THE RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

Working Paper No. 539, 2000

Occupational Choice and Incentives: The Role of Family Background

by Anna Sjögren

IUI, The Research Institute of Industrial Economics
P.O. Box 5501
SE-114 85 Stockholm
Sweden

Occupational Choice and Incentives: the role of family background

Anna Sjögren*
annas@iui.se
IUI, The Research Institute of Industrial Economics,
Box 5505, 114 85 Stockholm, Sweden.

27 September 2000

Abstract

A model of occupational choice and human capital investment is developed and tested. The model allows family background to influence occupational choice via access to economic resources, differences in costs of schooling, and ability uncertainty. The model predicts that people are more sensitive to economic incentives when considering occupations that are different from the parental occupation. It also predicts that the occupational choice of individuals from poor background is more sensitive to economic incentives than the occupational choice of well off individuals. These implications are confirmed on Swedish data using a mixed multinomial logit framework, explicitly accounting for unobserved heterogeneity.

Key words: economic incentives, family background, intergenerational mobility, mixed multinomial logit, occupational choice

JEL classification: C25, J24, J62

*I am grateful to John Dagsvik, Tore Ellingsen, Casey Mulligan, Mårten Palme, Kenneth Train for valuable comments. Sune Karlsson's help with computer capacity saved me weeks. Financial support from the Stockholm School of Economics and the Bank of Sweden Tercentenary Foundation is ackowledged.

1. Introduction

Positive correlation between the earnings of parents and offspring has emerged as one of the stylized facts in empirical labor economics. Recent research, surveyed in Solon (1998), has estimated this correlation to ranging between 0.2 and 0.5.[1] The traditional human capital approach to intergenerational transmission of earnings capacity and economic status formulated in Becker and Tomes (1979, 1986), offers at least two explanations for this observed intergenerational earnings correlations. The first, and most emphasized, is access to capital for human capital investments. A second explanation is that innate abilities or capacity to earn income (apart from the human capital one can accumulate through investments) are influenced by parents. This earning capacity is partly genetically transmitted, and partly influenced through upbringing, both purposely and as a form of externality from the parent's human capital and abilities.[2]

Another empirical regularity, emphasized in the Sociological literature on intergenerational mobility, is that people tend to choose an occupation not too different from that of their parents. This occupational choice pattern is illustrated in Table 1, which presents a transition matrix of occupations of father and offspring compiled from data from the Swedish Level of Living Survey (1991). The transition matrix clearly rejects the hypothesis that father's and off-spring's occupations are independent.[3] This tendency to choose occupations not too different from that of the parents does not find its explanation in traditional human capital models of mobility since these do not analyze heterogeneous human capital and hence disregard occupations.

This paper investigates how family background may influence earnings capacity by suggesting that there is an information externality from parental occupation which makes individuals better at assessing their earnings prospects in occupations that are similar to the parental occupation. This mechanism induces family background to influence the importance people place on economic incentives when making educational and occupational choices.[4]

---

[1] See e.g. Björklund and Jäntti (1997) and Mulligan (1997).

[2] The finding that various measures of socioeconomic background enter significantly into wage equations, when schooling has been controlled for, can be interpreted as support for this second explanation for transmission of earnings capacity, See Haveman and Wolfe (1995) for a survey of studies.

[3] See section 4 for a description of the data set as well as the selections made for this study.

[4] This information externality is first modelled in Sjögren (1998)

Table 1: **Transition matrix: father's occupation (down) vs individual's occupation (across)**

	prof	soc	adm	sale	agr	trans	prod	serv	N of f's
prof	**154**	**87**	**101**	**46**	10	19	73	33	523
	(87.4)	(81.8)	(94.3)	(38.6)	(17)	(36.9)	(113.1)	(54)	
soc	**10**	**11**	**12**	**7**	0	1	7	**7**	55
	(9.2)	(8.6)	(9.9)	(4.1)	(1.8)	(3.9)	(11.9)	(5.7)	
adm	**73**	**50**	**78**	**30**	2	15	42	23	313
	(52.3)	(48.9)	(56.4)	(23.1)	(10.2)	(22.1)	(67.7)	(32.3)	
sale	**49**	35	**70**	**46**	5	18	40	24	287
	(47.9)	(44.9)	(51.8)	(21.2)	(9.3)	(20.2)	(62.0)	(29.6)	
agr	107	128	115	43	**86**	56	**189**	**104**	828
	(138.3)	(129.5)	(149.3)	(61.1)	(26.9)	(58.4)	(179)	(85.5)	
trans	52	**60**	66	23	8	**52**	66	**41**	368
	(61.5)	(57.5)	(66.4)	(27.2)	(12)	(25.9)	(79.6)	(38)	
prod	222	227	255	98	23	**120**	**434**	**160**	1539
	(257.1)	(240.7)	(277.6)	(113.6)	(50)	(108.5)	(332.7)	(158.9)	
serv	28	**41**	**50**	12	3	6	36	**29**	205
	(34.2)	(32.1)	(37)	(15.1)	(6.7)	(14.4)	(44.3)	(21.2)	
other	9	**20**	13	6	0	**10**	**24**	**14**	96
	(16)	(15)	(17.3)	(7.1)	(3.1)	(6.8)	(20.8)	(9.9)	
N of ind's	704	659	760	311	137	297	911	435	4214

prof=professional, soc=social/medical,adm=administrative, sale=sales, trans= transport and communication, prod=production, serv=services, f's=fathers. Figures in brackets indicate expected number in the category, given row-column independence. Bold figures indicate over representation in the category. χ^2-test of row-column independence: χ^2= 440.4. Independence rejected at P= 0.00.

I develop a model of educational investments and occupational choice, inspired by Roy (1951). The model allows family background to influence educational and occupational choices in three ways: The first two mechanisms, differences in economic resources available for human capital investments, and differences in costs of schooling due to imperfect capital markets or information cost are well known from traditional human capital models. The third mechanism, introduced here, works via family background related uncertainty about individual ability to work and succeed in different occupations. I assume that individuals do not have perfect information about their ability for different occupations, and that ability uncertainty is greater when considering occupations distant from the parents' occupation. Hence, risk averse individuals are more reluctant to choose distant occupations than familiar ones. The model generates a number of implications for how family background influences the valuation of economic incentives under different assumptions regarding risk aversion.

The implications of the model are tested on data from the Swedish Level of Living Survey (LNU 1991). In particular, I test if family background influences the way in which occupational choices of people are determined by economic incentives, such as earnings prospects, earnings dispersion, return to education and return to experience. The empirical approach is inspired by previous empirical studies of occupational choice, e.g. Boskin (1974), Robertson and Symons (1990), Orazem and Mattila (1991), and Flyer (1997). In contrast to these studies, in which individual ability heterogeneity is either completely ignored or controlled for using rough measures, this study uses econometric techniques, i.e. MMNL, to account for unobserved heterogeneity. This estimation method, discussed thoroughly in McFadden and Train (1997), takes into account that there are observable factors that influence the individual's occupational choice by introducing a random element in the estimated coefficients.

In general, the empirical findings support the implications that can be drawn from the model. I find that people are more sensitive to economic incentives and reluctant to take chances when considering unfamiliar occupations, and poor background increases sensitivity to economic incentives. In particular, the results indicate that return to education is of greater importance for people from less well off background, while people with educated parents tend to go for occupations requiring long educations, caring less about the actual returns to education. Furthermore, while less well off background implies greater disliking of wage dispersion, wage dispersion is actually regarded as positive by individuals with university educated parents.

The paper is organized as follows. In Ssection 2, I develop a model of occupational choice and derive some testable implications. Section 3 presents an empirical specification. I describe the data and construct incentive variables using wage regressions in Section 4 . Section 5 presents results of MMNL estimations of family background determined differences in attitudes toward incentives. Section 6 concludes.

