DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

Firm Specific Gender and Ethicity Pay Differentials in Britain

Stephen Pudney and Nikolaos Theodoropoulos

Discussion Paper 2006-09

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: ++357-2-892430, Fax: ++357-2-892432
Web site: http://www.econ.ucy.ac.cy

FIRM-SPECIFIC GENDER AND ETHNICITY PAY DIFFERENTIALS IN BRITAIN

Stephen Pudney* and Nikolaos Theodoropoulos**
*** Institute for Social and Economic Research (*ISER*), University of Essex, UK**
**** Department of Economics, University of Cyprus, Cyprus and Centre for Research and Analysis of Migration, Department of Economics, University College London, UK**

ABSTRACT Using matched employer-employee data we examine firm-specific gender and ethnicity pay differentials in Britain. We estimate an econometric earnings model using the partially-observed pay variable provided in the data and test the normality assumption that underlies the usual interval regression technique. We then estimate alternative specifications allowing for firm-specific random effects, using a semi-parametric finite mixture estimator. The empirical estimation reveals a 22% (13%) weekly (hourly) gender pay gap and a 28% (19%) weekly (hourly) pay race gap. Strikingly, although significant and sizeable the firm-specific effects are not correlated with other variables that may act as indirect indicators of pay differentials.

KEYWORDS: matched employer-employee data, pay differentials, random effects, semi-parametric finite mixture estimator

JEL CLASSIFICATION: C31, C34, C35, J7

Address for correspondence: Nikolaos Theodoropoulos, Department of Economics, University of Cyprus, P.O. Box 20537, CY-1678, Cyprus. E-mail: n.theodoropoulos@ucy.ac.cy

Acknowledgments This research is based on data from the 1998 Workplace Employee Relations Survey (WERS98). The survey was jointly sponsored by the Department of Trade and Industry, the Advisory, Conciliation and Arbitration Service (ACAS), the Economics and Social Research Council (ESRC) and the Policy Studies Institute (PSI). As with other surveys in the series, the National Centre for Social Research (SCPR), was commissioned to conduct the survey fieldwork on behalf of the sponsors. WERS98 is deposited and available from the Data Archive at Essex University. Neither the sponsors nor the Data Archive has any responsibility for the analysis of interpretation of the material contained in this paper. We are grateful to Sarah Brown, Martin Hoskins and John Sessions for useful comments and suggestions. We would like also to thank seminar and conference participants at Essex, Keele, Leicester, at the XIIth Summer School of the European Economic Association and the 4th IZA Summer School in Labor Economics, at the 2002 ASSET (Cyprus), at the 2005 WPEG (York) and at the 2006 EALE conference in Prague. We would also like to thank the Royal Economic Society conference grant scheme for financial support.

1. Introduction

Ethnic and gender pay differentials have been heavily researched by labour economists. A common consensus is that both gender (Blau and Kahn, 2003; O'Neil and Polachek, 1993) and ethnic pay differentials (Chandra, 2003; Smith and Welch, 1989) have decreased over the last three decades. This decline is the result of many changes, including equal pay, anti-discrimination and affirmative action legislations (see Zabalza and Tzannatos, 1985 for the UK and Donohue and Heckman, 1991 for the US).

There is evidence that gender pay differentials have also been affected by welfare reforms, incomes policies and changes in industrial structure, education and labour force participation and job mobility, fertility and family structure (Borooah and Lee, 1988; Goldin, 2004; Goldin and Katz, 2000). Changes in the comparator male earnings distribution have also been important (Gosling *et al*., 2000). For ethnic minorities, convergence in years and quality of schooling has been found for recent cohorts in the US and in the UK (see Card and Krueger, 1992 and Dustmann and Theodoropoulos, 2006 respectively). Despite these relative gains, gender and the ethnicity pay gaps are remarkably persistent (Blau and Kahn, 2006; Darity and Nembhard, 2000).

