Generalized switching regression analysis of private and public sector wage structures in Germany

Dustmann, C.; van Soest, A.H.O.

Publication date:
1995

Contents

Generalized Switching Regression Analysis of Private and Public Sector Wage Structures in Germany

Christian Dustmann
and
Arthur van Soest

March 1995

Abstract

This paper analyzes wage structures in the public and the private sector for Germany. The data contains a rich set of variables on parents' characteristics that we use as instruments. We extend the empirical literature in this field by endogenizing education level and hours worked, and by using life cycle wage differentials in the structural selection equation. We show that these extentions significantly improve the standard model. Moreover, they lead to considerably different parameter estimates. We compute conditional and unconditional wage predictions for the various specifications using model simulations. We find that, on average, potential wages in the private sector exceed those in the public sector. Those actually working in the public sector, would do somewhat better in the private sector, while those working in the private sector would earn much less in the public sector.

*We are grateful to Richard Disney and an anonymous CentER referee for useful comments. Research of the second author is made possible by a fellowship of the Netherlands Royal Academy of Arts and Sciences (KNAW).

†University College, Gower Street, London, WC1E 6BT, UK.

‡Tilburg University, P.O.Box 90153, 5000 Le Tilburg, Netherlands.

1 Introduction

In all industrialized countries and many developing countries the state has an important role as employer. In 1980, total public sector employment accrued to an average of 24.2 percent of non–agricultural employment for OECD - countries. Public sector employment may occur at different levels, like the central government, state and local authorities and financial and nonfinancial public enterprises. Many services supplied by the public sector are not only essential, but also monopolized. As a consequence, employment conditions of public employees are not regulated by market forces, but rather determined by politicians or politically important interest groups (including that of public employees themselves).[1] Wages and working conditions in the public sector differ generally from those in the private sector. Devices which relate wages to productivity are often missing.[2] Moreover, if public sector employment is substantial, wage and employment conditions of the public sector may also affect conditions in the private sector.[3] In recent years, public sector employment conditions and the efficiency of the public sector versus the private sector has become an important policy issue. As a consequence, economists have become increasingly interested in analyzing conditions in public and private sectors. One main research interest is the wage structure in both sectors.

Relying on the assumption that in equilibrium someone with a given vector of human capital characteristics should receive the same wage in both sectors, some studies compare estimated coefficients of wage regressions and predicted wages for the two sectors.[4] But as a consequence of differences in occupations and requirements, sector choice may not be random. This leads to a selection problem in estimating wage equations. The more recent literature usually takes account of this by estimating switching regression models. Such an analysis has been performed for various countries. Van der Gaag and Vijverberg (1988) analyze public-private sector employment for Ivory Coast. Their OLS estimates are considerably different from estimates accounting for selection bias. Similar models are, for example, estimated for the US (Belman and Heywood, 1989), Peru (Stelcner et al., 1989), Italy (Brunello and Rizzi, 1993), and the Netherlands (Theeuwes et al., 1985, van Ophem, 1993, and Hartog and Oosterbeek, 1993). Pedersen et al. (1990) analyze public – private sector earnings in Denmark using panel data. Zweimüller and Winter-Ebmer (1994) analyze wage structures for Austria. Gindling (1991) and Terrell (1993) provide

[1] Whether or not the decision of the policy maker represents in this context the interest of the median voter is an open issue in the literature (see, for instance, Downs (1957) and Reder (1975)).

[2] In many industrialized countries, public sector wages are bargained between public sector unions and the government. See Holmlund (1993) and De Fraja (1993) for theoretical models. The bargaining outcome between public sector unions and employers is often influenced by agreements in the private sector. For instance, Jacobson and Ohlsson (1994) analyze the long-run relationship between government wages and private sector wages in Sweden and find that the private sector is the wage leader.

[3] Jacobsen (1992) provides some evidence for this hypothesis.

[4] See, for instance, Smith (1976), Gunderson (1979), Shapiro and Stelcner (1989) and Peng (1992).

evidence for Costa Rica and Haiti, respectively.

These studies address the following questions: Firstly, what determines the selection in the private and the public sector. Secondly, what are the differences in the pay structure between the two sectors. And thirdly, what are the conditional and unconditional wages differentials between the two sectors. Conclusions vary considerably across the countries analyzed. One may conclude from this that wage structures, incentives and selection mechanisms between public and private sector differ across countries, which is reasonable in view of the divergent institutional settings for private and public sector occupations in different countries.

However, sometimes conclusions differ also between studies for the same country. A reason for this is that results are very sensitive to model assumptions. We show in this paper that some assumptions frequently made in this literature are questionable and that small changes in the specification of the model may lead to different conclusions. At the same time, we provide a first analysis of pay structures in the public and the private sector for Germany. We use data from the German Socio Economic Panel. The data provide background information on parents' social and economic status. This equips us with a rich set of instruments and enables us to allow for endogeneity of some of the usual regressors in wage and choice equations.

A crucial question in this type of models is identification. Non-parametric identification of structural selection and wage equations requires exclusion restrictions on variables in both equations. Many studies use different education measures in the wage equations and in the selection equation, and/or use age in one equation and potential experience in the other. For instance, Belman and Heywood (1989) use continuous measures of education in the wage regression and degrees in the selection equation. Van der Gaag and Vijverberg (1988) and Stelcner et al. (1989) go exactly the opposite way. Identification is in these cases obtained by imposing different restrictions on the way the same information enters selection and wage equations.

Education is a crucial variable in all these studies, not only for identification, but also as a determinant of sectoral choice. Nearly all studies find that a high level of education increases significantly the probability of working in the public sector. We argue, however, that the treatment of education in all these studies is problematic. Since most occupations are not equally available in both sectors, one should expect that human capital in the two sectors is not readily transferable and specific occupations in both sectors require specific types of education. It is therefore likely that individuals meet the sector decision together with part of their educational choice. In other words, an individual who wants some occupation in the public sector, may choose the necessary education simultaneously. We therefore endogenize education. We find that exogeneity of this variable in the selection equation is strongly rejected. This changes the conclusions about the effect of education on sectoral choice. It also questions the use of education as identifying device. We find that the assumption of exogeneity of education in the wage regressions leads to a downward

bias of the educational coefficients.

Following Hartog and Oosterbeek (1993), we include hours worked as a regressor in the wage equation. However, we allow hours of work to be endogenous. Again, we find that exogeneity is strongly rejected.

In all studies using switching regressions, the structural selection equation is estimated with the *current* predicted wage differential as additional regressor. The underlying assumption is that individuals may change sectors at any point in time. However, we have argued already that the decision to join a sector is often not easily reversible. It then should be seen as a long-termed decision, based on a comparison of expected lifetime incomes. We construct a life cycle wage differential and allow the sectoral choice to depend on both the current and the life cycle wage differential. Our empirical results do not unambiguously indicate which of the two measures to choose.

Finally, we compare conditional and unconditional wage differentials between sectors, using simulations. We show that predictions differ between models which do and do not take sector selection into account. The general conclusions about differences in conditional and unconditional wage differentials, appear to be rather robust with respect to the other specification assumptions.

The paper is structured as follows. In the next section we give a short overview on the public and the private sector in Germany and discuss the data. Section 3 describes the econometric model. The results are discussed in section 4. Section 5 gives some conclusions.

2 Institutional Background and Data

Public and Private Sector in Germany

The public sector in Germany distinguishes between two types of employees: civil servants proper (*Beamte*) and blue and white collar public sector employees. Civil servants are usually appointed for lifetime after an initial probation period. Their rights are regulated by law, and they are excluded from collective bargaining. Civil servants include those attached to the essential functions of the state (defense, police force, law and order) and account for about 41% of all state employees. Other state employees have a less special relationship with the state. Their working contracts may well be temporary. Their rights are regulated by negotiated agreements between unions and employers, as in the private sector. They have the right to negotiate wages. Still, although civil servants are not allowed to get actively involved in wage negotiations, civil servants' wage increases are strongly linked to the results of wage negotiations of other public sector employees. The type of work done by civil servants and other state employees is not very different, except for some jobs reserved for civil servants only. Pay scales are the same and apply to all

public sector workers at the federal, state and local level (see Brandes et al, 1990). We shall not distinguish the two categories of public sector employees in the empirical work.