2. A Human Capital Approach to Occupational Choice, Incentives and Family Background

Each individual lives in two periods, as students and as workers. Students live off money received from their parents net of what ever they have to pay for their education.[5] Workers live off their own earnings. To keep the model simple, loans and savings are disregarded and the possibility of receiving bequests or inheritance is ignored. Hence, all individuals in the model are more or less capital constrained.[6] The individual chooses the amount of education and the occupation that give the highest life time utility. Utility of an individual who chooses occupation j depends on the level of consumption while a student, $c_{1j,}$ and on consumption while working in occupation j, c_{2j}. This second period consumption is subject to uncertainty because the individual does not know for sure how well he will succeed in the chosen occupation. Life time utility, given that occupation j is chosen:

$$U(c_{1j}, c_{2j}) = u(c_{1j}) + \gamma E\left[u(c_{2j})\right], \quad \gamma > 0,\ u' > 0,\ u'' < 0. \tag{2.1}$$

The individual influences his level of consumption in both periods through the choice of occupation and through the amount of education he invests as a student. The budget constraint while a student is:

$$Y_1 - k_j H_j = c_{1j}, \tag{2.2}$$

where Y_1 is the money the individual receives from his parents. This amount may differ across individuals according to the income and generosity of the parents. H_j is the human capital investment associated with choosing occupation j, k_j is the per unit cost of human capital investment in occupation j. This cost is likely

[5]The structure of this model is inspired by the intergenerational model of Mulligan (1997,1999) and by Willis' (1986) version of the occupational choice model described in Roy (1951). A slightly different version of the model is analyzed in Sjögren (1998).

[6]Mulligan (1997) thoroughly analyses the effects of capital constraints.

to be higher for people from poor or uneducated background. It may also depend on the occupation under consideration since some types of education require more effort than others.

The budget constraint in period two is:

$$Y_{2j} = c_{2j}, \tag{2.3}$$

where Y_{2j} is the earnings of the individual in occupation j. These earnings depend on the amount of human capital, H_j, the individual has invested in, the individual's endowment of occupation j specific ability A_j, the wage rate, W_j, and on the occupation specific parameter β_j determining the ability sensitivity of earnings in occupation j:

$$Y_{2j}(A_j) = W_j H_j A_j^{\beta_j}. \tag{2.4}$$

The individual has a prior belief about his ability for each occupation such that

$$\ln A_j \sim N(\overline{a}_{j,}\rho_j). \tag{2.5}$$

I also assume that the individual is better at assessing his ability in occupations similar to that of his parents than at assessing how successfully he would make it in unfamiliar occupation. Hence, the variance, ρ_j, of the prior distribution is assumed to be larger the more distant from the parental occupation is occupation j. I further assume that in the population at large, occupation specific ability is log normally distributed such that

$$\ln A_j \sim N(0,1) \text{ for all } j. \tag{2.6}$$

In line with Roy (1951), this implies that if individuals were randomly assigned to education level and occupation, the log of earnings are also normally distributed, $\ln Y_{2j} \sim N(\ln\left(W_j\overline{H_j}\right), \beta_j)$, where $\overline{H_j}$ is the average education of those in the occupation. With random assignment there would be a tight link between the earnings dispersion within an occupation and the sensitivity of earnings to ability.

However, individuals *choose* the occupation which give them the highest life time utility. That is, an individual chooses occupation i if expected utility in occupation i is higher than expected utility in occupation j, i.e.:

$$E\left[U_i(H_i^*)\right] - E\left[U_j(H_j^*)\right] \geq 0, \text{ for all } j \neq i, \tag{2.7}$$

where H_i^* and H_j^* are the optimal human capital investments associated with each occupational choice.

The optimal human capital investment is obtained by maximizing with respect to H the expression for lifetime utility, 2.1, subject to the budget constraints 2.2 and 2.3. Assume a constant relative risk aversion utility function:

$$u(c) = \frac{1}{1-\frac{1}{\sigma}} c^{1-\frac{1}{\sigma}}, \tag{2.8}$$

where the coefficient for relative risk aversion, $\frac{1}{\sigma} > 0$.[7]

The optimal human capital investment, given that the individual chooses occupation j, is thus determined by maximizing the following expression for lifetime utility with respect to H_j.

$$E\left[U_j\right] = \frac{1}{1-\frac{1}{\sigma}} \left(Y_1 - k_j H_j\right)^{1-\frac{1}{\sigma}} + \frac{\gamma}{1-\frac{1}{\sigma}} \int_{-\infty}^{\infty} \left(W_j H_j A_j^{\beta_j}\right)^{1-\frac{1}{\sigma}} f(a_j) da_j, \tag{2.9}$$

where $\left(\frac{1}{\gamma} - 1\right) \geq 0$, is the rate of time preference.[8]

The first order condition is:

$$k_j u'(c_{1j}) = \gamma \int_{-\infty}^{\infty} u'(c_{2j}) W_j A_j^{\beta_j} f(a_j) da_j, \tag{2.10}$$

which states that the value of foregone consumption in period one must on the margin equal the value of the gain in consumption in period two. This first order condition can be manipulated to obtain an expression for the optimal human capital investment given that the individual chooses occupation j:[9]

$$H_j^* = \frac{Y_1}{k_j + \left(\frac{k_j}{\gamma}\right)^{\sigma} \left(W_j \overline{A}_j^{\beta_j} \xi_j\right)^{1-\sigma}}, \tag{2.11}$$

where

$$\xi_j = e^{\frac{\beta_j^2}{2}\rho_j\left(1-\frac{1}{\sigma}\right)} \tag{2.12}$$

is an uncertainty factor. The uncertainty factor, $\xi > 1$ if risk aversion is moderate ($\frac{1}{\sigma} < 1$). The uncertainty factor, $\xi < 1$ if risk aversion is strong ($\frac{1}{\sigma} > 1$).

[7] The higher $\frac{1}{\sigma}$ the more risk averse the individual.

[8] Recall that $\ln A_j = a_j$ and that $a_j \sim N(0,1)$

[9] See appendix A.1 for details.

Comparative statics on this optimal H, presented in appendix A.2, show that the optimal human capital investment increases with Y_1 and decreases with the cost, k. However, an increase in W, induces higher human capital investment only if the individual is moderately risk averse. In the case of strong risk aversion, the high wage is a substitute for human capital investment. Higher uncertainty about ability, ρ, induces higher human capital investment. Increased sensitivity to ability, β, within an occupation causes the strongly risk averse to invest in more human capital unless expected ability, and hence expected earnings, are very high. Unless they expect to be very able, the strongly risk averse have an incentive to insure themselves by investing in more education. When risk aversion is moderate, human capital investments increase as a result o increased ability sensitivity unless expected ability is very low because of the complementarity between ability and education.

2.1. Expected life time utility in an occupation

Now, derive the following expression for expected life time (indirect) utility given the optimal human capital investment and given that the individual chooses occupation j:

$$E\left[U_j^*\right] = \frac{1}{\left(1-\frac{1}{\sigma}\right)}\left(\left(Y_1 - k_j H_j^*\right)^{1-\frac{1}{\sigma}} + \gamma\left(W_j H_j^* \overline{A}_j^{\beta_j}\xi_j\right)^{1-\frac{1}{\sigma}}\right). \quad (2.13)$$

Substitute for the optimal human capital investment from 2.11 and rearrrange to get:

$$E\left[U_j^*\right] = \left(1-\frac{1}{\sigma}\right) Y_1^{1-\frac{1}{\sigma}}\left\{1+\gamma^{\sigma} k_j^{1-\sigma}\left(W_j \overline{A}_j^{\beta_j}\xi_j\right)^{\sigma-1}\right\}^{\frac{1}{\sigma}}. \quad (2.14)$$

The envelope theorem, permits the derivation of comparative statics on the expression for expected utility disregarding the effect on $E\left[U_j^*\right]$ of changes in H^*. These comparative statics, presented in appendix A.3, show that an increase in the amount of resources received from the parents, Y_1, increases utility whereas an increase in the cost, k_j, of human capital investments reduces utility. Increasing the wage, W_j, improves the utility associated with choosing an occupation. The effect of an increase in the uncertainty about ability, ρ_j, depends on the degree of risk aversion. Utility decreases when risk aversion is strong and increases when risk aversion is moderate. An increase in the sensitivity of earnings to ability, β_j, increases expected utility if ability is high enough and reduces expected utility if ability is low. However, the cut off ability, at which expected utility remains unchanged, depends on the degree of risk aversion. If risk aversion is strong, higher β_j, will increase utility only if ability is very high. With moderate risk aversion, utility will increase unless ability is very low.

