Firm-Size Wage Differentials in the Netherlands

Hessel Oosterbeek
Mirjam van Praag

ABSTRACT. An unsolved problem in modern labor economics is the positive relation between the size of the firm in which a worker is employed and his wage. One line of research that has been developed quite recently in this field is the application of endogenous switching regression models. In this paper we utilize such a model to investigate firm-size wage differentials in the Netherlands. The principal findings are that larger firms pay higher returns on schooling whereas smaller firms tend to reward IQ. Combined with the finding that high IQ-workers are sorted into the largest firms, the results are consistent with a model of job screening. Furthermore, we find that employed sons of self-employed fathers are more likely to work in small firms and that wage prospects for all types of workers are indeed most favorable in larger firms.

1. Introduction

A still unsolved anomaly in modern labor economics is the stylized fact that workers in large firms earn higher wages than those in small firms. The fact is an anomaly since it is inconsistent with the mechanism that firms adjust the size of their work-force until the value of marginal product equals the exogenous wage rate. Several explanations have been put forward. In their famous article, Brown and Medoff (1989) investigate the contributions of different factors to the explanation of the observed phenomenon with reference to the U.S. Factors which they mention are: (i) worker quality; (ii) working conditions; (iii) unionization, and (iv) imperfect information effects (efficiency wages and monitoring). Although each of these factors is relevant, the authors conclude that there "remains a significant wage premium for those working for large employers" (p. 1027). Using a data set referring to West Germany, Schmidt and Zimmermann (1991) come to a similar conclusion.

After the publication of the Brown and Medoff article, several papers have appeared that apply a member of the family of endogenous switching regression models to the relation between firm-size and wages. These models focus on the possibility that workers are sorted into the size of firm that best suits their personal characteristics. These characteristics need not be observable for either the firm or the researcher in order to influence both the probability of being sorted into a certain firm-size class and the potential wage rate within this class. Idson and Feaster (1990) estimate a switching regression model in which they distinguish five different firm-size categories and hence five different wage regimes. After estimating the model, they decompose the size-wage gap into three effects: an endowment effect, a coefficient effect and a selection bias effect. The main conclusion is that the calculations show that "selectivity considerations act to reduce the wage gap between small and large firms" (p. 116). These results refer to the U.S. A similar analysis has been performed by Main and Reilly (1993) in relation to Britain. Including more explanatory variables in their selection equation than Idson and Feaster, they report "no evidence of non-random sorting of workers across plant size" (p. 125).

In the present paper, we employ essentially the same methodology as Idson and Feaster, and Main and Reilly. The contribution of this paper is twofold. First and most importantly, our analysis relates to the Netherlands, and as far as we are aware this is the first paper that addresses the firm-size wage relation for that country. A reason why the Dutch case is of special interest is that collective agreements in an industry do not only apply to the firms represented in the bargaining process, but – due to mandatory extension – also

Final version accepted on August 15, 1994

University of Amsterdam
Amsterdam, the Netherlands

to all other firms in the industry. Secondly, we utilize a data set that is particularly suitable to estimate an endogenous switching regression model. Besides information regarding firm-size and wages, our data set contains statistics concerning social background and IQ. We think that this kind of information is indispensable for the type of model involved. In a sense, the whole idea underlying self-selection models comes down to dealing with the problem of mis-specification. If the researcher has all the relevant information available, no factors crucial to both the wage equation and the selection function will be omitted.

The remainder of this paper is divided into four sections. Section 2 provides a brief outline of the model and the estimation procedure. Section 3 describes the data set and explains our choice of variables. Section 4 presents and discusses the empirical results, and Section 5 sums up the main conclusions.

2. Model and estimation procedure

The standard approach to analyzing wage data for different contingencies where the observed contingency might be endogenous with respect to the wage rate is to apply a version of the endogenous switching regression model. Which specific version of this model is employed depends on the nature of the state variable. The most well-known version is that where the selection variable is dichotomous, in which case the selection equation is specified as a probit-model.[1] More involved models use (for instance) multinominal probit or logit models.