Compared with the private sector, entrance to the public sector is more formalized and based on educational certificates (see Brinkmann, 1976). In the public sector, employees receive special training not only for the occupation as such, but for each different post in their career group. As a consequence of this substantial amount of job-specific education, strong ties develop between employee and employer and entrance into the public sector from private sector occupations is difficult. Moreover, age regulations restrict entrance to civil servant occupations (see Brandes et al, 1988). Blossfeld and Becker (1989) describe the various restrictions on changes between sectors. They conclude that institutional constraints hinder changes between sectors in both directions considerably, particularly with increasing seniority. This is confirmed by the low numbers of transitions: in each of the years 1984 till 1987, less than one percent of all employees changed sector. The public sector in Germany expanded rapidly in the 60's and 70's. The expansion was mainly induced by an extension of the welfare state and the corresponding expansion of social, educational and medical services. Between the 1950 and 1992 the number of employees in the public sector increased from 2.2 Mio to more than 4.95 Mio.[5] The share of public sector employees on the total number of salary earners increased from 12.5 percent in 1965 to 18 percent in 1980 and stabilized since then at about 20 percent. In 1985, 82.9 percent of all public sector employees were employed full-time.

The social security system for civil servants is slightly different from that applying to other public sectors employees. Compared with public sector employees, private sector employees are similarly secured in case of illness. However, in case of invalidity, death and seniority, public sector employees are substantially better protected than private sector employees. For instance, while pensions of civil servants amount to about 80 percent of their latest net income, and pensions of other public employees even to about 105 percent, private sector employees who have paid their contributions to the pension scheme for 45 years, only receive about 72 percent of their latest net income (see Krause, 1981).

Data and Variables

The empirical analysis is based on data from the German Socio-Economic Panel, which includes about 6000 households. 4500 of these have household heads with German nationality, 1500 with foreign nationality. This study is restricted to the first subsample. We combine information from the first (1984) and third (1986) wave of the panel.[6] We

[5] All numbers refer to West Germany only.

[6] We construct most variables from the first wave which provides information about *actual* labor market experience, allowing us to distinguish between cohort and experience effects in the selection equation. From the third wave we obtain information about the family background and parental characteristics when the individual was 15 years old.

restrict the analysis to male individuals who are in dependent employment at the time of the interview. Table A1 in the appendix describes the variables used for the analysis. After excluding all individuals with missing values in relevant variables (see table A2), the sample used for the present analysis reduces to 1428 observations, with 972 employed in the private sector and 456 in the public sector.

Summary statistics are given in table A3. The first two columns give means and, if appropriate, standard errors for the pooled sample of private and public sector employees. The other columns refer to the subsamples of private and public employees. The education level variable is ordered, with values 0 to 5. It is constructed from detailed information about educational background. From table A3, it is clear that public sector employees have a stronger educational background than private sector employees, on average. Public sector employees are somewhat older and slightly more experienced than private sector employees. Private sector employees work, on average, 8 hours per month more than public sector employees, whereas their hourly earnings are lower.

A number of family background variables are used. These reflect the labor market status of the father when the child was aged 15, whether the mother participated in the labor market or not, age of father and mother when the individual was born, and education level of father and mother. We find, for example, that fathers of those in the public sector more often worked in the public sector than fathers of those in the private sector.

3 Econometric Model

We present the main features of the empirical model. Details are provided in the appendix. For each individual, we explain the choice between public and private sector work and the hourly wage rate in the chosen sector. In previous research, an important variable for the sector choice has been the individual's education level. As said above, this is likely to be endogenous. We add an equation to explain this variable, in which education levels of both parents are the main explanatory variables. Following Hartog and Oosterbeek (1993), we allow the wage rate to depend upon hours worked. Since hours worked, according to labor supply theory, may depend upon the wage rate, we allow for endogeneity of hours worked, and add an hours equation. The model thus consists of five equations, explaining education level, the choice between private and public sector (selection), hours of work, and wage rates for the public and private sector.

Education Level

Education levels are coded from 0 to 5, in ascending order (see section 3). An appropriate model is an ordered probit specification (Maddala, 1983, chapter 2):[7]

$$E^* = X_E\,\beta_E + u_E\,; \quad E = j \text{ if } m_{j-1} < E^* \leq m_j \text{ for } (j = 0, ..., 5)\,, \tag{1}$$

where E denotes the educational level attained, and E^* is a latent variable. The vector of explanatory variables X_E contains information concerning the parents (see estimation results in next section for details). Assumptions on the error term u_E are given below. The boundaries satisfy $-\infty = m_{-1} < m_0 < ... < m_4 < m_5 = \infty$. By means of normalization, we assume $m_0 = 0.5$ and $m_4 = 4.5$; m_1, m_2 and m_3 are parameters to be estimated.

Wage Rates

Potential before tax earnings per hour worked are modeled for the private and public sector separately:

$$\ln W_j = X_W\,\beta_j + DE\,\alpha_{Ej} + \alpha_{Hj} \ln H + u_j\,,\; j = 0, 1\,, \tag{2}$$

where $j = 0$ and $j = 1$ denote the public and private sector, respectively. The vector DE contains five dummy variables for the five highest education levels. H denotes hours worked per four weeks. Exogenous variables such as age and experience are included in X_W. Properties of the error terms u_0 and u_1 are given below.

Selection

We model the binary choice between private sector ($S = 0$) and public sector work ($S = 1$). The structural form of this equation is given by

$$S^* = X_S\,\beta_S + DE\delta_E + \delta_c \Delta_c\, lnW + \delta_l \Delta_l\, lnW + u_S\,, \tag{3}$$

where

$$S = \begin{cases} 0 & : \quad S^* < 0 \\ 1 & : \quad S^* \geq 0\,. \end{cases}$$

X_S is a vector of explanatory variables, including, for example, the father's occupational group. $\Delta_c lnW$ and $\Delta_l lnW$ are the current and life cycle log wage differentials

[7] For notational convenience, the index indicating the individual is omitted throughout.

between potential public and private sector wage rates. [8] As argued above, the sector choice decision is not easily reversible. This implies that individuals may not base their choice on the current wage differential, but on the discounted lifetime wage differential: the difference between the discounted (net present) value of past, current and future earnings over the whole working life. We have approximated the life cycle wage differential by an average of wage predictions at various stages of the life cycle (see appendix).

Hours worked

We add a structural form equation for hours worked:

$$lnH = X_H \beta_H + \gamma_S S + \gamma_W lnW + u_H . \tag{4}$$

Here W is the before tax wage rate in the chosen sector, i.e. W_0 if $S = 0$ and W_1 if $S = 1$. The sector dummy S itself is included to reflect institutional differences between the two types of jobs.

Distribution of Error Terms

The vector of error terms $u = (u_E, u_S, u_0, u_1, u_H)'$ is assumed to be independent of all explanatory variables in X_E, X_S, X_W and X_H,[9] and multi-variate normal with mean zero and covariance matrix Σ. By means of normalization, $\Sigma(2,2) = Var(u_S)$ is set equal to one. For practical purposes, u_H is assumed to be independent of the other error terms. Endogeneity of hours worked thus only comes about through the systematic part of the hours equation. $\Sigma(3,4) = \mathrm{Cov}(u_0, u_1)$ is not identified. The other elements of Σ (four variances and five covariances) can be estimated.

Identification

For model identification, we need various exclusion restrictions. In the wage equations, we exclude all parental characteristics, thus assuming that correlation between ability and observed parental characteristics only comes about through selection and education level. In the selection equation, we exclude education levels of the parents to identify endogeneity of education, but we do retain the parents' occupational group variables. To identify the impact of the wage differential, we exclude actual experience. Only those overidentifying restrictions were imposed which were not rejected by the data. An exception is the hours of work equation, which we consider as auxiliary only and do not give it much emphasis.

[8] To be precise, we used the systematic part of the log wage differentials only. See appendix for details.

[9] In the appendix it is explained that endogeneity of experience induced by education level differences is allowed for.

Estimation

The model is estimated by maximum likelihood. The hourly wage rate and hours of work are observed for all individuals in the sample. The likelihood contribution of each individual can be written as the bivariate density of wage rate (in the observed sector) and hours of work, times the conditional probability of the observed education level and sector choice, given wage rate and hours of work. See appendix for details.