The expected utility of individual i associated with choosing occupation j as a function of incentives and personal characteristics can be summarized as follows:

$$E\left[U_{ij}^*\right] = F(\overset{+}{Y_{1i}}, \overset{-}{k_{ij}}, \overset{+}{W_j}, \overset{+-}{\beta_j}, \overset{+}{\overline{A}_{ij}}, \overset{+-}{\rho_{ij}}), \tag{2.15}$$

where the signs above the arguments of the function indicate how the marginal valuation of the occupation changes with marginal changes in the arguments.

2.2. Family background and the sensitivity to economic incentives

In this section I use comparative statics results summarized in expression 2.15 to derive predictions about differences in sensitivity to economic incentives due to family background. The effect of family background on the marginal valuation of economic incentives is defined as the derivative of the marginal utility of the incentive variable in question with respect to family background. That is: $\partial\frac{\partial E[U]}{\partial I}/\partial F$, where I is the economic incentive variable and F is the family background variable. I further define the sensitivity to an incentive as the magnitude of the absolute value of the marginal valuation. Thus, a person is said to become more sensitive to the incentive variable if $\partial\frac{\partial E[U]}{\partial I}/\partial F$ has the same sign as $\frac{\partial E[U]}{\partial I}$ and less sensitive if they have the opposite signs.

The incentive variables considered are the wage, W_j, and the sensitivity of earnings to ability, β_j. As family background variables I will consider, Y_{1i}, the economic support received from parents, k_{ij}, the cost of educational investments, and ρ_{ij}, which captures that the distance in terms of ability uncertainty of the occupation in question to the occupation of the parents.

2.2.1. The sensitivity to the wage rate

$$\begin{aligned} \frac{\partial\frac{\partial E[U]}{\partial W}}{\partial Y_1} &= \frac{\partial^2 E\left[U\right]}{\partial c_2^2}\frac{\partial c_2}{\partial H^*}\frac{\partial H^*}{\partial Y_1}\frac{\partial c_2}{\partial W} + \frac{\partial E\left[U\right]}{\partial c_2}\frac{\partial^2 c_2}{\partial W \partial H^*}\frac{\partial H^*}{\partial Y_1} \\ &= (1-\frac{1}{\sigma})\gamma\left(W_j H^*\right)^{-\frac{1}{\sigma}}\left(\overline{A}_j^{\beta j}\xi\right)^{(1-\frac{1}{\sigma})}\frac{\partial H^*}{\partial Y_1} \gtrless 0 \; if \; (1-\frac{1}{\sigma}) \gtrless 0 \end{aligned} \tag{2.16}$$

$$\frac{\partial \frac{\partial E[U]}{\partial W}}{\partial k} = \frac{\partial^2 E\left[U\right]}{\partial c_2^2}\frac{\partial c_2}{\partial H^*}\frac{\partial H^*}{\partial k}\frac{\partial c_2}{\partial W} + \frac{\partial E\left[U\right]}{\partial c_2}\frac{\partial^2 c_2}{\partial W \partial H^*}\frac{\partial H^*}{\partial k} \tag{2.17}$$

$$= (1-\frac{1}{\sigma})\gamma\left(W_j H^*\right)^{-\frac{1}{\sigma}}\left(\overline{A}_j^{\beta j}\xi\right)^{(1-\frac{1}{\sigma})}\frac{\partial H^*}{\partial k} \gtrless 0 \; if \; (1-\frac{1}{\sigma}) \lessgtr 0 \tag{2.18}$$

The expression in 2.16 shows that strongly risk averse individuals become less sensitive to the wage rate with increases in the income received from the parents. Furthermore, 2.17 shows that as the cost of education increases, the sensitivity to the wage rate is increased for the strongly risk averse. The reasons for these effects are that marginal utility of consumption decreases rapidly when risk aversion is strong and that the increase in consumption that is made possible by a higher wage is valued less when the level of consumption is high at the outset. When the individual is moderately risk averse, however, utility becomes more sensitive to the wage rate as the income received from the parents increases and less sensitive if the cost of education increases.

$$\frac{\partial \frac{\partial E[U]}{\partial W}}{\partial \rho} = \left(1-\frac{1}{\sigma}\right)\gamma\left(W_j H^*\right)^{-\frac{1}{\sigma}}\left(\overline{A}_j^{\beta j}\xi\right)^{(1-\frac{1}{\sigma})}\left(\frac{\partial H^*}{\partial \rho} + \frac{\beta^2}{2}(1-\frac{1}{\sigma})H^*\right) > 0 \tag{2.19}$$

2.19 shows that increased ability uncertainty increases the sensitivity to the wage rate.

2.2.2. The sensitivity to ability sensitivity (wage dispersion)

$$\frac{\partial \frac{\partial E[U]}{\partial \beta}}{\partial Y_1} = (1-\frac{1}{\sigma})\gamma H^{*-\frac{1}{\sigma}}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\frac{1}{\sigma})}\left(a+\beta\rho(1-\frac{1}{\sigma})\right)\frac{\partial H^*}{\partial Y_1} \tag{2.20}$$

$$\gtreqless 0 \quad if \quad \begin{cases} (1-\frac{1}{\sigma}) < 0 \text{ and } \overline{a}_j \lesseqgtr \beta\rho(\frac{1}{\sigma}-1) \\ (1-\frac{1}{\sigma}) > 0 \text{ and } \overline{a}_j \gtreqless \beta\rho(\frac{1}{\sigma}-1) \end{cases}$$

$$\frac{\partial \frac{\partial E[U]}{\partial \beta}}{\partial k} = (1-\frac{1}{\sigma})\gamma H^{*-\frac{1}{\sigma}}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\frac{1}{\sigma})}\left(\overline{a}_j+\beta\rho(1-\frac{1}{\sigma})\right)\frac{\partial H^*}{\partial k} \tag{2.21}$$

$$\gtreqless 0 \quad if \quad \begin{cases} (1-\frac{1}{\sigma}) < 0 \text{ and } \overline{a}_j \gtreqless \beta\rho(\frac{1}{\sigma}-1) \\ (1-\frac{1}{\sigma}) > 0 \text{ and } \overline{a}_j \lesseqgtr \beta\rho(\frac{1}{\sigma}-1) \end{cases}$$

Expression 2.20 implies that strongly risk averse individuals become less sensitive to ability sensitivity as the income received for the parents increases. In other words, the individual is less positive when ability is large enough and less negative when ability is low. 2.21 shows that when the cost of education increases, the sensitivity to ability sensitivity is increased. Moderately risk averse individuals, on the other hand, become more sensitive to ability sensitivity when the income received from the parents increases and less sensitive to ability sensitivity as the cost of education increases.

$$\begin{aligned}\frac{\partial \frac{\partial E[U]}{\partial \beta}}{\partial \rho} &= (1-\frac{1}{\sigma})\gamma\left(W_j H^* \overline{A}_j^{\beta j}\xi\right)^{(1-\frac{1}{\sigma})} \\ &\left[\left(\frac{\partial H^*}{\partial \rho}H^{*-1}+\frac{\beta^2}{2}(1-\frac{1}{\sigma})\right)\left(\overline{a}_j+\beta\rho(1-\frac{1}{\sigma})\right)+\beta\right] \\ &\gtreqless 0 \;\; if \;\; \overline{a}_j \gtreqless \phi\end{aligned} \tag{2.22}$$

2.22 implies that increased ability uncertainty tends to increases the sensitivity to ability sensitivity.[10]

2.3. Summarizing the implications of the model

If risk aversion is strong, individuals from less privileged background are more sensitive to incentives than are individuals from more privileged background. That is, they are more sensitive to the wage rate, and they are more positive towards wage dispersion if their ability is high enough, and more negative if ability is low. If risk aversion is moderate, individuals from less privileged background are less sensitive to incentives than are individuals from well off background.

Regardless of risk aversion, individuals are more wage sensitive and more sensitive to wage dispersion when considering occupations that are unfamiliar than when considering a familiar occupation.