Nevertheless, there remains the difficulty that any differential reward to observed human capital variables (such as education) can be interpreted as either discrimination or as the result of differences in the unobserved correlates of these variables. Neal and Johnson (1996) have argued that controlling for pre-market skill levels largely accounts for the US black-white wage gap (see Lang and Manove, 2006 for a contrary view). Apart from human capital, other attributes such as non-cognitive skills (Heckman and Rubinstein, 2001) and behavioural characteristics (Bowles *et al*., 2001) are strong predictors of earnings. Thus,

drawing a firm conclusion on discrimination is difficult because between-group productivity differentials cannot be observed (Altonji and Blank, 1999). For this reason, the pay gap is better labelled as "pay disadvantage" rather than as "pay discrimination" (Pudney and Shields 2000).

The above studies have all used datasets giving information only on the employee (individual or household data) or on the employer (plant or firm level data) despite the fact that labour market outcomes involve the matching of employees and employers. Matched employer-employee data make possible much more convincing evidence on discrimination (Hellerstein and Neumark, 2005). However, it is only recently that such datasets have become widely available (Hamermesh, 1999). Studies using linked employee-employer datasets (Abowd *et al*., 1999; Hellerstein *et al*., 1999; Meng, 2004) have found strong firm effects in explaining wage differentials, without linking these effects to gender and ethnicity.

In this paper, we examine firm-specific gender and ethnicity pay differentials in Britain using matched employer-employee data from the 1998 Workplace Employment Relations Survey (WERS98), which allows us to control for unobserved firm heterogeneity and identify its role in determining pay differentials related to gender and ethnicity. Following Pudney (2000) and Cardoso (2000) we incorporate gender and ethnicity dummies, interacted with firm specific effects.

The paper is organised as follows. We begin in Section 2 by describing the matched survey data, and provide a measurement of pay. In Section 3 we present our model and propose two alternative tests checking normality. Section 4 explores a semi-parametric finite mixture random effects estimator that allows for the interaction of non-normal workplace effects with individual worker's gender/ethnicity characteristics. Section 5 presents the

results from the random effects specification. Section 6 uncovers firm-specific indicators of disadvantage at the workplace. Finally, conclusions are drawn in Section seven.

2. The WERS98 data

2.1 Survey design

WERS98 is the fourth in a series of industrial relations surveys that began in 1980, and is the first that includes workplaces with fewer than 25 employees or contains a matched survey of employees. Interviews were conducted in 2,191 workplaces between October 1997 and June 1998 with a plant-level response rate of 80%. Workplaces were sampled from the Inter-Departmental Business Register (IDBR). The sample is stratified by workplace employment size and industrial sector. It is nationally representative, but excludes agriculture and mining.

We use information from two of the WERS98 questionnaires. The management questionnaire was implemented in a face-to-face interview with the most senior workplace manager with day-to-day responsibility for personnel matters. It gives information on the nature of the workplace, business conditions, product markets, union coverage, organisation and establishment and the composition of the labour force, particularly its occupational, gender and ethnic mix. The second element of our dataset is the employee survey, which was administered to 25 randomly-selected employees (or the whole workforce, where the establishment had fewer than 25 employees), using a self-completion questionnaire; 28,323 questionnaires were returned, giving a 64% response rate. The employee survey gives information on earnings, ethnicity, gender, educational attainment, age, job tenure, and occupational class.

We use the subset of establishments and workers who supplied complete information on critical variables and we restrict attention to full-time employees (at least 30 hours per week). This results in a sample of 20,345 workers, linked to a set of 1,727 firms. Summary statistics are presented in Table A1 in the Appendix.

2.2 Measurement of pay

The wage information asked of respondents in the employee questionnaire relates to the following question: "*How much do you get paid for your job here, before tax and other deductions are taken out? If your pay changes before tax from week to week because of overtime, or because you work different hours each week, think of what you earn on average*". Respondents were asked to place their pay level within 12 bands, chosen to approximate decile bands and the top and bottom 5% of the earnings distribution as estimated from the 1996 New Earnings Survey. The level of missing responses to this question was 1% (Cully *et al.*, 1999).

Empirical work on the labour market usually uses the hourly wage rate as a measure of the return to labour. In fact, for many workers there is no such thing as the hourly wage. The employment relationship is very often a complex relationship in which there is a package of rewards and constraints accepted by the employee, rather than a simple constant unit price market for workers' time. To investigate the robustness of our results to this issue, we use two alternative measures of the reward to labour. Our first measure of pay is the hourly wage, equal to ratio of weekly earnings to the number of working hours per week, including any overtime or extra hours.

