Layard (eds.), *Handbook of Labor Economics*, Amsterdam.

Willis, Robert and Sherwin Rosen, 1979, 'Education and Self-selection,' *Journal of Political Economy*, 87, pp. S7–S36.

Summary

AN EMPIRICAL ANALYSIS OF UNIVERSITY CHOICE AND EARNINGS

In this paper we analyze the relations between university choice and earnings. We estimate a model in which a choice function determines the decision to apply to a particular university and thereby the switch to a particular wage regime. The wage structures of the options in turn determine the lifetime earnings prospects associated with those options, and these prospective earnings influence university choice. The results reveal some notable differences between the wage structures of graduates from different universities. In the choice function we find that graduates from different universities differ with respect to social background, gender and the motives that they consider important in choosing a department. A test on the possibility of pooling alternatives establishes heterogeneity of departments. Finally, we find that earnings prospects are not a particularly important factor in the choice of a specific university.