In the case of firm size, the dependent variable is ideally measured as a continuous variable. In reality, however, most labor market surveys that we know of measure firm-size in a limited number of classes. In our data set we can distinguish the usual three firm-size classes: small (less than 25 employees), medium (between 25 and 500 employees) and large (more than 500 employees). In this case, the nature of the dependent variable is ordinal (small < medium < large). The natural choice is therefore to apply an ordinal probit model.[2]

Assume that the firm-size is a semiloglinear function of the exogenous variable vector Z, associated parameter vector γ and the error term u, hence

$$\log(\mathrm{FS}) = Z\gamma + u, \tag{1}$$

where u is assumed to follow a normal distribution with mean zero and variance σ_u^2. Let α_{j-1} be the lower and α_j be the upper threshold of firm-size category j. Then a firm in which individual i is employed belongs to the jth category if

$$\alpha_{j-1} < \log(\mathrm{FS}) < \alpha_j \quad (j = 1[\text{small}], 2[\text{medium}], 3[\text{large}]), \tag{2}$$

where $\alpha_0 \to -\infty$ and $\alpha_3 \to +\infty$. Define $\mathrm{FSC}_j = 1$ if an individual is employed by a firm in firm-size class j, $\mathrm{FSC}_j = 0$ otherwise. Hence $\mathrm{FSC}_j = 1$ if $\alpha_{j-1} < \log(\mathrm{FS}) < \alpha_j$. Since we are interested in explaining the probability that an individual will work in FSC_j, and since by assumption $u \sim N(0, \sigma_u^2)$, we can write for the probability that an individual will be employed by a firm in the firm-size class small, medium and large respectively:

$$\Pr(\mathrm{FSC}_1 = 1) = \Phi((\alpha_1 - Z\gamma)/\sigma_u), \tag{3a}$$

$$\Pr(\mathrm{FSC}_2 = 1) = \Phi((\alpha_2 - Z\gamma)/\sigma_u) - \Phi((\alpha_1 - Z\gamma)/\sigma_u), \tag{3b}$$

$$\Pr(\mathrm{FSC}_3 = 1) = 1 - \Phi((\alpha_2 - Z\gamma)/\sigma_u), \tag{3c}$$

where Φ is the cumulative standard normal distribution function.

For each firm-size class j we specify a wage equation in the form

$$\log(y_j) = X\beta_j + v_j \; (j = 1, 2, 3) \tag{4}$$

where we omit the subscripts i pertaining to individuals for readability.

We assume that the vector of error terms v_j and u follow a multivariate normal distribution: $E(v_j^2) = \sigma_j^2$, $E(u^2) = \sigma_u^2$, $E(v_k v_l) = \sigma_{kl}$ and $E(uv_j) = \sigma_{uj}$. Since Φ is an increasing function of its argument and since σ_u must be positive, we can directly interpret the signs of the estimated coefficients.

Instead of maximizing the likelihood function of the full model, it is possible to estimate the wage equations and the selection equation using a two-stage method. This method is a slightly modified version of the one proposed by Lee (1979). The first step in this method is to purify the estimates of the wage equations of the selection bias implied by the fact that the expected value of the error term does not equal zero for the sub-sample of respondents observed in a particular wage regime. To be more precise,

$$E(\ln y_j \mid X) = E(X\beta_j + v_j) = X\beta_j + E(v_j \mid FSC_j = 1) \quad (5)$$

This latter term is not equal to zero, but instead

$$\begin{aligned} E(v_j \mid FSC_j = 1) &= E(v_j \mid \alpha_{j-1} < FS < \alpha_j) \\ &= \sigma_{ju}\{[\phi((\alpha_{j-1} - Z\gamma)/\sigma_u) - \phi((\alpha_j - Z\gamma)]/ \\ &\quad [\Phi((\alpha_j - Z\gamma)/\sigma_u)) - \Phi((\alpha_{j-1} - Z\gamma)/\sigma_u)]\} \\ &= \sigma_{ju}\lambda_j, \end{aligned} \quad (6)$$

where φ is the density function of the standard normal. The λ_js can be extracted from the ordered probit estimation as discussed above. Hence, σ_{ju} is an additional estimable coefficient. The wage equations then become

$$\ln y_j = X\beta_j + \sigma_{ju}\lambda_j + w_j, \quad (7)$$

where $E(w_j \mid FSC_j = 1) = 0$ for $j = 1, 2, 3$. Hence, these wage equations can be consistently estimated by OLS. It should be noted that of the four parameters α_1, α_2, σ_u and γ_0 (the coefficient for the constant term in γ), only two can be identified. Usually σ_u is set equal to unity and γ_0 is set equal to zero. However, our equation (1) gives rise to more intuitive identification restrictions. Since we know the class intervals, we can set α_1 equal to log(25) and α_2 equal to log(500). The results of the analysis are not affected by this choice.[3]