[10] $\phi = \frac{2}{\left(\frac{1}{\sigma}-1\right)}\left(\frac{\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta_j}\xi_j\right)^{1-\sigma}+k}{\sigma\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta_j}\xi_j\right)^{1-\sigma}+k}\right)$
$+\beta\rho(\frac{1}{\sigma}-1)$

3. Empirical Specification and Estimation

3.1. A random utility model

The theoretical model is the point of departure for the specification of a mixed multinomial logit model of occupational choice. Equation 2.14 implies that choosing the occupation which maximizes expected utility is equivalent to maximizing $\left(\frac{W_j \overline{A}_{ij}^{\beta_j} \xi_{ij}}{k_{ij}}\right)$. This is used as a reduced form for occupational choice. Hence, I define individual i's random utility associated with choosing occupation j as:

$$V_{ij} \equiv \ln\left(\frac{W_j \overline{A}_{ij}^{\beta_j} \xi_{ij}}{k_{ij}}\right) + \varepsilon_{ij} = \ln W_j + \beta_j \ln \overline{A}_{ij} + \ln \xi_{ij} - \ln k_{ij} + \varepsilon_{ij}$$
$$= \mathbf{X}_j(\overline{\mathbf{\Phi}} + \mathbf{\Phi}_f + \boldsymbol{\mu}_{ij}) + \varepsilon_{ij}, \ (j = 1, ..., M), \quad (3.1)$$

where V_{ij} is the expected life time utility for an individual i, who has family background f, of choosing occupation j. $\mathbf{X}_j$ is a vector of attributes of occupation j reflecting the returns, risk and costs associated with choosing the occupation. Together, $\overline{\mathbf{\Phi}}$, $\mathbf{\Phi}_f$ and $\boldsymbol{\mu}_{ij}$ capture the individual's marginal valuation of economic incentives, as discussed in section 2.1.

The coefficient vector $\overline{\mathbf{\Phi}}$ captures the population mean of the marginal valuation of the occupational attributes and $\mathbf{\Phi}_f$ captures deviations from this mean marginal valuation that depend on the family background f of individual i. The theoretical model showed that the marginal valuation of incentives may differ across family background also when there are no systematic background differences in preferences *per se*. However, $\mathbf{\Phi}_f$, captures also systematic background related deviations in preferences. The coefficient vector $\boldsymbol{\mu}_{ij}$ captures stochastic individual deviations from this $(\overline{\mathbf{\Phi}} + \mathbf{\Phi}_f)$ that result from unobserved heterogeneity due to e.g. unobserved occupation specific ability or individual preference differences that are not shared for people of the same family background. The distribution of this individual heterogeneity is allowed to differ across family backgrounds. ε_{ij} is an individual and occupation specific random disturbance which is assumed to be i.i.d. extreme value. This ε_{ij} captures that an individual may have a special interest in a particular occupation which has nothing to do with ability or background or interest for other occupations.

3.2. Measures and empirical specification

The $\mathbf{X}_j$ vector should contain variables that reflect $\ln W_j + \beta_j \ln \overline{A}_{ij} + \ln \xi_{ij} - \ln k_{ij}$. That is, it should capture the returns and risks and costs associated with choosing an occupation. Measures of returns and risks are based on the earnings function, (2.4), which has it that $\ln Y_j = \ln W_j + \ln H_j + \beta_j a_j$. A measure of the length of education typically required in the occupation is included to reflect the occupation specific element of the cost of human capital investments, $\ln k_{ij}$.

Since (2.4) represents working life earnings in a two-period model it does not include work experience. This is inadequate in the empirical analysis because the return to experience may capture an element of the return to ability in the occupation, especially if ability is of importance for how much an individual is able to advance and learn on the job. There is of course also a genuine return to experience which needs to be taken into account.

Hence, I assume that the underlying wage structure in occupation j can be described by an ability augmented Mincer type earnings function:

$$\ln y_{ij} = \varphi_{0j} + \varphi_{1j} s_i + \varphi_{2j} x_i + \varphi_{3j} x_i^2 + \varphi_{4j} a_i + u_{ij} \tag{3.2}$$

where y is earnings, s is schooling, x is working experience, a is ability and u_{ij} is a random disturbance which is assumed to be i.i.d. in the population as a whole.[11] φ_{0j} corresponds to $\ln W_j$, φ_{1j} measures returns to schooling, φ_{2j} and φ_{3j} capture the return to experience, and φ_{4j} corresponds to β_j, which is linked to wage dispersion within the occupation.[12]

The return to schooling and experience can be argued to depend on the ability of the individual. From the model it also follows that investment in human capital or schooling is endogenous and depends on ability, background and returns. Estimates of φ_{1j}, φ_{2j}, φ_{3j} and φ_{4j} therefore all include elements of the return to ability. Since ability is uncertain, estimates would also capture elements of risk adjustment factor ξ_{ij}, which is linked to family background through ρ_{ij}.

Because there is no data on occupation specific ability, it is not possible to estimate each individuals' true expected returns associated with each occupation from data. The estimated contents of $\mathbf{X}_j$ are thus occupation specific, but not individual specific. The mixed multinomial logit model, handles this by instead allowing the estimated marginal valuations to vary across family background and

[11] See Willis (1986) for a discussion of earnings functions.

[12] See the argument of Roy (1951), presented in section 2.1.

individuals. Hence, the heterogeneity in these estimated marginal valuations reflect differences in preferences, heterogeneity in valuation and unmeasured heterogeneity in the measured returns, risks and costs. The estimation of $\mathbf{X}_j$ from earnings data is discussed below.

The theoretical model predicts that occupational choice will depend on family background by influencing the marginal valuation of economic incentives. The model also emphasizes different aspects of family background. First, family background influences (in the model it determines) the resources available for educational investments. Second, family background influences the cost of education because of possible imperfections in access to financial markets and because well educated parents may have better information about the educational system and thus face lower information costs. Third, family background will influence how well the individual can assess his ability to make it in different occupations. That is, the more distant from the family occupation, the poorer is the quality of information the individual has about his ability.

3.3. Estimation

The individual chooses occupation j if $V_{ij} \geq V_{in}$, for all n.[13] When ε_{ij} are i.i.d. extreme value, the choice probability conditional on $\boldsymbol{\mu}$, is:

$$L_j(\boldsymbol{\mu}) = \frac{\exp(((\overline{\boldsymbol{\Phi}} + \boldsymbol{\Phi}_f) + \boldsymbol{\mu}_{ij})\mathbf{X}_j)}{\sum_n \exp(((\overline{\boldsymbol{\Phi}} + \boldsymbol{\Phi}_f) + \boldsymbol{\mu}_{in})\mathbf{X}_n)} \tag{3.3}$$

regardless of the distribution of $\boldsymbol{\mu}$.[14] However, since $\boldsymbol{\mu}$ is not known, the unconditional probability of choosing occupation j, P_j, is the integral of 3.3 over all values of $\boldsymbol{\mu}$ weighted by the density of $\boldsymbol{\mu}$:

$$P_j = \int L_j(\boldsymbol{\mu}) f(\boldsymbol{\mu} \mid \Omega) d\boldsymbol{\mu}. \tag{3.4}$$

$f(\boldsymbol{\mu} \mid \Omega)$ is the density function of $\boldsymbol{\mu}$ and Ω are the fixed parameters of this density function. This integral generally has no closed form and hence needs to be approximated through simulation. Thus, for a given set of values for the parameters, Ω, a value of each element $\boldsymbol{\mu}$ is drawn from its distribution. The logit formula, $L_j(\boldsymbol{\mu})$, is calculated using this draw. This process is repeated, and

[13] We ignore the option of not choosing an occupation at all.

[14] See Brownstone and Train (1996).

the average of the resulting $L_j(\boldsymbol{\mu})$'s is the approximate choice probability. The sum over all individuals of the logs of this approximate choice probability is the simulated log-likelihood function and the estimated parameters are those that maximize this simulated log-likelihood function.[15]

McFadden and Train (1997) show that any random utility model can be approximated with a mixed logit through appropriate choice of explanatory variables and distributions for the random parameters. Furthermore, they state that an appropriate specification test for the MMNL model is a likelihood ratio test for omitted variables with the corresponding MNL model as the restricted model.[16]

4. Data and Measurement

The empirical analysis uses data extracted from the Swedish Level of Living survey (LNU).[17] The data set used contains information on age, occupation, years of education, type of education, hourly earnings in 1991, and family background variables such as father's education and occupation for a representative sample of 6773 individuals born between 1915 and 1973.