The second measure is total weekly earnings, excluding bonuses or other non-standard pay items. This can be justified under a different view of the employment relation and the underlying technology. Consider a stylised example. There is a sequence of production periods, each of length T, in which productive activity takes place. In each period, the employer requires a fixed set of activities, p, to be completed by the worker. Now suppose that the technology is sufficiently flexible that workers can deliver this contracted volume of activity at a rate and over a time span of their own choosing, provided it complies with the overall production timetable. Thus p can be decomposed as $p = t\,e$, where t is time spent doing productive work during the production period and e is intensity of effort during time at work. The time T-t is used as on or off the job leisure or 'social' time. The employer is indifferent between alternative (t, e) combinations provided $t \leq T$. The worker's problem is then to choose a utility-maximising combination (t, e) subject to the constraints $te = p$ and $t \leq T$. Workers with different tastes will choose different (t, e) combinations. Hours of work are essentially meaningless here: respondents might report them either as t or as the conventional standard length of the work period T, even if actual activity time t is less than T. In any case, the relevant return to productive activity is measured appropriately by total payment per contracted task completed during the production period, p. This is reflected directly by weekly or annual earnings. On the other hand, pay per reported hour may be contaminated by confusions between t and T.

3. The econometric model

3.1 The interval regression specification

We use a conventional semi-log regression model for individual pay. For worker i in firm h:

$$w_{ih} = \beta_0 + \boldsymbol{x}_{ih}\boldsymbol{\beta} + u_{0h} + \xi_{1ih}u_{1h} + \xi_{2ih}u_{2h} + \varepsilon_{ih} \qquad i = 1 \ldots m_h;\ h = 1 \ldots n \qquad (1)$$

where $\boldsymbol{x}_{ih}$ is a vector of observable covariates, ξ_{1ih} = 1 if worker *i,h* is female and ξ_{1ih} = 0 otherwise; ξ_{2ih} is a similar dummy if worker *i,h* is a member of the ethnic minority group. The unobservable variable ε_{ih} is the usual random disturbance term distributed with mean zero and unknown variance σ^2. In the hourly pay model, we retain log earnings as the dependent variable but include log working hours per week (including any overtime or extra hours) as a regressor, with its coefficient restricted to 1.

The unobservable variable u_{0h} is a general firm-specific wage premium; u_{1h} and u_{2h} are firm-specific gender and ethnicity wage differentials respectively. Conditional on $\boldsymbol{x}_{ih}$, ξ_{1ih} and ξ_{2ih}, we treat u_{0h}, u_{1h} and u_{2h} as random workplace effects, distributed randomly with unrestricted means and variances, subject to a mean-independence assumption $\mathrm{E}(\boldsymbol{x}_{ih,}u_{jh})= 0$ for j = 0, 1, 2. Given the cross-sectional nature of the data, any individual is only observed with one employer, so unobserved individual heterogeneity cannot be identified and is consigned to the error term. The firm-specific random effect u_{oh} captures unobserved firm effects common to all individuals at the establishment.

There is a complication induced by the design of the WERS98 questionnaire, since the (log) wage w_{ih} is observed only within ranges. Let the observed pay interval for worker *i*, *h* be $R_{ih} = (\underline{W}_{ih}, \overline{W}_{ih})$ and assume normality for the error term ε_{ih}. Then the log-likelihood for this model is

$$\ln L = \sum_{h=1}^{H} \ln\left(\Pr(w_{ih} \in R_{ih} \,|\, \boldsymbol{x}_{ih}, \xi_{1ih}, \xi_{2ih})\right) \qquad (2)$$

The relevant probability is

$$\Pr(w_{ih} \in R_{ih} \,|\, \boldsymbol{x}_{ih}, \xi_{1ih}, \xi_{2ih}) = E_{\mathbf{u}}\left[P_{ih}(\mathbf{u}_h)\right] \tag{3}$$

where:

$$P_{ih}(\mathbf{u}_h) = \Phi\left(\left[\overline{W}_{ih} - \lambda_{ih}\right]/\sigma\right) - \Phi\left(\left[\underline{W}_{ih} - \lambda_{ih}\right]/\sigma\right) \tag{4}$$

and $\lambda_{ih} = \beta_0 + \boldsymbol{x}_{ih}\beta + u_{0h} + \xi_{1ih}u_{1h} + \xi_{2ih}u_{2h}$ and $E_{\mathrm{u}}[\ .\ ]$ denotes the expectation with respect to the random effects distribution. The implementation of this maximum likelihood (ML) estimator requires some method of approximating this expectation.