3. Data and choice of variables

In 1952, a survey was conducted amongst children in the last grade of elementary school in the Dutch province of Noord-Brabant, in order to obtain information on ability, school performance, family background and related variables. In 1982, thirty years later, the 5,800 observations were still available. After an administrative search of local government population registers in order to update respondents' addresses, we were in a position to recontact 4,700 individuals in 1983. A questionnaire was sent to all of them, this time gathering data on education, household composition, labor market status and earnings. After two reminders had been sent, the remaining male non-respondents were approached by an interviewer. It was decided (because of budgetary restrictions) to approach men only, because of their higher labor force participation rate. This explains the fairly small proportion of women within the sample. The response rate, based on the 4,700 individuals located, was 58%.

The sample has some distinct features that are worth pointing out. Firstly, all individuals were born in 1939 or 1940. Consequently, the sample is very homogeneous with respect to age. Secondly, as already mentioned, all individuals were living in the same region, Noord-Brabant, when they were approximately 12 years old. Earlier analysis (Hartog and Pfann, 1985) indicated that Noord-Brabant can quite harmlessly represent the entire country, insofar as that many variables related to labor force participation, industrial structure and earnings are not significantly different from values observed in the Netherlands as a whole. Another interesting feature is the availability of early childhood variables.

To fill the sectors Z (the determinants of firm-size) and X (the determinants of wages), we have made the following choices. The size of firm that a worker joins is believed to be affected by innate human capital (IQ), general human capital (schooling) and specific human capital (company training). Including IQ and schooling is inspired by the theoretical model proposed by Garen (1985). He argues that because of higher monitoring costs, larger firms obtain less accurate information about their workers' abilities and therefore reward schooling more and ability less than smaller firms do. Knowing this, workers will join the firm-size class that best fits their schooling/ability-combination. Reasons to include the potentially endogenous variable of company training are the following. Firstly, our measure of company training relates to all past company training and is not restricted to training in the present firm. To the extent that it relates to training obtained with other firms, the variable is exogenous. Secondly, if company training is positively related to schooling and IQ, omitting company training would then bias the estimated coefficients for schooling and IQ. Finally, specific vocational training is one of the few regressor variables in the selection equation which is estimated by Idson and Feaster. Including a similar variable in our analysis allows us to compare the results. It is important, however, to note that the results reported in this paper are not different from those obtained from a version which excluded company training from the model.

Another determinant of the size of the firm by which a worker will be employed is the worker's marital status (dummies for not being married and having a working partner) as proxy for his or her stability or social integration, social background (measured by the father's education and whether the father was self-employed when the child was 12 years old) and the sector in which he works (dummies for agriculture and trade/hotels/repair). It is noted that Z contains both variables that relate to the worker and variables that relate to the employer. This mixture follows from the fact that the observed employment of a worker in a particular firm-size class requires both that the worker decided to work in that class and that there was an employer in that class who decided to hire that worker. Instead of observing both decisions separately, we observe the single indicator FSC_j. From this perspective, the selection equation is a reduced form equation in which it is quite natural to include both demand and supply side characteristics of labor (cf. Maddala 1986, p. 1677).

Regressors included in the wage equations, in addition to the human capital variables, are social background, number of working hours per week and two dummies related to differences in tax treatment (not being married and having no children). Working hours and tax dummies are included because the dependent variable within the wage equations is net hourly wage. This variable is calculated by dividing the reported after-tax earnings per period by the reported number of hours worked during the same period. As a very simple approximation for the possible non-homogeneity of the wage rate with respect to hours, we include hours in the list of explanatory variables of wages. The dependent variable in the selection equation is firm-size in categories as this is reported by the employee. It is sometimes argued that establishment size, besides firm-size, would provide an additional explanation for individual wages (Brown and Medoff, 1989). Unfortunately, since establishment size is not known, we cannot test the explanatory power of that variable. Descriptions, means (for the sub-samples used) and standard deviations are presented in Table I.

The fact that not every respondent participates in the labor market reduces the sample to 1,599 observations. After eliminating females, public sector workers, self-employed persons and respondents who failed to report the size of their firm, their wage or their number of working hours, we are left with 569 observations.