From this sample 4214 individuals of working age who have a registered occupation are selected. The normal retirement age in Sweden is 65, hence, the sample includes individuals between the ages of 20 and 65. Hourly earnings have been computed using data on total monthly earnings and hours worked. Individuals with hourly earnings exceeding SEK 900 are not included in the sample. The sample is further reduced by the exclusion of the occupational category agriculture. The reason for excluding individuals working in agriculture is that the earnings function estimated in order to generate occupation specific incentive variables fails completely to capture how agricultural earnings are determined. The reason for this failure is likely to be that earnings in agriculture are to a great extent explained by size and location of farm. The empirical analysis is therefore conducted without people working in agriculture. Hence, the empirical analysis is based on a

[15]The gauss code used for our estimations is generously made available on the home page of Kenneth Train, Department of Economics, University of California, Berkeley. Our estimtions are based on 200 repetitions.

[16]In this paper, we will assume that the random parameters are normally distributed. However, there may be situations where we do not want to allow the parameters to take different signs for different people. In such cases, it may be preferable to assume that parameters are log normally distributed.

[17]See Erikson and Åberg (1987) for a detailed description of the Swedish Level of Living Survey.

sample of 4077 observations. Definitions and measurements of the data analyzed are discussed in the rest of this section.

4.1. Occupational categories

Based on NYK85 occupational codes, 8 occupational categories are defined.[18]. These are professional, social, administrative, sales, agriculture, transport and communications, production and services. For a more detailed presentation see Table A.4 in the Appendix.

At this aggregate level, it is obvious that each occupational category spans a vast range of activities. This occupational classification rather than one based more closely on years of education is motivated by the concept of occupation used in Roy (1951). Occupations are occupations because they require different types of, or combinations of, abilities. Hence, the classification employed in this paper assumes that a nurse and a medical doctor use the same type of ability. The difference between them is that they have different amounts of education.

Table 4.1 presents some descriptive statistics for the occupational categories except agriculture, which is excluded from the analysis for reasons already discussed. It is clear from Table 4.1 that there are large gender differences between the different occupations. These differences indicate that the types of jobs held by men and women in each occupational category may differ.

[18]NYK85 = Nordic occupation classification 1985. The 8 occupations correspond to the 1-digit NYK85 categories, except that mining and production workers have been put in the same group.

Table 4.1:
Descriptive figures for 7 broad occupational categories

	male					female				
	obs	av w/h	s.d. w	ed yrs	s.d. ed.yrs	obs	av w/h	s.d. w	ed yrs	s.d. ed.yrs
prof	403	81.9	48.9	13.7	3.82	301	68.6	42.0	13.9	3.01
soc	88	77.8	54.0	14.9	4.1	571	66.1	62.0	11.5	2.8
adm	285	87.8	60.6	13.0	3.17	475	64.8	32.9	11.4	2.62
sales	159	78.2	84.8	11.6	2.89	152	45.5	40.6	10.6	2.92
trans	196	57.3	37.6	10.1	2.57	101	57.7	29.8	10.3	2.24
prod	743	62.4	46.5	10.2	2.38	168	41.8	52.4	9.2	2.65
serv	173	64.3	61.6	10.6	2.83	262	46.8	54.3	9.2	2.6
total	2047	71.3	54.9	11.6		2030	59.7	49.6	11.2	
m+f	4077	65.5	52.6	11.4						

obs=number of observations, av w/h=average hourly wage, s.d.w= standard deviation of wage, ed yrs=average years of education, s.d. ed yrs=standard deviation of average years of education, m=male, f=female.

4.2. Constructing incentive variables

The next step is to obtain measures of the occupational attributes in $\mathbf{X}_j$. The lack of data on occupation specific ability and the endogenous nature of schooling, makes estimation of 3.2 subject both to omitted variable and simultaneity biases. Biased estimates are a serious problem if there is reason to believe that the biases in the estimates are systematic across occupations in a way that systematically distorts a comparison of returns across occupations.[19] Since I have not found arguments for why this should be the case, I proceed with estimating occupation

[19] See Willis (1986) for a discussion of ability and biases in estimation of wage equations.

specific wage regressions:

$$\ln y_{ij} = \varphi_{0j} + \varphi_{1j} s_i + \varphi_{2j} x_i + \varphi_{3j} x_i^2 + \varphi_{4j} u_{ij}. \tag{4.1}$$

The estimate of φ_{0j} is used as a measure of the hourly wage rate in the occupation, φ_{1j} measures the return to schooling , φ_{2j} and φ_{3j} capture the return to experience. Although the return to experience would be better described by both φ_{2j} and φ_{3j}, only φ_{2j} will be entered into the final occupational choice estimations in order to economize on the number of estimated parameters. The variance of the regression residuals, φ_{4j}^2, assuming that u_{ij} is N(0,1), is assumed to capture earnings differences that result from factors other education and experience, e.g. differences in ability and is used as a measure of wage dispersion in the occupation. s_i is the number of years of schooling of individual i and x_i measures individual i's experience. Experience is measured as the individual's age, less seven (school starting age), less years of education. The large gender differences in the types of occupations held within these broadly defined occupational categories that became evident in Table 4.1, motivate estimation of separate parameters for men and women. The estimated gender specific incentive variables are presented in Table 4.2.

Table 4.2:
Incentive variables for 7 broad occupational categories

	male				female			
	W	Wdisp	Edpr	Expr	W	Wdisp	Edpr	Expr
	φ_0	φ_4	φ_1	φ_2	φ_0	φ_4	φ_1	φ_2
prof	3.87	0.073	0.024	0.025	3.73	0.116	0.035	0.009
soc	3.36	0.068	0.058	0.028	3.86	0.078	0.025	0.005
adm	3.63	0.09	0.038	0.041	3.66	0.048	0.036	0.017
sales	3.38	0.152	0.060	0.037	3.35	0.062	0.058	0.020
trans	3.87	0.044	0.024	0.014	3.93	0.029	0.017	0.013
prod	4.14	0.078	0.036	0.012	3.69	0.202	0.045	0.002
serv	3.49	0.084	0.052	0.018	3.75	0.062	0.023	0.013
mean	3.68	0.084	0.042	0.025	3.71	0.085	0.034	0.011
m+f	3.69	0.084	0.038	0.018				

W=wage rate, Wdisp=wage dispersion, Edpr=return to education
Expr=return to experience, m=male, f=female.

4.3. Family background variables

Background effects connected to having rich or poor parents, informational advantages of people with well educated parents, and effects of having well off or poor parents on the cost of educational investments are captured with a set of background dummies based on the fathers education.[20] I define three educational background categories, basic education, intermediate education and university education. Basic education corresponds to having attended only compulsory school.[21]

[20] Using both mother and father would probably add to the picture. However, this would result in twice as many background variables apart from causing a problem with missing observations. Many mothers in the sample lack an occupation, probably because beeing a housewife is not classified as an occupation.

[21] During the 1950's and 1960's compulsory school was extended from seven to nine years.

Intermediate education corresponds to having attended more than compulsory education, but not at university level.

I use occupation dummies to capture background effects related to the distance between occupations. Hence, occupations are classifies in terms of familiar or unfamiliar. An occupation is familiar if it belongs to the same occupational category as the father's occupation and unfamiliar if it belongs to a different occupational category. The occupational categories are based on the pattern found in the transition matrix in Table 1, which identifies three groups of occupations between which intergroup occupational mobility tends to be low. The first occupational background category, which is referred to as mixed, contains transport and communications and services. The second background category, is referred to as white collar and contains professionals, social sector, administration and sales. The third category, referred to as blue collar, contains production workers and agriculture.

5. Empirical Results

The results are obtained from estimations of the following model:

$$\begin{aligned} V_{ij} \;=\; & (\overline{\Phi}_1+\overline{\mu}_{1i})W_j+(\overline{\Phi}_2+\overline{\mu}_{2i})Wdisp_j+(\overline{\Phi}_3+\overline{\mu}_{3i})Edpr_j+ \\ & (\overline{\Phi}_4+\overline{\mu}_{4i})Expr_j+(\overline{\Phi}_5+\overline{\mu}_{5i})Ed_j+ \\ & (\Phi_{1f}+\mu_{1if})D_fW_j+(\Phi_{2f}+\mu_{2if})D_fWdisp_j+(\Phi_{3f}+\mu_{3if})D_fEdpr_j+ \\ & (\Phi_{4f}+\mu_{4if})D_fExpr_j+(\Phi_{5f}+\mu_{5if})D_fEd_j+\varepsilon_{ij}. \end{aligned} \tag{5.1}$$

The incentive variables included are the wage rate, W, i.e. the φ_0-estimate, wage dispersion, $Wdisp$ (φ_4), return to education, $Edpr$ (φ_1), return to experience, $Expr$ (φ_2). I also include the average years of education required in the occupation, Ed, as a proxy for the costs of choosing the occupation. However, the length of education required in the occupation also gives information about the level of earnings in the occupation.