4 Estimation Results

We consider a variety of different specifications. Differences between models mainly relate to different zero restrictions on Σ and on coefficients of explanatory variables, the inclusion of the current or the life cycle wage differential in the selection equation. Table 1 presents the models we have estimated and their likelihood values. Likelihood ratio tests based on this table lead to the following conclusions.

[Table 1 about here]

First, exogeneity of education level ($\text{Cov}(u_E, u_0) = \text{Cov}(u_E, u_1) = \text{Cov}(u_E, u_S) = 0$) is strongly rejected. [10] Exogeneity of education level is a maintained assumption in the majority of the empirical literature (see section 1). Second, exogeneity of hours worked ($\gamma_W = \gamma_S = 0$) is also strongly rejected, suggesting that allowing for endogeneity of hours worked is a significant improvement compared to, for example, Hartog and Oosterbeek (1993). Third, we find that neither the current nor the life cycle wage differential are significant in the selection equation. The 'reduced form' model in which wage differentials are not included is not rejected against models with one or both wage differentials. Moreover, once the wage differentials are included, the educational dummies do not add anything.[11] Finally, we find that the selectivity problem is always present: the hypothesis that selection is exogenous in the wage equations ($\delta_c = \delta_l = \text{Cov}(u_0, u_S) = \text{Cov}(u_1, u_S) = 0$) is strongly rejected, no matter whether we allow for endogeneity of education and hours worked or not. We examine these findings in more detail below. We focus the discussion of the parameter estimates on the most general model (model 1). For the selection equation and the wage equations, we also present some alternative estimates.

Hours Worked and Education

Table 2 reports the results of the education and the hours equation in model 1. In the equation explaining the individuals' education level, both the father's and the mother's

[10] Unless stated otherwise, we use a significance level of 5 percent (two-tailed test).

[11] If we include education level itself (treated as a cardinal variable), we find the same result: a t-value of -0.3.

education have a significant positive effect, with that of the father being more important. The occupational category of the father is also important: if the father was a civil servant, this has a strongly imroves education (no father or father is blue collar worker is the reference group). Age is introduced as a polynomial of order five to account for cohort effects in a flexible way. Figure 1 shows the probability of a high education level as a function of age (other variables are set equal to their sample means). It appears that educational achievement is relatively low for those who received their education during the second world war.

[Table 2 about here]

From the hours equation, it appears that hours worked for pay is negatively related to log wages. This is a typical income effect. Hours worked for pay are lower in the public than in the private sector. An explanation is that paid overtime is less common in the public than in the private sector. Over the life cycle, hours worked increase with age until about age 38, and decrease after that. House owners tend to work more hours than renters, which probably reflects a financial constraint. The number of children is positively related to hours worked. This may be related to the fact that in households with (many) children, the wife is more likely to withdraw from the labor force if the husband has a comparative advantage in the market sector. The husband then works more to compensate for the income loss (see Becker, 1981, chapter 2). Unexpectedly, the amount of interest income has a positive and significant effect on hours worked, while a dummy indicating that interest income is received has a negative impact. The latter is significant at the 10 percent level only.

Sectoral Choice

In table 3, we present the results for the selection equation in the most general model (model 1, column 1) and in two alternative specifications. Column 2 is the reduced form model without wage differentials in the selection equation, but allowing for endogeneity (model 4 in table 1) of education and hours worked. Column 3 refers to the prototype model in which education and hours worked are exogenous (model 12). The father's type of occupation substantially influences the sector decision of the offspring. Those whose father was a civil servant have a significantly higher probability to have a public sector job than those whose father had a blue-collar private sector job (the reference group). Those whose father was self-employed or had a white-collar job in the private sector have an intermediate position in this respect.

[Table 3 about here]

In all models presented, we include education as a series of dummy variables. In the general model, education level plays no role for sector choice. In the reduced form model, the education variables are jointly significant. Comparison with the general model suggests that in this case, the negative signs of some of the dummy variables simply

reflect the wage effect. In the first two models, the correlation coefficient between errors in education equation and selection equation is significant. Exogeneity of education is rejected, indicating that educational choices and sector choices are made simultaneously, as suspected in the discussion in section 1. If we do not allow for endogeneity (model 12), education level dummies are jointly significant at the 5 percent level. The parameter estimates in model 12 strongly differ from those in model 4. The positive relation between education level and selection into the public sector is much stronger. This remains the case if the wage differentials are included, and seems completely due to setting $\mathrm{Cov}(u_E, u_S)$ to zero. The positive effect of education found in other empirical studies might thus be due to unobserved factors, correlated with both education level and selection, instead of to education level itself.

This result questions the conclusions drawn in other studies about the effect of education on sectoral choice. Moreover, it suggests that many estimates of wage equations might also be biased, since correction for selectivity is often identified through the use of different specifications of the educational variables in selection and wage equations.

We now turn to the wage differential. All studies use the *current* predicted wage differential as additional regressor. To allow for the possibility that sectoral choice may be a long-termed decision, we include both the current and the expected lifetime wage differential. In model 1, both wage differentials have the expected positive sign, but are insignificant. If one of the wage differentials is removed from the equation, the other remains positive and insignificant (models 2 and 3 in table 1). The current wage differential has a somewhat larger t-value than the life cycle one. Our rather crude way of approximating the life cycle wage differential, ignoring, for example, differences in unemployment risks, or benefit or pension rights, could explain this.

Another explanation could be the lack of identifying restrictions on the selection equation. If, for example, we impose that education level enters linearly in the selection equation or exclude age, both wage differentials become significant. We conclude that our findings in this respect are not robust with respect to the chosen specification. Our results suggest that the choice of identifying restrictions is crucial and could explain differences between the various studies in the literature. For example, Hartog and Oosterbeek (1993) find a positive and strongly significant impact of the current wage differential. They include years of schooling and general education in the selection equation, versus dummies for education levels attained in the wage equations, thus adding additional identifying constraints. Van Ophem (1993), also using Dutch data, finds an insignificant impact of the wage differential. His identifying restrictions are much weaker: only tenure is excluded from the selection equation.

In figure 2, we have sketched the estimated probability of working in the public sector as a function of age and education level for model 1. The public sector probability is highest for the two lowest education levels and for the highest level.[12] The public sector

[12] As expected from table 3, the findings for model 4 are similar. Using model 12, however, would lead

probability gradually increases with age for the younger cohorts, but drops strongly for the oldest cohort. As we shall see below, this cannot be explained from the wage differential. It is probably a cohort effect. Note that those in the oldest age cohort started their career just after the second world war.

Wage Equations

Estimation results for public and private sector wage equations are given in tables 4 and 5. Again, we present three specifications: the general model (model 1), the model in which education level and hours worked are exogenous (model 9), and the OLS results (model 12). In model 1, exogeneity of education in the wage equation is rejected for the private sector, but not for the public sector. Both correlation coefficients are negative. An explanation for this might be incomplete measurement of the education level. A measurement error on education level combined with the positive impact of education on wages may explain both the negative correlation and the downward bias of the estimated impact of education in models 9 and 12, for the same reasons as in a simple linear model.

[Tables 4, 5 about here]

In models 1 and 9, we find that the correlation between errors in selection and wage equation is insignificant for the private sector, but positive and strongly significant for the public sector. [13] This reinforces the positive selection effect of the wage differential. Along the lines of Roy (1951), it indicates that the mean wage of those who have chosen to work in the public sector is larger than the expected public sector wage of an arbitrary individual with the same observed characteristics. We return to this in the next subsection. Positive selection into the public sector, but no selection into the private sector, is also found by van Ophem (1993) and Hartog and Oosterbeek (1993). Van der Gaag and Vijverberg (1988) find a positive selection into both sectors.

The effect of log hours worked on the hourly wage rate is significantly negative for both sectors, with elasticity below one. It increases when endogeneity of hours worked is ignored. Like Hartog and Oosterbeek (1993) we find that the effect is stronger in the public than in the private sector. An explanation could be measurement errors in hours worked, since wage rates are computed as the ratio of monthly earnings and hours worked. For all specifications and both sectors, the wage increases monotonically with education level. In the general model, the impact of education level is stronger than in the models which do not allow for endogeneity of education.

We have included a quadratic function of both age and actual experience. Including age terms apart from experience terms leads to a significant improvement of the likelihood.

---

to the conclusion that the probability of public sector employment increases with education level.

[13] The same results were found with other models allowing for these correlations (models 2 - 8) and remained valid when the systematic part of the equations was changed by, for example, using different age, experience, or education functions.