The figures in Table I already reveal some patterns. The (log) wage rate does indeed increase with firm-size. The same is true for average IQ, years of schooling and amount of company training. We estimated a simple log wage equation

TABLE I
Descriptive statistics of variables by firm-size category; means and standard deviations

Variable	Included in*	Small	Medium	Large
Net hourly wage rate		14.08 (3.84)	15.68 (8.05)	17.10 (6.06)
Log net hourly wage rate		2.61 (0.26)	2.68 (0.37)	2.78 (0.34)
IQ	s, w	99.68 (12.74)	100.75 (13.11)	105.25 (13.73)
Schooling; years of post-primary schooling	s, w	3.07 (2.72)	3.45 (2.97)	4.48 (3.85)
Company training; in full-time years	s, w	0.04 (0.18)	0.15 (0.54)	0.34 (0.68)
Father's education; 1 = lowest, . . . , 6 = highest	s, w	2.29 (0.48)	2.24 (0.43)	2.45 (0.85)
Self-employed father	s, w	0.28 (0.45)	0.26 (0.44)	0.18 (0.39)
Not married	s, w	0.06 (0.23)	0.09 (0.29)	0.08 (0.27)
No children	w	0.09 (0.28)	0.15 (0.35)	0.15 (0.35)
Working partner	s	0.28 (0.45)	0.26 (0.44)	0.25 (0.43)
Working hours	w	42.56 (7.81)	43.05 (6.77)	42.28 (5.91)
Agriculture	s	0.02 (0.14)	0.01 (0.09)	0.00 (0.00)
Trade/hotels/repair	s	0.27 (0.44)	0.13 (0.33)	0.06 (0.23)
λ	w	−2.29 (0.36)	−0.09 (0.37)	1.00 (0.32)
Observations		138	246	185

* w indicates that the variable is included in the wage equation, and s that the variable is included in the selection equation.

on the full sample using IQ, years of schooling, company training, father's education, self-employed father, not married, no children, working hours and dummies for medium and large firm-size as regressors. This gives coefficients of 0.067 and 0.094 for the respective firm-size dummies, both with t-ratios well above 2.0. This establishes that *other (observed) things being equal*, wages increase with firm-size.

Other notable results in Table I relate to the declining percentages of people working in the trade/hotels/repair sector or who are sons of self-employed persons as firm-size increases. The average levels of father's education, marital status and working hours do not vary systematically with firm-size.

Explanatory variables frequently included in related literature are tenure, union-membership and – to capture compensating differentials – job characteristics such as shift work and heavy work. These variables are not available in our data-file. Union-membership is less of a problem in the Dutch context than it would be in other institutional settings. Whereas in other countries individual wages vary to quite a large extent according to whether or not workers are members of unions, such a relation is not present in the Netherlands (cf. Van den Berg and Groot, 1992) since agreements between unions and employers apply to all workers in a sector (union and non-union alike).

4. Estimation results

Table II presents the results of the first step of the analysis, the estimation of the ordered probit selection model. The results confirm the picture already emerging after inspecting Table I: the probability of ending up in a large firm increases with years of schooling, company training and IQ, and decreases with the dummies for self-employed father, agricultural sector and the trade/hotels/repair-sector.[4] The positive relation between all types of human capital (IQ, school and company training) and firm-size corroborates the hypothesis that human capital and physical capital are complements (assuming that large firms are more capital intensive than small firms).[5]

The findings in Table II are consistent with the implications of the theoretical literature relating to firm-size, which suggest that workers are matched to firms of differing size on the basis of the cost of monitoring employees (Oi, 1983a, b). According to Oi, firms of larger (staff) size face different costs with regard to monitoring the productive performance per employee at the margin of production. He assumes that a larger work force incurs greater monitoring costs in hiring an additional worker. This implies that the marginal monitoring costs within a large firm exceed those within a smaller firm. Given that workers differ with respect to the amount of monitoring required, large firms should display a greater propensity to hire workers with characteristics predisposing them not to need too much monitoring. And these characteristics are likely to be reflected in education. Hiring a marginal worker with more education can be seen as ensuring productivity, where monitoring is less vital in order to ensure a given level of productivity. The positive effect of education on selected employer size can also be explained by reference to an internal labor market view. Large firms are attractive to workers possessing a large amount of human capital because they offer more opportunities for internal promotion.