D_f is a set of dummy variables measuring background. The dummies included in the models are presented in Table 5. $\overline{\Phi}$ and $\overline{\mu}$ are the marginal valuations of the background category for which there is no dummy included. Φ_f and μ_f are the deviations from $\overline{\Phi}$ and $\overline{\mu}$ of background category f.

Table 5:
Estimated models of occupational choice as a function of incentives

Model	D_f
1	None
2	Father has intermediate/university education
3	Occupation is unfamiliar

5.1. Model selection

Each model is estimated using two specifications, an unrestricted (MMNL) and a restricted (MNL) to serve as a test of the random coefficient specification. The values of the log-likelihood function for each specification are presented in Table 5.1:

Table 5.1:
The value of the log-likelihood function

D	MNL	MMNL
None	-7501	-7776
Father's education	-7441	-7424
Occupational distance	-7406	-7394

Evaluation of the estimated models on the basis of likelihood ratio tests selects the models with background effects over the models without background effects and MMNL estimation over MNL estimation in all three models with background effects on the 5 per cent significance level.[22] The log-likelihood function values further indicate that out of the models with background effects, occupational distance produces the best estimation results.

5.2. Estimation results

The results from the estimations of occupational choice models with background effects are presented in Tables 5.2.1 and 5.2.2. In Table 5.2.1, I present the results

[22] We can note that the results of the MNL estimation of the model without background effects conform qualitatively with the results presented by Orazem and Mattila (1991). However, it should also be noted that the MNL estimations with or without background effects fail to satisfy tests of the independence of irrelevant alternatives assumption. The reason for failure may well be the presence of unaccounted for unobserved heterogeneity.

of estimations allowing coefficients to depend on father's education. The first third of Table 5.2.1 reports the estimates for people whose fathers have only basic education. The middle section of the table reports how the reactions of people whose fathers have intermediate education *differ* from the reactions of people whose fathers have basic education. The third part of the table reports how the reactions of people whose fathers have university education *differ* from the reactions of those whose fathers have basic education.

The point estimates of the coefficients for those with educated fathers are simply the sums of the basic education coefficient estimates of the marginal valuation of an incentive variable and the difference estimates for the corresponding incentive variable for the categories with educated fathers. Significant difference estimates imply that the marginal valuation of the particular incentive variable of individuals with educated fathers differs significantly from the marginal valuation of those whose fathers have basic education. Estimates significant at the 10 per cent level are shown in bold face letters.

The results of the coefficient estimates are very much in line with the predictions of the theoretical model. As expected, the results show that people value the wage rate positively. The positive valuation of the average years of education indicates that the positive effect of high level of income associated with lengthy education is stronger than the negative aspect of high costs of education. This positive coefficient could also reflect a positive valuation of the non-pecuniary returns associated with occupations requiring long education.

It is interesting to note that the return to education is less important the more educated the father, but that average length of education is regarded as more positive the more educated the father. This indicates that if the father is well educated the child will also want a long education, caring less about the actual return and cost of this education and more about that education can be regarded as a "ticket" to high wage, high status occupation.

The result regarding wage dispersion is also interesting. Wage dispersion is regarded as negative by all, except by the university background people. It supports the idea that people from well off background can afford to take chances, either because their future income is diversified enough or because they have better information about where in the wage distribution they will end up, or in fact because their average ability is high enough to make them positively disposed toward wage dispersion.

Table 5.2.1:

Occupational choice with educational background effects

Father's Education	Explanatory variable		beta	stdev	t-value	prob
Basic	W	$\overline{\Phi}_1$	**3.04**	**0.21**	**14.77**	**0.00**
		$s(\overline{\mu}_1)$	0.04	1.06	0.04	0.48
	Wdisp	$\overline{\Phi}_2$	**-3.41**	**0.76**	**-4.46**	**0.00**
		$s(\overline{\mu}_2)$	0.18	6.26	0.03	0.49
	Edpr	$\overline{\Phi}_3$	**26.92**	**3.11**	**8.64**	**0.00**
		$s(\overline{\mu}_3)$	0.27	18.56	0.01	0.49
	Expr	$\overline{\Phi}_4$	**-11.24**	**6.16**	**-1.83**	**0.03**
		$s(\overline{\mu}_4)$	**64.83**	**10.12**	**6.40**	**0.00**
	Ed	$\overline{\Phi}_5$	**0.06**	**0.02**	**3.81**	**0.00**
		$s(\overline{\mu}_5)$	0.01	0.16	0.04	0.48
Intermediate	W	Φ_1	-0.42	0.44	-0.97	0.17
		$s(\mu_1)$	0.03	2.15	0.02	0.49
	Wdisp	Φ_2	1.67	1.51	1.11	0.13
		$s(\mu_2)$	0.11	13.19	0.01	0.50
	Edpr	Φ_3	**-18.46**	**6.79**	**-2.72**	**0.00**
		$s(\mu_3)$	0.07	35.73	0.00	0.50
	Expr	Φ_4	9.32	11.14	0.84	0.20
		$s(\mu_4)$	**79.94**	**23.28**	**3.43**	**0.00**
	Ed	Φ_5	**0.18**	**0.03**	**5.39**	**0.00**
		$s(\mu_5)$	0.00	0.31	0.00	0.50
University	W	Φ_1	**-1.74**	**1.06**	**-1.65**	**0.05**
		$s(\mu_1)$	0.71	3.17	0.22	0.41
	Wdisp	Φ_2	**3.90**	**3.11**	**1.25**	**0.10**
		$s(\mu_2)$	0.33	26.77	0.01	0.50
	Edpr	Φ_3	**-31.34**	**19.60**	**-1.60**	**0.05**
		$s(\mu_3)$	0.36	62.54	0.01	0.50
	Expr	Φ_4	8.68	18.44	0.47	0.32
		$s(\mu_4)$	**73.40**	**50.81**	**1.44**	**0.07**
	Ed	Φ_5	**0.28**	**0.06**	**4.92**	**0.00**
		$s(\mu_5)$	0.00	0.53	0.00	0.50
	LogL-fn	-7424.32				

4077 observations, 200 draws. s(μ)=standard deviation of μ,
W=wage rate, Wdisp=wage dispersion, Edpr= return to education
Expr=return to experience, Ed=average years of education

The valuation of return to experience is negative which can be interpreted to indicating that people have time preferences. People with educated fathers are, however, significantly less negative about return to experience, indicating that they can afford to wait for pay off on their occupational choice. The only variable with significant variation in its coefficient, $s(\overline{\mu})$ i.e. with a significant variance component, is return to experience. An interpretation of this result is that the actual return to experience an individual can expect will be highly dependent on the ability to learn on the job and to climb up the career ladder. An alternative interpretation of the negative coefficient estimate on return to experience is thus that return to experience is viewed as uncertain. Moreover, there are background effects in the variation in the coefficient on return to experience, indicating that the presence of unobserved heterogeneity is background dependent.