The age variables may reflect both cohort and life cycle effects. Surprisingly, age plays no role in the public sector, but is significant in the private sector. Differences between the three models are quite small in this respect. In figure 3, we have sketched the expected wage rates in both sectors as a function of education level and age. Experience is replaced by its best linear prediction, given age and education level. Other variables are set equal to their sample means. Public sector wages increase with age, while private sector wages show a much flatter pattern. For young workers, the private sector pays much better than the public sector. For the oldest age group, the difference is negligible.

Wage Differentials

We consider unconditional and conditional wage predictions according to the estimated models. The unconditional (public or private sector) wage prediction is defined as the (average) predicted value of the (public or private) wage rate for an arbitrary individual in the population. The conditional wage prediction is the weighted population average of the wage predictions of all individuals in the sample, where the weights are the estimated sector probabilities (see, e.g., Heckman, 1990, for a discussion of various definitions). To take account of the full structure of the error terms, we have computed the wage predictions using a simulation of the complete model. See appendix for details. The results for all models in table 1 are reported in table 6.

[Table 6 about here]

The first three columns refer to wages in the private sector, the last three columns to wages in the public sector. The first and fourth columns present predicted log wages in the two sectors for an average individual (*unconditional*). According to all models, the private sector pays better than the public sector, on average. For the private sector wage, all models yield similar predictions. For the public sector wage however, differences between the models are much larger. In particular, the models in which selectivity is not taken into account (models 10-12, $\mathrm{Cov}(u_S, u_0) = \mathrm{Cov}(u_S, u_1) = 0$) yield higher public sector wage predictions than the other models, with differences of more than 20 percent.

The other columns refer to *conditional* wage predictions. Column 2 refers to wages in the private sector of an individual who has chosen to work in the private sector. These predictions are similar for all models. All models are able to reproduce the average log wage rate in the private sector rather well. Similarly, column 6 refers to public sector wages of public sector workers. Again, all models reproduce the average log wage in the public sector reasonably well. The most interesting predictions are those in columns 3 and 5, which have no observed sample equivalent. Column 3 refers to the wages that public sector workers could have received when they would have worked in the private sector. Column 5 presents potential public sector wages of private sector workers. These columns reveal much larger differences between the various models than columns 2 and 6. Compared to those in the most general model (model 1), potential public sector wages

of private sector workers in the model not allowing for endogeneity of education level or hours worked (model 9) are about 9 percent lower (column 2). Surprisingly, the most general model and the most restrictive models (models 11 and 12) lead to very similar predictions of potential public sector wages. This is, however, not the case for potential private sector wages of public sector workers (column 5). Here models 1 and 9 yield similar results, while the outcomes of the most restrictive models exceed those of model 1 by more than 35 percent.

A common observation of all models is that predicted *unconditional* wages are higher in the private than in the public sector: an average individual faces higher wage prospects in the private than in the public sector (columns 1 and 4). Those who selected themselves into the private sector (columns 2 and 5) are those with lower potential wages in both sectors. This is due to both observed and unobserved characteristics: from figures 2 and 3 for example, we conclude that age is negatively correlated with selection into the private sector, but positively with wages in both sectors. Accordingly, those who work in the public sector do better in both sectors than the average individual.[14]

Now consider the conditional wage differentials. Those in the private sector would be much worse off in the public sector (columns 2 and 5). According to model 1, the difference would be about 30 percent, on average. The difference is much smaller in models that do not allow for selectivity (in model 12, about 6 percent). The endogenous selection implies that those in the public sector are those with, on average, the smaller wage differential between the two sectors. Thus selection works in the 'right' direction. This result is stable across specifications. For those who actually work in the public sector, the average wage differential in model 1 is less than 7 percent (columns 3 and 6). For models 7 and 9, this wage differential has the opposite sign. If wages were the only criterion for sector choice, then the results for model 1 indicate that the choice of many public sector workers is not rational. However, choosing the public sector is not purely based on wage considerations. For example, job security, entitlements to unemployment benefits, pension rights, and many other non-monetary job characteristics play a role. In section 2, we have discussed some differences in benefit and pension entitlements between the sectors that may compensate the wage disadvantage from choosing a public sector occupation.

[Table 7 about here]

Wage advantages between sectors are often found to vary across educational categories or age groups. Van der Gaag and Vijverberg (1982), for instance, find that unconditional expected wages in the private sector are high for those with low education, but low for those with high education. Hartog and Oosterbeek (1993) find that public sector occupations have higher wage prospects for all educational groups, while van Ophem (1993) reports that the unconditional wage advantage of the public sector diminishes

[14]Many empirical studies ignore the selection effects of observed characteristics and focus only on unobserved characteristics.

with increasing age. Table 7 presents conditional and unconditional log wages for different educational and age groups for model 1. The results speak for themselves.

Tables 6 and 7 are based upon the point estimates. In table 8, we present 90 percent confidence intervals for the wage differentials, taking account of the errors in the parameter estimates (see appendix). The lefthand panel shows that the average unconditional public sector wages are significantly (at the two-sided 10 percent level) lower than average unconditional private sector wages for all education and age groups. Differences are diminishing with age, and lowest for the lowest educational groups. The latter confirms what we saw in figure 3 for a representative individual.

[Table 8 about here]

The middle and righthand panel in table 8 refer to conditional differentials. For all age and education categories, we find that those in the private sector would be significantly worse off in the public sector, on average. According to the righthand panel, those with low education level who chose the public sector do significantly better in the public sector. For public sector workers in the oldest age group, the differential is insignificant. For all other groups, the average wage differential is significantly negative, i.e. in terms of current wages, people would be better off in the private than in the public sector. The average across all groups, is significantly negative also. The size of the conditional differential is modest for public sector workers compared to private sector workers. The precision of the estimated conditional differentials is larger for private sector than for public sector workers. [15]

5 Summary and Conclusion

In this paper we analyze public - private sector choice and wage structures in both sectors for Germany. We estimate generalized switching regression models, extending the standard model in this field in various directions.

We use information on education and occupational status of the parents as additional instruments. We thus avoid ad hoc exclusion restrictions and are able to allow for endogeneity of education. We find that the common conclusion that higher levels of education increase the probability of choosing a public sector job, is only valid if we assume that education is exogenous in the selection equation. Allowing for endogeneity improves the model significantly, and implies that the effect on the sector choice drops to zero in terms of size and significance level. The positive correlation between education and choice is solely driven by unobserved characteristics that affect both in the same direction.

[15] Note that we can only consider the average differential and not the individual variation in the differentials. For example, we cannot predict the fraction of workers for whom the differential is positive, because the covariance between errors in the two wage equations ($\text{Cov}(u_0, u_1)$) is not identified.

To take account of the fact that sectoral choices are long-termed decisions, we use the current as well as an imputed life cycle wage differential in the selection equation. Both have the right sign, but significance levels are low. The results slightly favor the use of the current wage differential, which may be due to the rough way in which we construct the life cycle wage differential. Further research on this based on the use of panel data is on our agenda. Both wage differentials become significant, however, if we include education as a level variable in the selection equation, instead of using dummies. Thus the precision of the estimates of the wage differential coefficients depends strongly on the identifying restrictions, which may well explain different results in the literature.

Not allowing for endogeneity of the educational level results in a substantial downward bias in the estimated impact of education level on the private sector wage. For all educational groups, we find that unconditional wages are initially higher in the private sector, but this advantage levels out with age. Unconditional public sector wages are lower for all educational and age categories than unconditional private sector wages.

On average, conditional wages of public sector employees are somewhat higher in the private sector than in the public sector. Considering separate age and education groups, we find that only those in the public sector with lowest education level are doing significantly better in the public sector. The conditional wage differential is insignificant for the oldest age group; other age and education groups would do significantly better in the private sector. This may be explained by other monetary or non-monetary differences between public and private sector occupations in Germany, not explicitly included in our analysis.

Our analysis suggests that differences in the results of the many studies on public-private sector choice and pay structures may partly be explained by different identifying assumptions. Relaxing the assumption of exogeneity of crucial variables like education, may considerably change some of the conclusions. On the other hand, the results on conditional and unconditional wage differentials are found to be remarkable stable through different specifications, as long as selection is taken into account.

References

Becker, G. (1981): *A Treatise on the Family*, Cambridge: Harvard University Press, 1981.