TABLE II
Estimation results of the selection equation

Variable	Coefficient	t-value
Constant	2.328	(2.27)*
Schooling	0.104	(2.54)*
Company training	0.936	(4.18)*
Not married	0.390	(0.91)
IQ	0.019	(1.98)*
Self-employed father	–0.515	(1.94)**
Father's education	0.305	(1.44)
Working partner	–0.272	(1.05)
Agriculture	–2.836	(2.18)*
Trade/hotels/repair	–1.916	(5.63)*
σ_u	2.364	(18.59)*
Number of observations	569	
Loglikelihood	–565.6	

* Significant at the 5% level;
** Significant at the 10% level.

Table III presents the results of the second part of the analysis, the wage equations corrected for selectivity by an additional regressor. The three wage regimes are estimated by means of OLS. The dependent variable is the logarithm of net wage rate per hour.

TABLE III
Estimation results wage equations

Variable	Small firms	Medium firms	Large firm
Constant	2.294 (9.86)*	2.736 (11.50)*	2.935 (8.74)*
IQ/100	0.287 (1.62)	0.276 (1.50)	0.113 (0.63)
Schooling	0.021 (2.47)*	0.042 (4.85)*	0.047 (6.18)*
Company training	–0.004 (0.03)	–0.051 (0.94)	–0.035 (0.80)
Father's education	0.122 (2.75)*	0.057 (1.11)	0.033 (1.23)
Self-employed father	–0.031 (0.67)	0.013 (0.26)	0.007 (0.13)
No child	0.007 (0.06)	0.039 (0.43)	0.031 (0.38)
Not married	–0.219 (1.51)	–0.108 (0.94)	–0.282 (2.64)*
Hours per week	–0.008 (2.89)*	–0.014 (4.34)*	–0.011 (2.98)*
λ	–0.017 (0.26)	–0.074 (0.82)	–0.080 (0.61)
Observations	138	246	185
Standard error or regression	0.234	0.330	0.266
Adjusted R^2	0.18	0.20	0.39

Absolute t-values in parentheses; * significant at the 5% level; ** significant at the 10% level.

The results in Table III reveal some nice regularities.[6] First of all we note that the results with respect to the intercept, the schooling coefficient and the coefficient for IQ tend to be consistent with Garen's (1985) prediction that "large firms will reward schooling more and measured ability less, and their wage schedule will have a larger intercept term" (p. 716). Indeed, we find that the intercept term rises with firm-size class. Moreover, the rate of return on schooling doubles from small to medium firms, and is half a percent higher in large firms than in medium firms. Also – albeit, just below the conventional 10% level threshold – a higher IQ is about equally rewarded in both the small and medium-sized firms, whereas a higher IQ produces no wage gain in large firms. Garen's predictions are derived from a model in which the key assumption is that monitoring costs rise with firm-size. Because of these higher costs, large firms obtain less accurate information about their workers' abilities and therefore have to rely more heavily upon such indicators of ability as schooling.

The fact that workers with a higher IQ are sorted into larger firms, which do not pay for this characteristic directly, indicates that large firms are quite successful in using schooling as a screening device. By providing a higher rate of return on schooling, large firms are able to attract workers with a large amount of schooling and hence, because of the positive correlation between schooling and IQ, workers with high levels of IQ.[7] In contrast to our results, Main and Reilly (1993) find that "the estimated returns to education [are] highest in 'small' plants with the effect in 'large' plants insignificant". This finding contradicts Garen's prediction. However, Main and Reilly did not include an ability measure in their wage equations and, with higher rewards for ability in small firms and a positive association between schooling and ability, their finding may be due to omitted variable bias.

Furthermore, the results in Table III indicate that the amount of company training that a worker has obtained has no effect on his wage rate. This finding is consistent with the human capital view of on-the-job training only if most (or all) of the company training is firm-specific. In that case, it may be that the firm bears the full cost and receives the full return.

An interesting finding is that the father's education does affect the wages of employees who have joined a small firm, whereas it has no influence on wages in the other two firm-size classes. This result can be explained along similar lines to the finding for IQ. Father's education (as the researcher's measure for a person's cultural capital) is positively related to the worker's own level of schooling.[8] Large firms do not observe a worker's cultural capital but they do value it and pay for it by rewarding the individual's own schooling. Small firms are able to observe a worker's cultural capital and can pay for it directly. Presumably the indirect reward in large firms and the direct reward in small firms balance, as father's education does not enter the selection equation with a significant coefficient.