Table 5.2.2:

Occupational choice with distance to occupation effects

Father's occupation	Explanatory variable		beta	stdev	t-value	prob
Familiar	W	$\overline{\Phi}_1$	**2.51**	**0.19**	**13.33**	**0.00**
		$s(\overline{\mu}_1)$	0.01	1.11	0.01	0.50
	Wdisp	$\overline{\Phi}_2$	**-1.54**	**0.98**	**-1.57**	**0.06**
		$s(\overline{\mu}_2)$	0.37	6.11	0.06	0.48
	Edpr	$\overline{\Phi}_3$	**10.66**	**3.97**	**2.68**	**0.00**
		$s(\overline{\mu}_3)$	0.88	19.17	0.05	0.48
	Expr	$\overline{\Phi}_4$	**15.36**	**6.13**	**2.50**	**0.01**
		$s(\overline{\mu}_4)$	**69.21**	**8.80**	**7.86**	**0.00**
	Ed	$\overline{\Phi}_5$	**0.17**	**0.02**	**6.86**	**0.00**
		$s(\overline{\mu}_5)$	0.00	0.15	-0.01	0.49
Unfamiliar	W	Φ_1	**0.15**	**0.08**	**1.82**	**0.03**
		$s(\mu_1)$	0.00	0.14	0.02	0.49
	Wdisp	Φ_2	**-1.70**	**1.35**	**-1.26**	**0.10**
		$s(\mu_2)$	0.14	5.80	0.02	0.49
	Edpr	Φ_3	**9.67**	**4.26**	**2.27**	**0.01**
		$s(\mu_3)$	0.60	13.22	0.05	0.48
	Expr	Φ_4	**-33.53**	**7.26**	**-4.62**	**0.00**
		$s(\mu_4)$	11.56	20.97	0.55	0.29
	Ed	Φ_5	**-0.07**	**0.03**	**-2.36**	**0.01**
		$s(\mu_5)$	0.00	0.05	0.01	0.49
	LogL-fn	-7393.67				

4077 observations, 200 draws. s(μ)=standard deviation of μ, W=wage rate, Wdisp=wage dispersion, Edpr=return to education, Expr=return to experience, Ed=average years of education

Table 5.2.2 presents the results of estimations when people are allowed to have different marginal valuations of incentives when they consider familiar and unfamiliar occupations. The first half of the table presents the estimates when considering a familiar occupation and the second half presents how marginal valuation *differs* when considering an unfamiliar occupation.

The general pattern from the previous estimations reappears with one exception. The coefficient on return to experience is positive for familiar occupations. This, again, supports the idea that return to experience is connected to uncertainty about ability. When considering a familiar occupation the uncertainty involved is small. As predicted by the theoretical model, people are more sensitive to incentives when they consider unfamiliar occupations. In particualr, people are more positive towards the wage rate, they dislike wage dispersion more strongly, they are more interested in return to education, they dislike return to experience and they are less willing to go for a long education.

As in the previous model, return to experience is the only variable with significant variation in its coefficient. There is, however, no significant difference in this heterogeneity when familiar or unfamiliar occupations are considered.

6. Conclusions

I have developed and tested a simple model of how family background influences the individual's occupational choice. The results suggest that understanding occupational mobility is indeed important for understanding the intergenerational transmission of earnings and income. The theoretical model shows how family background influences people's marginal valuations of economic incentives when making occupational choices. The empirical results verify the existence of family background effects.

In particular, I find that people are more sensitive to economic incentives when considering unfamiliar occupations than when considering familiar ones. Furthermore, the results show that people with poorly educated parents are more sensitive to economic incentives since they to a greater extent are attracted to occupations with high wage rates and high returns to education. I also find evidence that individuals in this group are more risk averse in their occupational choices since they are more negative towards wage dispersion and returns to experience than are people with well educated parents.

The results of the theoretical and empirical analyses suggest that policies affecting the incentive structure on the labor market are not neutral in terms of their effect on people from different family background. Poor return to education due to high taxes, compressed wage structure, or costly education financing will deter individuals from less well situated background to a larger extent than it will affect well off individuals. Likewise, unstable rules regarding education financing or income taxation that increase the uncertainty about future income or costs of education may have a greater impact on the occupational choices of the less well situated than on the choices of the well off.

References

[1] Becker, G. S. and N. Tomes (1979), "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility", *Journal of Political Economy*, vol. 87, no. 6, 1153-1189.

[2] Becker, G. S. and N. Tomes (1986), "Human Capital and the Rise and Fall of Families", *Journal of Labor Economics*, vol. 4, no. 3, pt. 2, S1-S39.

[3] Björklund, A. and M. Jäntti (1997), "Intergenerational Income Mobility in Sweden Compared to the United States", *American Economic Review*, vol. 87, no. 5, 1009-1018.

[4] Boskin, M. J. (1974), "A Conditional Logit Model of Occupational Choice", *Journal of Political Economy,* vol. 82, 389-398.

[5] Brownstone, D. and K. Train (1996), "Forecasting New product Penetration with Flexible Substitution Patterns", forthcoming in *Journal of Econometrics.*

[6] Erikson, R. and R. Åberg (1987), *Welfare in Transition - Living Conditions in Sweden 1968-1981,* Clarendon Press, Oxford.

[7] Flyer, F. A. (1997), "The Influence of Higher Moments of Earnings Distributions on Career Decisions", *Journal of Labor Economics*, vol. 15, no. 4. 689-713.

[8] Haveman, R. and B. Wolfe (1995), "The Determinants of Children's Attainments: A Review of Methods and Findings", *Journal of Economic Literature*, vol. XXXIII, December, 1829-1878.

[9] McFadden, D. L. and K. Train (1997), "Mixed Multinomial Logit Models for Discrete Response", working paper, Department of Economics, University of California, Berkeley.

[10] Mulligan, C. B. (1999), "Galton vs the Human Capital Approach to Inheritance", *Journal of Political Economy*, vol. 107(6 Part 2), S184-S224.

[11] Mulligan, C. B. (1997), *Parental Priorities and Economic Inequality*, The University of Chicago Press.

[12] Orazem, P. F. and J. P. Mattila (1991), "Human Capital, Uncertain Wage Distributions, and Occupational and Educational Choices", *International Economic Review*, vol. 32, no. 1(Feb), 103-122.

[13] Robertson, D. and J. Symons (1990), "The Occupational Choice of British Children", *The Economic Journal*, vol. 100, 828-841.

[14] Roy, A. D. (1951), "Some thoughts on the distribution of earnings", *Oxford Economic Papers*, 3:135-146.

[15] Sjögren, A. (1998), "The Effect of Redistribution on Intergenerational Mobility: Does wage equality nail the cobbler to his last?", *Working Paper Series in Economics and Finance,* no. 239, Stockholm School of Economics.

[16] Solon, G. (1999), "Intergenerational Mobility in the Labor Market", in *Handbook of Labor Economics,* vol 3A, ch 29, 1761-1800, eds O. Ashenfelter and D. Card, Elsevier Science B.V.

[17] Willis, R. J. (1986), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions" in *Handbook of Labor Economics,* vol. 1, ch. 10, 526-602, eds O. Achenfelter and R. Layard, Elsevier Science B.V.

A. Appendix

A.1. The individual's optimization problem

The individual maximizes the following utility function:

$$\max_{H_j} U(c_1, c_{2j}) = u(c_1) + \gamma E\left[u(c_{2j})\right], \tag{A.1}$$

where

$$E\left[u(c_{2j})\right] = \int_{-\infty}^{\infty} u(c_{2j}) f(a_j) da_j. \tag{A.2}$$

The first order condition is:

$$u'(c_1) = \gamma \int_{-\infty}^{\infty} u'(c_{2j}) W_j A_j^{\beta_j} f(a_j) da_j. \tag{A.3}$$

Given the chosen utility function:

$$(Y_1 - k_j H_j)^{-\frac{1}{\sigma}} k_j = \gamma \frac{H_j^{-\frac{1}{\sigma}} W_j^{1-\frac{1}{\sigma}}}{\sqrt{2\pi\rho}} \int_{-\infty}^{\infty} (A_j^{\beta_j})^{1-\frac{1}{\sigma}} \exp\left\{-\frac{1}{2\rho}(a_j - \overline{a_j})^2\right\} da_j \quad \text{(A.4)}$$

Because $(A_j^{\beta_j})^{1-\frac{1}{\sigma}} = \exp\left\{(1-\frac{1}{\sigma})\beta_j a_j\right\}$ I can write the integral:

$$I = \int_{-\infty}^{\infty} \exp\left\{\frac{2\rho(1-\frac{1}{\sigma})\beta_j a_j - (a_j - \overline{a_j})^2}{2\rho}\right\} da_j. \quad \text{(A.5)}$$