Belman, D. and J. S. Heywood (1989): "Government Wage Differentials: A Sample Selection Approach," *Applied Economics*, 21, 427-438.

Blossfeld, H.P., R. Becker (1989): "Arbeitsmarktprozesse zwischen Oeffentlichem und Privatwirtschaftlichem Sektor," *MittAB*, 2, 233-247.

Brandes, W. et al (1988): *Der Staat als Arbeitgeber*, dfg-research report, Universitaet-GH Paderborn.

Brandes, W. et al (1990): *Der Staat als Arbeitgeber. Daten und Analysen zum öffentlichen Dienst in der Bundesrebublik.*, Campus-Verlag, Frankfurt.

Brinkmann, G. (1976): *Aufgaben und Qualifikation der Oeffentlichen Verwaltung*, Goettingen: Otto Schwarz.

Brunello, G. and D. Rizzi (1993): "I Differenziali Retributivi Nei Settori Pubblico e Privato in Italia: Un'Analisi Cross Section," *Politica Economica*, 9, 339-366.

Downs, A. (1957): *An Economic Theory of Democracy*, New York: Harper & Row.

De Fraja, G. (1993): "Unions and Wages in Public and Private Firms: A Game-Theoretic Analysis," *Oxford Economic Papers*, 45, 457-469.

Gindling, T.H. (1991): "Labor Market Segmentation and the Determination of Wages in the Public, Private-Formal, and Informal Sectors in San Jose, Costa Rica," *Economic Development and Cultural Change*, 39, 585-605.

Greene, W. H. (1993): *Econometric Analysis*, MacMillan, MacMillan Publishing Company, New York.

Gunderson, M. (1979): "Earnings Differentials between the Public and Private Sectors," *Canadian Journal of Economics*, 12, 228-242.

Hartog J. and H. Oosterbeek (1993): "Public and Private Sector Wages in the Netherlands," *European Economic Review*, 37, 97-114.

Heckman, J. (1990): "Varieties of Selection Bias," *American Economic Review Papers and Proceedings*, 80, 313-318.

Holmlund, B. (1993): "Wage Setting in Private and Public Sectors in a Model with Endogenous Government Behavior," *European Journal of Political Economy*, 9 ,149-162.

Hundley, G. (1991): "Public- and Private Sector Occupational Pay Structures," *Industrial Relations*, 30, 416-434.

JACOBSEN, J.P. (1992): "Spill over Effects from Government Employment," *Economic Letters*, 39, 101-104 .

JACOBSON, T. AND H. OHLSSON (1994): "Long-Run Relations between Private and Public Sector Wages in Sweden," *Empirical Economics*, 19, 343-360.

KRAUSE, P. (1981): "Landesbericht Bundesrepublik Deutschland", in *Soziale Sicherung im Öffentlichen Dienst*, H.F. Zacher, M. Bullinger and G. Igl (eds), Duncker and Humbold, Berlin.

MADDALA, G. (1993): *Limited Dependent and Qualitative Variables in Econometrics* , Cambridge: Cambridge University Press.

PENG, Y. (1992): "Wage Determination in Rural and Urban China: A Comparison of Public and Private Industrial Sectors," *American Sociological Review*, 57, 198-213.

PEDERSEN, P., J. SCHMIDT-SOERENSEN, N. SMITH, AND N. WESTERGAARD-NIELSEN(1990): "Wage Differentials between the Public and Private Sectors," *Journal of Public Economics*, 41, 125-145.

ROY, A. D. (1951): "Some Thoughts on the Distribution of earnings," *Oxford Economic Papers*, 3, 135-146.

REDER, M. (1975): "The Theory of Employment and Wages in the Public Sector," in *Labor in the Public and Private Nonprofit Sectors*, ed. D. Hamermesh, Princeton: Princeton University Press, 1-48.

SHAPIRO, D. M. AND M. STELCNER (1989): "Canadian Public-Private Sector Earnings Differentials, 1970-1980," *Industrial Relations*, 28, 72-81.

SMITH, S. (1976): "Pay Differentials between Federal Government and Private Sector Workers," *Industrial and Labor Relations Review*, 29, 233-257.

STELCNER, M., VAN DER GAAG, J. AND W. VIJVERBERG (1989): "A Switching Regression Model of Public-Private Sector Wage Differentials in Peru: 1985-1986," *Journal of Human Resources*, 24, 545-559.

TERRELL, K. (1993): "Public - Private Wage Differentials in Haiti," *Journal of Development Economics*, 42, 293-314.

THEEUWES, J., KOOPMANS, C., VAN OPSTAL, R. AND H. VAN REIJN (1985): "Estimation of Human Capital Accumulation Parameters for the Netherlands," *European Economic Review*, 29, 233-257.

VAN OPHEM (1993): "A Modified Switching Regression Model for Earnings Differentials between the Public and Private Sectors in the Netherlands," *Review of Economics and Statistics*, 215-224.

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," *Review of Economics and Statistics*, 70, 244-252.

Zweimüller, J. and R. Winter-Ebmer (1994): "Gender Wage Differentials in Private and Public Sector Jobs," *Journal of Population Economics*, 7, 271-285.

Table 1: Model specifications and Likelihoods

Model	No Restrictions	Specification	Log Likelihood
1	0	—	-1679.08
2	1	$\delta_{Wl} = 0$	-1679.34
3	1	$\delta_{Wc} = 0$	-1680.62
4	2	$\delta_{Wl} = \delta_{Wc} = 0$	-1680.63
5	5	$\alpha_E = 0$	-1680.04
6	7	$\alpha_E = 0, \delta_{Wl} = \delta_{Wc} = 0$	-1697.17
7	3	$\Sigma(1,2) = \Sigma(1,3) = \Sigma(1,4) = 0$	-1688.14
8	2	$\gamma_S = \gamma_E = 0$	-1729.62
9	5	$\Sigma(1,2) = \Sigma(1,3) = \Sigma(1,4) = 0,$ $\gamma_S = \gamma_E = 0$	-1738.42
10	2	$\Sigma(2,3) = \Sigma(2,4) = 0$	-1683.26
11	7	$\Sigma(1,2) = \Sigma(1,3) = \Sigma(1,4) = 0,$ $\gamma_S = \gamma_E = 0,\ \Sigma(2,3) = \Sigma(2,4) = 0$	-1743.56
12	9	$\Sigma(1,2) = \Sigma(1,3) = \Sigma(1,4) = 0,$ $\gamma_S = \gamma_E = 0,\ \Sigma(2,3) = \Sigma(2,4) = 0,$ $\delta_{Wl} = \delta_{Wc} = 0$	-1744.61

Table 2: Education Level and Hours Worked, Model 1

Education Level			log Hours worked		
Variable	estimate	t-value	Variable	estimate	t-value
constant ed	9.0608	0.71	constant	5.2657	101.17
EDLEV_F	0.4212	9.15	M	0.0003	0.03
EDLEV_M	0.2210	3.97	HEAD	0.0315	1.80
AGE/10	-15.8435	-0.92	DINTINC	-0.0292	-1.67
(AGE/10)**2	10.4728	1.18	LINTINC+1	0.0063	2.11
(AGE/10)**3	-3.0001	-1.36	HOWN	0.0192	2.13
(AGE/10)**4	0.3885	1.46	CHILD	0.0062	0.75
(AGE/10)**5	-0.0187	-1.49	AGE/10	0.0467	1.54
AGE_F_BIRTH	0.0135	1.56	(AGE/10)**2	-0.0061	-1.68
AGE_M_BIRTH	-0.0026	-0.25	PUBLIC	-0.0461	-5.11
<2_PARENTS	-0.1956	-1.36	LHEARN	-0.0880	-5.13
F_SELF	0.3409	2.73	sigma hours	0.1223	76.21
F_CIVIL	0.8536	5.85			
F_WHITE	0.5613	4.56			
M_EMPL	-0.0718	-0.87			
sigma edl	1.3177	41.13			
bound 2	2.9358	52.33			
bound 3	3.7846	72.48			
bound 4	4.1131	92.95			