Whether the father was self-employed when the son was 12 years old is a factor with no effect on the wage level. If we interpret the dummy "self-employed father" as an indicator of – what Idson and Feaster label – the "individual's preference for independence", this result indicates that this preference is rewarded in none of the firm-size classes. In fact, since sons of self-employed fathers have a higher probability of working in small firms and since small firms pay lower wages, it seems as though workers with a preference for independence are prepared to accept a lower wage rate in order to exercise their independence.

The two "tax dummies" included indicate that workers in medium-sized firms are capable of shifting their different tax treatment to the employer, with single males in small and large firms having a net wage loss. Another interpretation of the "not married" (when 43 years old) dummy is that this variable measures whether a person is socially integrated or is a stable person. A person who is unstable or socially unintegrated (not married) will be at a disadvantage whether he works in a small or a large firm. The coefficients for the number of working hours are about the same in the three wage equations, all indicating that longer working hours decrease the average net wage rate. This result might be due to the progressiveness of income taxes.

None of the coefficients for the correction terms (λ_js) differ significantly from zero. This implies that OLS results would not have suffered from selection bias. Apparently there are no unobserved (by the researcher) characteristics that affect the worker's selection into a specific firm-size class and are at the same time related to the wage the worker could obtain in that firm-size class.

Finally, we note that there is no systematic pattern with respect to the standard error of the wage regression and the firm-size class. This outcome is at odds with Garen's "tentative hypothesis" that, given schooling and other observable variables, the dispersion of wages will be smaller among larger firms, since workers cannot be differentiated easily in such firms and this fact should be reflected in their wages (p. 736).

The results in Table III allow us to decompose the wage differential between employees in different firm-size classes into an endowment effect, a coefficient effect and a selection bias effect. The decomposition method is basically the same as that used in discrimination analysis. The average estimated logarithmic wage per hour within each size category is

$$E(\ln y_j \mid FSC_j = 1) = \beta_j' X_j^* + \sigma_{uj}\lambda_j^* \quad \text{for } j = 1, 2, 3. \tag{8}$$

Where X_j^* denotes $(1/n_j)\sum_{i=1}^{nj} X_{ij}$ and λ_j^* denotes $(1/n_j)\sum_{i=1}^{nj}\lambda_{ij}$. In turn, n_j denotes the number of individuals i belonging to FSC_j. Hence, the total average wage difference between FSC_j and FSC_k is

$$\begin{aligned} &E(\ln y_j \mid FSC_j = 1) - E(\ln y_k \mid FSC_k = 1) \\ &= \beta_j' X_j^* + \sigma_{uj}\lambda_j^* - \beta_k' X_k^* - \sigma_{uk}\lambda_k^* \\ &= (\beta_j' X_j^* - \beta_k' X_k^*) + (\sigma_{uj}\lambda_j^* - \sigma_{uk}\lambda_k^*) \\ &= (A + B) + C; \quad j = 1, 2, 3; \quad k = 1, 2, 3, \end{aligned} \tag{9}$$

where A + B denotes some combination of differences in characteristics and rewards to these characteristics, and C the difference in selection bias. For the decomposition of the first two effects, we use a decomposition as applied by Oaxaca (1973):

$$\begin{aligned} &(\beta_j' X_j^* - \beta_k' X_k^*) \\ &= 0.5(\beta_k + \beta_j)'(X_j^* - X_k^*) + \\ &\quad 0.5(\beta_k - \beta_k)'(X_j^* + X_k^*) \\ &= A + B \end{aligned} \tag{10}$$

where term A now reflects the wage differences as a consequence of different average endowments of characteristics determining the wage rate and term B represents these average differences between wages in firm classes as a consequence of different reward for endowments. If we express each of these terms as a percentage of the total (not logarithmic) wage rate difference and also express these wage rate differences across firm-size classes as percentage, we obtain Table IV.

Let us first look at these results net of the selection effect, i.e. wage differences amongst classes as if assignment to classes had occurred purely on the basis of observed characteristics. Comparing medium to small shows that the observed difference is entirely due to the fact that medium-

TABLE IV
Decomposition of wage rate differences in percentages

Wage rate differences in percentages as a consequence of differences in:	Medium compared to small	Large compared to small	Large compared to medium
Endowments of characteristics	0	8	6
Rewards for these characteristics	16	28	11
Selection bias	–5	–15	–8
Total wage difference per hour after tax	11	21	9

sized companies reward endowments more than small companies. Differences in observed characteristics play no role. Comparing large to small firms reveals that 75% of the effect is a consequence of different wage regimes, and 25% a result of different endowments. And a comparison of large and medium-sized firms shows that the wage difference is one-third due to the fact that large companies are in a position to attract workers with more favorable endowments, and two-thirds due to differing wage regimes. The endowments positively rewarded are most strongly represented within the large firms and least strongly within the small firms. Employers in large firms apparently pay the best wages given the same distribution of endowments.