Separating out terms which do not contain the integrand, rewrite the integral:

$$I = \int_{-\infty}^{\infty} \exp\left\{\frac{-(\overline{a_j})^2}{2\rho}\right\} \exp\left\{\frac{-a_j^2 + 2[\rho(1-\frac{1}{\sigma})\beta_{j+}\overline{a_j}]a_j}{2\rho}\right\} da_j. \quad \text{(A.6)}$$

Complete the square in the second exponent by multiplying and dividing by $\exp\left\{\frac{-[\rho(1-\frac{1}{\sigma})\beta_{j1+}\overline{a_j}]^2}{2\rho}\right\}$ in second and first exponent respectively:

$$\begin{aligned} I &= \int_{-\infty}^{\infty} \exp\left\{\frac{[\rho(1-\frac{1}{\sigma})\beta_{j+}\overline{a_j}]^2 - (\overline{a_j})^2}{2\rho}\right\} \quad \text{(A.7)} \\ &\quad \exp\left\{\frac{-a_j^2 + 2[\rho(1-\frac{1}{\sigma})\beta_{j+}\overline{a_j}] - [\rho(1-\frac{1}{\sigma})\beta_{j+}\overline{a_j}]^2}{\rho}\right\} da_j, \end{aligned}$$

which simplifies to:

$$\begin{aligned} I &= \int_{-\infty}^{\infty} \exp\left\{\frac{[\rho(1-\frac{1}{\sigma})\beta_j]^2 + 2(\overline{a_j})\rho(1-\frac{1}{\sigma})\beta_j}{2\rho}\right\} \quad \text{(A.8)} \\ &\quad \exp\left\{\frac{-(a_j + [\rho(1-\frac{1}{\sigma})\beta_{j+}\overline{a_j}])^2}{2\rho}\right\}. \end{aligned}$$

Moving the first exponent out of the integral gives:

$$\begin{aligned} (Y_1 - k_j H_j)^{-\frac{1}{\sigma}} k_j &= \gamma H_j^{-\frac{1}{\sigma}} W_j^{1-\frac{1}{\sigma}} \exp\left\{\rho(1-\frac{1}{\sigma})^2 \frac{\beta_j^2}{2} + \overline{a_j}(1-\frac{1}{\sigma})\beta_j\right\} \quad \text{(A.9)} \\ &\quad \frac{1}{2\pi\rho} \int_{-\infty}^{\infty} \exp\left\{\frac{-(a_j + [\rho(1-\frac{1}{\sigma})\beta_{j+}\overline{a_j}])^2}{2\rho}\right\} da_j. \end{aligned}$$

Now, use the fact that $\frac{1}{\sqrt{2\pi\rho}}\int\limits_{-\infty}^{\infty}\exp\left\{\frac{-(a_j+[\rho(1-\frac{1}{\sigma})\beta_{j+}\overline{a_j}])^2}{2\rho}\right\}da_j$ is the integral of a normal distribution with mean $[\rho(1-\frac{1}{\sigma})\beta_{j+}\overline{a_j}]$ and variance ρ. The integral from $-\infty$ to ∞ of a normal distribution is always equal to one. This gives:

$$(Y_1 - k_j H_j)^{-\frac{1}{\sigma}} k_j = \gamma H_j^{-\frac{1}{\sigma}} W_j^{1-\frac{1}{\sigma}} \exp\left\{\rho(1-\frac{1}{\sigma})^2\frac{\beta_j^2}{2} + \overline{a_j}(1-\frac{1}{\sigma})\beta_j\right\}. \quad \text{(A.10)}$$

Rearrange to obtain an expression for the optimal human capital investment given that the individual chooses occupation j:

$$H_j^* = \frac{Y_1}{k + \left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\right)^{1-\sigma} e^{\frac{\rho}{2}\beta_j^2(1-\sigma)(1-\frac{1}{\sigma})}}. \quad \text{(A.11)}$$

A.2. Comparative statics on $\mathbf{H}^*$

$$\frac{\partial H^*}{\partial Y} = \left(k + \left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{1-\sigma}\right)^{-1} > 0,$$

$$\frac{\partial H^*}{\partial k} = -Y\frac{1+\sigma k_j^{-1}\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{1-\sigma}}{\left(k+\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{1-\sigma}\right)^2} < 0,$$

$$\frac{\partial H^*}{\partial W} = -Y\frac{(1-\sigma)W_j^{-1}\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\sigma)}}{\left(k+\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{1-\sigma}\right)^2} \gtrless 0 \quad if \quad (1-\tfrac{1}{\sigma}) \gtrless 0,$$

$$\frac{\partial H^*}{\partial \beta} = -Y\frac{(1-\sigma)\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\sigma)}\left(\overline{a}_j+\beta\rho(1-\frac{1}{\sigma})\right)}{\left(k+\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\sigma)}\right)^2} \gtreqless 0$$

$$if \quad \begin{cases} \left(1-\frac{1}{\sigma}\right) > 0 \text{ and } \overline{a}_j \gtreqless \beta\rho\left(\frac{1}{\sigma}-1\right) \\ \left(1-\frac{1}{\sigma}\right) < 0 \text{ and } \overline{a}_j \lesseqgtr \beta\rho\left(\frac{1}{\sigma}-1\right) \end{cases}$$

$$\frac{\partial H^*}{\partial a} = -Y\frac{(1-\sigma)\beta\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\sigma)}}{\left(k+\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\sigma)}\right)^2} \gtrless 0 \quad if \quad (1-\tfrac{1}{\sigma}) \gtrless 0$$

$$\frac{\partial H^*}{\partial \rho} = -Y\frac{(1-\sigma)\frac{\beta}{2}^2\left(1-\frac{1}{\sigma}\right)\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\sigma)}}{\left(k+\left(\frac{k_j}{\gamma}\right)^{\sigma}\left(W_j\overline{A}_j^{\beta j}\xi\right)^{(1-\sigma)}\right)^2} < 0.$$

A.3. Comparative statics on E(U)

$$\frac{\partial E[U]}{\partial Y} = (Y - kH^*)^{-\frac{1}{\sigma}} > 0,$$
$$\frac{\partial E[U]}{\partial k} = -(Y - kH^*)^{-\frac{1}{\sigma}} H^* < 0,$$
$$\frac{\partial E[U]}{\partial W} = \gamma W_j^{-1} \left(H^* W_j \overline{A}_j^{\beta j} \xi\right)^{(1-\frac{1}{\sigma})} > 0,$$
$$\frac{\partial E[U]}{\partial \beta} = \gamma \left(H^* W_j \overline{A}_j^{\beta j} \xi\right)^{(1-\frac{1}{\sigma})} \left(\overline{a}_j + \beta\rho(1 - \tfrac{1}{\sigma})\right) \gtreqless 0 \quad if \quad \overline{a}_j \gtreqless \beta\rho\left(\tfrac{1}{\sigma} - 1\right)$$
$$\frac{\partial E[U]}{\partial a} = \gamma \left(H^* W_j \overline{A}_j^{\beta j} \xi\right)^{(1-\frac{1}{\sigma})} \beta > 0,$$
$$\frac{\partial E[U]}{\partial \rho} = \gamma \left(H^* W_j \overline{A}_j^{\beta j} \xi\right)^{(1-\frac{1}{\sigma})} \frac{\beta^2}{2}(1 - \tfrac{1}{\sigma}) \gtrless 0 \quad if \quad (1 - \tfrac{1}{\sigma}) \gtrless 0.$$

A.4. Occupational classification

Table A.4:
Occupations classified

Occupation	NYK85	Description
professional	001-099	Professional, technical and related work: technical, scientific, educational, religious, law, literary, journalistic, artistic
social/ medical	101-199	Health, nursing and social work: Medical, nursing, physiotherapy, dental, pharmaceutical, social,. health protection
administrative	201-299	Administrative, managerial and clerical work: Public and business administration, accounting, clerical, IT, economics and statistics
sales	311-399	Sales work: sales (business services, assets and goods), purchasing
agriculture	400-449	Agricultural, forestry and fishing work: Agriculture, horticulture, livestock and forestry management and work, wildlife protection, fishing
transp&com	601-649	Transport and communications work: Drivers, Train and flight personnel, telecom's workers.
production	501-599 701-891	Mining and production work: mining, textile, leather, metal processing, machine, electrical, wood, chemical, food, construction, painting.
service	901-989	Service work: Civilian protection, lodging and catering, cleaning, laundring, military.