Note: bound 1 = 0.5; bound 5 = 4.5

Table 3: Selection equation

	model 1		model 4		model 12	
Variables	parameter	t-value	parameter	t-value	parameter	t-value
constant pp	2.1047	0.13	2.2335	0.13	6.7201	0.44
AGE/10	-7.4239	-0.33	-9.1658	-0.41	-14.9415	-0.70
(AGE/10)**2	6.8308	0.58	7.3072	0.61	9.8622	0.86
(AGE/10)**3	-2.3831	-0.77	-2.4004	-0.78	-2.9507	-0.98
(AGE/10)**4	0.3639	0.94	0.3550	0.91	0.4131	1.09
(AGE/10)**5	-0.0205	-1.07	-0.0195	-1.03	-0.0220	-1.18
F_SELF	0.2545	2.42	0.2590	2.44	0.1233	1.14
F_CIVIL	0.5667	4.23	0.5841	4.33	0.4047	3.29
F_WHITE	0.2615	2.30	0.2337	2.05	0.0546	0.50
ED_LEVEL						
ED_LEVEL1	0.2102	0.29	-0.3882	-1.97	0.0714	0.52
ED_LEVEL2	0.0595	0.08	-0.2111	-0.73	0.5670	3.74
ED_LEVEL3	-0.2383	-0.28	-0.6218	-1.87	0.2305	1.12
ED_LEVEL4	0.4555	0.35	-0.7544	-2.08	0.1656	0.81
ED_LEVEL5	-0.1916	-0.21	0.0142	0.03	1.0952	6.01
wage diff curr	3.5508	1.24	0.0000	----	0.0000	----
wage diff lifec	3.2579	0.82	0.0000	----	0.0000	----
rho ed pp	0.3219	3.41	0.3121	3.11	0.0000	----

Table 4: Log Private Sector Wage Rate

	model 1		model 9		model 12	
Variables	parameter	t-value	parameter	t-value	parameter	t-value
const wpriv	3.0837	5.93	4.1821	13.92	4.1204	13.85
ED_LEVEL1	0.2218	5.09	0.1515	4.30	0.1537	4.43
ED_LEVEL2	0.4087	6.47	0.2759	6.18	0.2953	7.71
ED_LEVEL3	0.5148	7.31	0.3858	7.36	0.3960	7.75
ED_LEVEL4	0.6887	8.93	0.5581	9.32	0.5691	9.64
ED_LEVEL5	0.8041	8.82	0.5998	7.99	0.6432	13.22
M	0.0981	4.69	0.1012	4.86	0.0778	2.88
SIZE	0.0351	2.40	0.0361	2.44	0.0304	1.63
EXP/10	0.1699	3.16	0.1806	3.27	0.1807	3.31
(EXP/10)^2	-0.0441	-3.61	-0.0471	-3.88	-0.0469	-3.91
AGE/10	0.3605	3.73	0.3730	3.98	0.4120	4.69
(AGE/10)^2	-0.0341	-3.12	-0.0364	-3.52	-0.0404	-4.14
LHWMONTH	-0.2912	-2.90	-0.4974	-9.20	-0.4943	-9.10
sigma w	0.2689	51.97	0.2678	26.61	0.2645	68.39
rho w pp	-0.0290	-0.09	-0.2222	-0.82	0.0000	-----
rho w ed	-0.3064	-2.33	0.0000	----	0.0000	-----

Table 5: Log Public Sector Wage Rate

	model1		model 9		model 12	
Variables	parameter	t-value	parameter	t-value	parameter	t-value
const wpub	4.4153	8.29	5.2295	12.22	5.5145	12.97
ED_LEVEL1	0.1276	2.09	0.0706	1.48	0.0617	1.37
ED_LEVEL2	0.3573	4.45	0.2773	5.29	0.1877	3.94
ED_LEVEL3	0.4460	4.23	0.3522	4.44	0.3007	4.06
ED_LEVEL4	0.4901	3.91	0.3711	4.42	0.3485	4.56
ED_LEVEL5	0.8149	7.33	0.6852	11.30	0.5187	10.56
M	0.0837	3.30	0.0841	3.31	0.1319	4.61
SIZE	0.0324	1.74	0.0308	1.66	0.0496	2.12
EXP/10	0.2702	4.50	0.2810	4.57	0.2362	4.25
(EXP/10)^2	-0.0681	-4.89	-0.0721	-5.06	-0.0587	-4.79
AGE/10	0.0994	0.71	0.0947	0.67	0.0727	0.57
(AGE/10)^2	0.0070	0.43	0.0081	0.48	0.0056	0.39
LHWMONTH	-0.5441	-6.30	-0.6951	-11.14	-0.6711	-10.70
sigma w	0.2837	15.72	0.2924	16.11	0.2191	36.68
rho w pp	0.6963	8.74	0.8042	14.10	0.0000	----
rho w ed	-0.1326	-1.28	0.0000	----	0.0000	----

Table 6: simulated unconditional and conditional wages; sample means

model	log wage private sector			log wage public sector		
	all	private	public	all	private	public
1	2.971	2.927	3.052	2.664	2.487	2.986
2	2.993	2.938	3.128	2.673	2.532	3.018
3	2.991	2.945	3.086	2.684	2.522	3.024
4	3.003	2.927	3.142	2.674	2.450	2.991
5	2.989	2.935	3.113	2.671	2.523	3.014
6	3.035	2.955	3.233	2.707	2.579	3.023
7	2.942	2.927	2.977	2.641	2.481	2.996
8	2.992	2.952	3.078	2.670	2.513	3.009
9	2.944	2.936	2.961	2.634	2.485	2.986
10	2.989	2.940	3.084	2.914	2.864	3.012
11	2.974	2.935	3.047	2.898	2.857	2.976
12	2.973	2.939	3.058	2.900	2.871	2.972

Note: Definition of models: see table 1.
all: all workers (unconditional);
private/public: workers in private/public sector (conditional)

Table 7: simulated unconditional and conditional wages,
by education level (ED_LEVEL) and AGE; model 1

	log wage private sector			log wage public sector		
	all	private	public	all	private	public
ED_LEVEL=1	2.694	2.683	2.724	2.455	2.335	2.801
ED_LEVEL=2	2.891	2.881	2.918	2.558	2.436	2.867
ED_LEVEL=3	3.044	3.021	3.069	2.765	2.562	2.985
ED_LEVEL>=4	3.297	3.244	3.349	2.998	2.759	3.233
AGE<30	2.743	2.715	2.831	2.363	2.249	2.728
29<AGE<40	3.055	3.022	3.120	2.668	2.516	2.971
39<AGE<50	3.094	3.059	3.155	2.790	2.619	3.088
AGE>49	2.915	2.915	3.015	2.821	2.614	3.035

Table 8: 90 percent confidence intervals unconditional and conditional
wage differentials; model 1.

[(public sector wage rate/private sector wage rate)-1]100%

	all workers		private sector		public sector	
ED_LEVEL=1	-22.2	-19.3	-30.0	-26.6	0.6	9.9
ED_LEVEL=2	-28.3	-26.8	-35.6	-34.1	-6.1	-2.5
ED_LEVEL=3	-25.3	-22.6	-37.4	-34.6	-8.9	-4.9
ED_LEVEL=4	-26.1	-23.3	-39.8	-37.1	-10.9	-6.4
AGE<30	-32.9	-30.4	-38.4	-36.0	-12.4	-7.4
29<AGE	-31.5	-28.9	-39.5	-37.0	-13.8	-9.5
39<AGE	-25.7	-24.0	-34.8	-32.9	-7.2	-3.7
AGE>49	-15.7	-12.9	-28.7	-25.6	-0.0	4.5
all	-26.3	-25.3	-35.4	-34.4	-6.3	-4.3