Since we know whether an individual is working within a certain firm-size category, we are in a position to look at the effect of unobserved characteristics on wages, i.e. the selection bias. This reveals that selection on unobserved grounds tends to reduce all wage gaps, though the selection effect was shown to be not statistically significant.

An alternative mode of expressing the wage differentials by firm-size class is to calculate the predicted wages of a person observed in a particular class in each of the other classes. These calculations can take account of the selection term, in which case we speak of conditional predicted wages, or can neglect the selection term, in which case we speak of unconditional predicted wages. Table V reports the results of this exercise.

TABLE V
Predicted wages of FSC_i-employees in FSC_j

Wage rate with row characteristics and column coefficients	Small	Medium	Large
Unconditional			
Small	13.21	14.68	16.62
Medium	13.16	14.73	16.70
Large	14.22	15.89	17.92
Conditional			
Small	13.72	17.40	20.02
Medium	13.18	14.85	16.85
Large	14.00	14.80	16.59

For all endowment groups, both the unconditional and the conditional predicted wages peak within the wage regime of the large firm-size class. That implies that from a wage maximization point of view all workers would prefer to work in this firm-size class. Which firm-size class attracts the workers with the most valuable characteristics is less clear. The unconditional predicted wages in the top half of the Table suggest that workers in large firms possess the best endowment-bundle as they earn the highest wage in every wage regime. The conditioned predicted wages, however, point to the conclusion that workers who are actually employed in small firms would have earned the highest wage in both the large and medium-sized firms. The different conclusions from the two parts of the Table are caused by the relatively large negative coefficients of the selection terms in the medium and large firms wage regimes. Although neither of these coefficients is statistically significantly different from zero, they both have this major impact.

5. Conclusion and summary

An intriguing question in the empirical literature is why observably equal workers earn different wages if they work in firms of different sizes. The larger the firm, the higher the level of wages. In the literature a large number of different explanations have been put forward, but a significant wage differential remains unexplained. The research presented in this paper further explores the possible contribution of heterogeneity of workers. To that end we estimated an endogenous switching regression model and distinguished three firm-size classes: small, medium and large.

The empirical results are quite interesting. The results confirm the thesis that workers who require less monitoring (i.e. who are educated more highly and more able) are employed in larger firms. Moreover the wage schedules that firms of different sizes employ are different, with large firms paying higher rewards for schooling and small firms tending to pay better for ability. The combination of these results supports a screening view with respect to hiring decisions: large firms are incapable of observing workers' ability, but by offering a higher rate of return on schooling they succeed in attracting the most able workers.

We decomposed the wage differential between each pair of firm-size classes for average workers into an endowment effect, a reward effect and a selection effect. The observed differential between small and medium-sized firms can be attributed entirely to the reward effect, while that between small and large firms is mainly caused by that effect. Finally, the decomposition of the wage difference between medium-sized and large firms shows that large firms pay higher rewards for certain characteristics than medium-sized firms and that workers in large firms are endowed with the preferred characteristics. These findings are reiterated by our calculations of the conditional and unconditional predicted wages. These calculations indicate that for all types of workers predicted wages – either conditional or unconditional – are highest in large firms.

Acknowledgement

We gratefully acknowledge comments by Joop Hartog, Arie Kapteyn, Coen Teulings and two anonymous referees.

Notes

[1] Examples include: Lee (1978) with respect to union membership; Willis and Rosen (1979) with respect to college choice; Rees and Shah (1986) with respect to self-employment; and Van der Gaag and Vijverberg (1988) with respect to public sector employment.

[2] Studies that use the ordered probit structure include: Jiminez and Kugler (1988), studying the choice of training duration; Hartog (1988) and Van Ophem, Hartog and Vijverberg (1993), investigating job level allocation, and Idson and Feaster (1990), studying firm size selection.

[3] By imposing a similar restriction Björklund and Moffitt (1987) identify the variance of the probit equation. Since they use this trick in the context of a dichotomous decision variable which is better suited to a structural interpretation, they are able to calculate the welfare gains from the distinct choices.

[4] Dummy variables for other industries do not significantly affect this probability. It might be argued that the industry dummies are endogenous. In that respect it should be noted that our results remain unchanged after the exclusion of the industry dummies from the selection equation.

[5] A comparable claim with respect to the effect of schooling and training on firm size is made by Idson and Feaster (1990), although their results fail to support their conclusion. Their descriptive statistics show that the mean amount of special vocational training increases with firm size category, but this result does not stand up in a multivariate analysis.

[6] We tested for total differences between wage regimes by means of chow tests. The test for structural differences in sets of coefficients estimated on subsets of a sample reveals that all equations which we have used are significantly different from each other at a 1% level [Chow(1vs2) = 22.7, Chow(1vs3) = 76.1 and Chow(2vs3) = 11.7; these values are compared to the critical value of the F(10, 549) distribution].

[7] Using the same dataset, Oosterbeek (1990) reports a derivative of actual years of schooling with respect to IQ of 0.09, with a t-value of 10.8.

[8] Again using the same dataset, Oosterbeek (1990) reports a derivative of actual years of school with respect to father's education of 0.88, with a t-value of 5.1.

References

Björklund, A. and R. Moffitt, 1987, 'The Estimation of Wage Gains and Welfare Gains in Self-Selection Models', *The Review of Economics and Statistics* **69**, 42–49.

Brown, C. and J. Medoff, 1989, 'The Employer Size-Wage Effect', *Journal of Political Economy* **97**, 1027–1059.

Garen, J. E., 1985, 'Worker Heterogeneity, Job Screening, and Firm-size', *Journal of Political Economy* **93**, 715–739.

Hartog, J., 1988, 'An Ordered Response Model for Allocation and Earnings', *Kyklos* **41**, 113–141.

Hartog, J. and G. A. Pfann, 1985, Vervolgonderzoek Noordbrabantse Zesdeklassers (University of Amsterdam, Amsterdam).

Idson, T. and D. Feaster, 1990, 'A Selectivity Model of Employer-Size Wage Differentials', *Journal of Labor Economics* **8**, 99–122.

Jimenez, E. and B. Kugler, 1988, 'The Earnings Impact of Training Duration in a Developing Country; an Ordered Probit Selection Model of Colombia's Servicio Nacional de Aprendizaje (SENA)', *The Journal of Human Resources* **22**, 228–247.

Lee, L. F., 1978, 'Unionism and Wage Rates: A Simultaneous Equation Model with Qualitative and Limited-Dependent Variables', *International Economic Review* **19**, 415–433.

Lee, L. F., 1979, 'Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables', *Econometrica* **47**, 977–996.

Maddala, G. S., 1986, 'Disequilibrium, Self-Selection, and Switching Models', in Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics, Volume III*, Amsterdam: North-Holland.

Main, B. G. M. and B. Reilly, 1993, 'The Employer Size-Wage Gap: Evidence for Britain', *Economica* **60**, 125–142.

Oaxaca, R., 1973, 'Male-Female Wage Differentials in Urban Labour Markets', *International Economic Review* **14**, 693–709.

Oi, W., 1983*a*, 'Heterogeneous Firms and the Organization of Production', *Economic Inquiry* **21**, 147–171.

Oi, W., 1983*b*, 'The Fixed Employment Cost of Specialized Labour', in J. Triplett (ed.), *The Measurement of Labour Costs*, Chicago: University of Chicago Press.

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

Rees, H. and A. Shah, 1986, 'An Empirical Analysis of Self-Employed in the U.K.', *Journal of Applied Econometrics* **1**, 95–108.

Schmidt, C. M. and K. F. Zimmermann, 1991, 'Work Characteristics, Firm-Size and Wages', *The Review of Economics and Statistics* **73**, 705–710.

Van den Berg, A. and W. Groot, 1992, 'Union Membership in the Netherlands: A Cross-Sectional Analysis', *Empirical Economics* **17**, 537–564.

Van der Gaag, J. and W. Vijverberg, 1988, 'A Switching Regression Model for Wage Determinants in the Public and Private Sectors of a Developing Country', *The Review of Economics and Statistics* **70**, 224–252.

Van Ophem, H., J. Hartog and W. Vijverberg, 1993, 'Job Complexity and Wages', *International Economic Review* **34**, 853–872.

Willis, R. and S. Rosen, 1979, 'Education and Self-Selection', *Journal of Political Economy* **87**, S7–S36.