Appendix: Data description

Table A1: Explanation of Variables

Variable	
AGE	Age of Individual
EXP a)	Actual Labor Market Experience of Individual
ED_LEVEL0	Dummy; 1 if basic or intermediate schooling (Haupt/Realschule)
ED_LEVEL1	Dummy; 1 if basic schooling and apprenticeship
ED_LEVEL2	Dummy; 1 if intermediate schooling and apprenticeship
ED_LEVEL3	Dummy; 1 if high school (Gymnasium, Fachhochschule)/high school and apprenticeship
ED_LEVEL4	Dummy; 1 if engineering school or higher specific school
ED_LEVEL5	Dummy; 1 if university
ED_LEVEL	Ordered variable on education, calculated from information about degree
TEARN	Total Monthly Gross Earnings
HWMONTH b)	Hours worked per month for pay
LHWMONTH	Log of Hours worked per month for pay
HEARN	Hourly earnings
LHEARN	Log of Hourly earnings
INTINC	Interest income per month
LINTINC+1	Log(Interest income per month+1)
DINTINC	Dummy; 1 if interest income
M	Dummy; 1 if married
BLUE	Dummy; 1 if blue collar
WHITE	Dummy; 1 if white collar
CIVIL	Dummy; 1 if civil servant
HOWN	Dummy; 1 if individual house owner
HEAD	Dummy; 1 if individual head of household
PUBLIC	Dummy; 1 if employed in public sector
SIZE	Dummy; 1 if town larger than 100 000 inhabitants
CHILD	Dummy; 1 if child younger than 16 in household
F_ERW	Dummy; 1 if father employed when individual was 15
F_SELF	Dummy; 1 if father self employed
F_WHITE	Dummy; 1 if father white collar
F_BLUE	Dummy; 1 if father blue collar
F_CIVIL	Dummy; 1 if father civil servant
M_EMPL	Dummy; 1 if mother employed when individual was 15
<2_PARENTS	Dummy; 1 if grown up with father or mother only
EDLEV_F	Ordered education variable, father (constructed from degree information)
EDLEV_M	Ordered education variable, mother (constructed from degree information)
AGE_F_BIRTH	Age of father when individual was born
AGE_M_BIRTH	Age of mother when individual was born

a) Constructed from a biographical scheme; after the age of 15.

b) Two variables on hours worked are available: normal hours worked, and actual hours worked including overtime. Furthermore, the individual was asked whether overtime work was paid for. The variable on hours worked used here measures *earnings-effective* hours worked and was constructed as follows: If the individual reported overtime hours and overtime work is paid for, HWMONTH is constructed on the basis of this measure. In other cases, HWMONTH is constructed on the basis of normal hours worked.

Table A2: Missing Information

	Number Observation:	Percent Public:
Males in dependent employment	1809	29.96
No Missings in:		
AGE_F_BIRTH, AGE_M_BIRTH	1749	30.36
EDLEV_F, EDLEV_M	1722	30.48
F_ERW, M_EMPL	1564	31.39
ED_LEVEL	1557	31.34
TEARN, HWMONTH	1428	31.93

Table A3: Descriptive Statistics

	Pooled		Private		Public	
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
PUBLIC	0.319		0		1	
AGE	39.711	11.302	38.895	11.426	41.451	10.841
EXP	19.668	11.718	19.387	11.796	20.267	11.541
EDLEV	3.263	2.144	2.902	1.884	4.032	2.443
EDLEV_F	2.210	1.688	2.131	1.652	2.377	1.752
EDLEV_M	1.475	1.171	1.400	1.102	1.635	1.293
SIZE	0.556	0.496	0.549	0.497	0.572	0.495
AGE_F_BIRTH	31.622	7.052	31.513	7.089	31.855	6.974
AGE_M_BIRTH	28.397	5.981	28.344	6.059	28.508	5.815
HWMONTH	167.924	23.956	170.543	24.752	162.342	21.133
TEARN	3366.385	1279.640	3341.042	1310.932	3420.405	1209.868
HEARN	20.294	7.892	19.750	7.640	21.455	8.294
LHEARN	2.945	0.355	2.919	0.353	3.001	0.351
INTINC	193.441	701.630	176.326	650.588	229.925	799.187
DINTINC	0.247		0.235		0.274	
HOWN	0.417		0.389		0.476	
M	0.764		0.757		0.780	
BLUE	0.462		0.584		0.203	
WHITE	0.367		0.410		0.276	
CIVIL	0.166		0.002		0.517	
CHILD	0.455		0.463		0.438	
HEAD	0.897		0.871		0.953	
F_SELF	0.143		0.134		0.162	
F_CIVIL	0.103		0.075		0.164	
F_WHITE	0.141		0.134		0.155	
M_EMPL	0.412		0.422		0.390	
<2_PARENTS	0.069		0.068		0.070	
NOBS	1428		972		448	

Appendix: Model Specification and Likelihood

Model

We address the issue of endogeneity of experience by rewriting some of the equations and by adding an equation explaining experience. The equations (1) (education level), (3) (selection) and (4) (hours worked) remain unchanged.

Experience: if we allow for endogeneity of education level, it is also natural to allow for endogeneity of experience, since experience will be negatively affected by education level. We add a regression equation to explain experience (Exp):

$$Exp = X_{Exp}\beta_{Exp} + DE\tau_E + u_{Exp}\,. \tag{5}$$

DE is the vector of educational dummies (see section 4). X_{Exp} contains age and other exogenous variables. Assumptions about u_{Exp} determine whether this equation should explicitly be taken into account (see below).

Wage Rates: to take account of experience explicitly, we change the notation of (2):

$$\ln W_j = g_W(X_W, E, Exp; \beta_j) + \alpha_{Hj}\ln H + u_j\,, j = 0, 1\,. \tag{6}$$

X_W now excludes experience. g_W is a given function (for example a linear combination of education dummies, Exp and Exp^2).

Wage differentials in selection equation: in (3), we include both the current and the life cycle wage differential. There are two ways to define the current wage differential:

$$\Delta_{c1}\ln W = g_W(X_W, E, Exp; \beta_1) - g_W(X_W, E, Exp; \beta_0) + (\alpha_{H1} - \alpha_{H0})\ln 160\,, \tag{7}$$

$$\Delta_{c2}\ln W = \Delta_{c1}\ln W + u_1 - u_0\,. \tag{8}$$

$\Delta_{c1}\ln W$ does not depend upon u_1 or u_0. It is the difference between predicted log wages for a standard working month, not taking account of potential knowledge of u_1 and u_0. We denote these log wages by $\ln WP_1$ and $\ln WP_0$.

$\Delta_{c2}\ln W$ is preferable from an economic point of view, since the individual will also take account of u_1-u_0. Assume for the moment that δ_l=0 (no life cycle wage differential in (3)). Then, if we use $\Delta_{c2}\ln W$, we can write the selection equation (3) in terms of $\Delta_{c1}\ln W$:

$$S^* = X_S\,\beta_S + DE\delta_E + \delta_c\Delta_{c1}\,lnW + v_S\,, \tag{9}$$

where $v_S = u_S + \delta_c(u_1 - u_0)$. This shows that (7) and (8) lead to observational equivalent models, unless specific assumptions on the covariances of u_S with u_E, u_1 and u_0 are made.

For the life cycle wage differential, again, two definitions can be suggested.

$$\Delta_{L1}\ln W = \ln \mathrm{NPV}(WP_1) - \ln \mathrm{NPV}(WP_0)\,, \tag{10}$$

$$\Delta_{L2}\ln W = \Delta_{L1}\ln W + \mathrm{NPV}(u_1 - u_0)\,, \tag{11}$$

where $\mathrm{NPV}(WP_0)$ and $\mathrm{NPV}(WP_1)$ are the net present values of predicted wages in the two sectors, for some given discount rate. The pattern of log wages as a function of age and experience is taken into account. In fact, we approximate NPV by the average value of predicted wages at five equidistant points of time during an individual's working life. This is a rather rough approximation. We assume that individuals do not change sector or lose their job, and that cohort effects do not affect wages. Working with the life cycle wage differential requires an assumption about the time persistence of the error terms. If the u_1 and u_0 of one individual are uncorrelated over time, they approximately average out, and (10) is a reasonable approximation. If u_1 and u_0 remain constant over time, (11) is to be used. Since, however, $\mathrm{NPV}(u_1 - u_0)$ will be a linear function of u_1 and u_0, (10) and (11) lead to observational equivalent models, as in the case of the current wage differential. We therefore work with (10).

Distribution of error terms: To the assumptions in section 4 on $u = (u_E, v_S, u_0, u_1, u_H)'$, we add the assumption that u_{Exp} is independent of u and of all exogenous variables.

Likelihood Contributions

We use the selection equation given in (9). The case with the life cycle wage differential is similar. Consider someone working in the public sector ($S = 1$). E, Exp, W_1 (=W) and H are observed, W_0, E^* and S^* are not observed. Denote, for given parameter values, the 'residuals' of (5), (6) and (4) by e_{Exp}, e_W and e_H. For given parameters, these are the realizations of u_{Exp}, u_W and u_H. Denote the density of x conditional on y by $f_{x|y}$. The likelihood contribution can be written as

$$L = \int_{R(E)} \int_{R(S)} f_{E^*,Exp,S^*,lnW_1,lnH|X}(E^*, Exp, S^*, lnW_1, lnH)\, dS^* dE^*\,. \tag{12}$$

Here R(E) is the region of possible values of E^* for the observed value of E (for given parameters m_j). R(S) is defined likewise ($R(S) = [0, \infty)$). X contains all exogenous variables. (12) can be rewritten in terms of residuals:

$$L = \int_{RU(E)} \int_{RU(S)} \sqrt{|Det|}\, f_{u_E,v_S,u_1,u_H,u_{Exp}}(u_E, v_S, e_1, e_H, e_{Exp}) dv_S du_E\,. \tag{13}$$

Here RU(E) and RU(S) are the ranges of the errors u_E and v_S corresponding R(E) and R(S). Det is the Jacobian term:

$$Det = 1 - \alpha_{H1}\gamma_W \,. \tag{14}$$

The integral in (13) can be written as a density times a bivariate conditional probability:

$$L = \sqrt{|Det|} f_{u_1,u_H,u_{Exp}}(e_1, e_H, e_{Exp}) \int_{RU(E)} \int_{RU(S)} f_{u_E,v_S|u_1,u_H,u_{Exp}}(u_E, v_S) dv_S du_E \,. \tag{15}$$

The integral is a bivariate cumulative normal probability (see, e.g., Greene, 1993, p. 76). Since u_{Exp} is independent of the other errors, L can be written as follows.

$$L = \sqrt{|Det|} f_{u_{Exp}}(e_{Exp}) f_{u_1,u_H}(e_1, e_H) \int_{RU(E)} \int_{RU(S)} f_{u_E,v_S|u_1,u_H}(u_E, v_S) dv_S du_E \,. \tag{16}$$

The factor related to u_{Exp} is the only factor containing the parameters in (5) and contains no other parameters. As a consequence, (5) can be ignored while estimating the rest of the model; the model can be estimated as if Exp was exogenous. The intuition is that endogeneity of Exp is only caused by endogeneity of E, and this is taken into account.

Simulations

To compute the wage differentials in section 6, we have used the model to simulate all endogenous variables, taking exogenous variables as given. We have treated experience as an endogenous variable. For this purpose, we specify and estimate (5). We use a linear equation, and assume normality of u_{Exp}. The estimated relation is

$$Exp = -15.22 - 1.10DE1 - 3.03DE2 - 4.40DE3 - 6.61DE4 - 8.62DE5 + 0.93AGE + 0.69M + u_{Exp} \tag{17}$$

DE1-DE5 are the education level dummies (0 excluded). M is 1 if married, 0 otherwise. All variables are significant at the 5 percent level. The estimate of $\sigma(u_{Exp})$ is 4.22; the R^2 is 0.87.

The simulation now works as follows. For each individual, we draw 10 values of all error terms from the estimated sixvariate normal distribution of error terms. Using actual values of exogenous variables, we recursively compute E from (1), Exp from (17), and S from (3). For each sector separately, we solve (2) and (5) to compute $\ln W_0$ and $\ln W_1$. We then average over draws per individual and over individuals in the (sub)population. In this way we construct tables 6 and 7, based on the parameter estimates. For table 8, we repeated the exercise 100 times, for 100 draws of the parameter vector from its estimated asymptotic distribution. The confidence intervals are bounded by the 5 and 95 percentiles of the 100 outcomes.

6 Figures

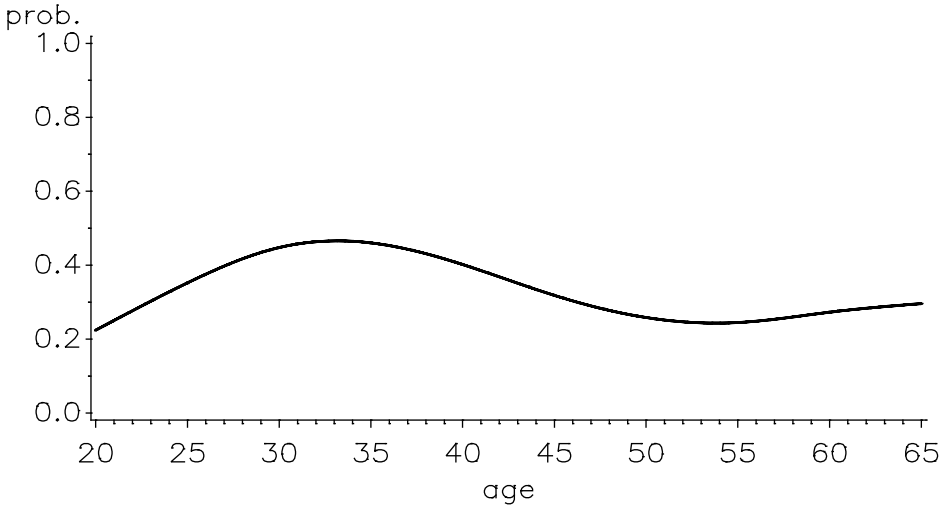

Fig. 1: Probability Education Level >= 2

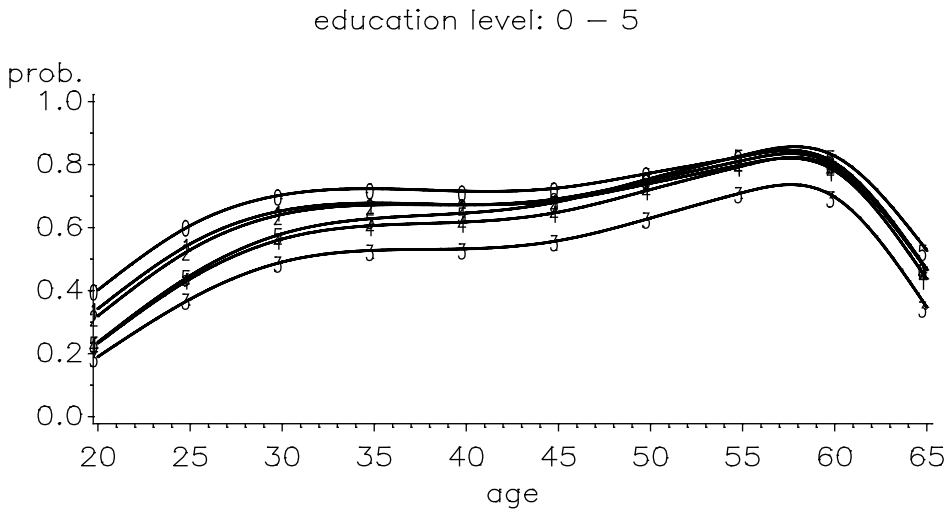

Fig. 2: Probability Public Sector

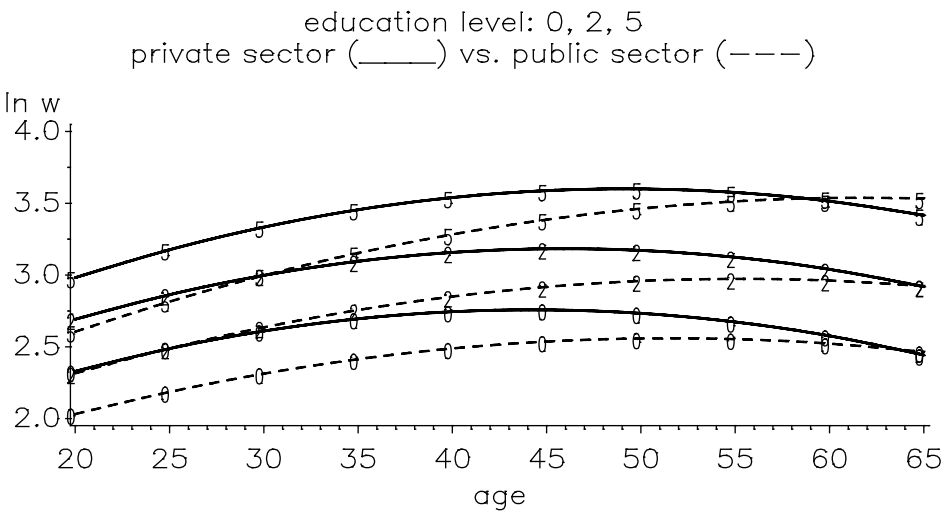

Fig. 3: Log Wage Rates