3.2 Constant ethnicity and gender differentials

Let the means of u_{1h} and u_{2h} be μ_1 and μ_2. In that case, they can be treated as constant parameters to be estimated, with the dummy variables ξ_{1ih} and ξ_{2ih} absorbed into the vector $\boldsymbol{x}_{1ih}$. Define $u_{1h}^{*} = u_{1h} - \mu_1$ and $u_{2h}^{*} = u_{2h} - \mu_2$. If we treat the composite variable $v_{ih} = u_{0h} + \xi_{1ih}\,u_{1h}^{*} + \xi_{2ih}\,u_{2h}^{*}$ as a residual and estimate equation (1) by Gaussian interval regression (IR) techniques, the resulting estimates may be inconsistent for two separate reasons: non-normality and heteroskedasticity. Since the model is nonlinear, departures from normality of v_{ih} produce inconsistency. Secondly, the variance of the composite error $v_{ih} + \varepsilon_{ih}$ depends on gender and ethnicity, so the model is heteroskedastic. Again, the standard IR estimator is inconsistent under heteroskedasticity, whether or not the normality assumption is correct. To address these issues, we allow explicitly for the existence of between-firm variation in ethnic and gender pay differentials and we allow the three firm effects to have a general non-normal joint distribution.

3.3 The form of the pay distribution

The normality assumption often fails in the tails of the distribution (Chay and Honore, 1998). We propose two new tests for non-normality in the IR model. The first examines the effect of aggregating a number of pay intervals at the top or bottom ends of the pay scale, using a likelihood-based technique. Re write the conditional probability of observing the *j*th pay interval as $P_{jih}(\theta) = \Phi((W_j - \lambda_{ih})/\sigma) - \Phi((W_{j-1} - \lambda_{ih})/\sigma)$ where θ is the parameter vector and the W_j are the boundaries of the pay intervals. The likelihood element for a representative observation is:

$$l_{ih} = \prod_j P_{jih}(\theta)^{y_{jih}} \tag{5}$$

where $y_{jih} = 1$ if w_{ih} is in the *j*th pay range and 0 if not. Consider a subset *S* of the pay ranges in the upper or lower tail of the pay distribution and decompose the likelihood element as:

$$\begin{aligned} l_{ih} &= \left\{ P_{Sih}(\theta)^{y_{Sih}} \prod_{j \notin S} P_{jih}(\theta)^{y_{jih}} \right\} \left\{ \frac{\prod_{j \in S} P_{jih}(\theta)^{y_{jih}}}{P_{Sih}(\theta)} \right\}^{y_{Sih}} \\ &= l_{ih}^{*}(\theta) \times l_{ih}^{**}(\theta) \end{aligned} \tag{6}$$

where $y_{sih} = 1$ if $W_{ih} \in S$ and 0 otherwise.

Note that the components $l_{ih}^{*}(\theta)$ and $l_{ih}^{**}(\theta)$ are both likelihoods in their own right and can be maximised separately to give alternative estimates of $\boldsymbol{\theta}$. We follow the approach of Ruud (1984) and test the specification of the model by carrying out a likelihood ratio test of H_0: $\boldsymbol{\theta}^{*} = \boldsymbol{\theta}^{**}$, using the following statistic:

$$\chi^2 = -2\left[\sum_h \sum_i l_{ih}(\hat{\theta}) - \sum_h \sum_i l_{ih}^{*}(\hat{\theta}^{*}) - \sum_h \sum_i l_{ih}^{**}(\hat{\theta}^{**}) \right] \tag{7}$$

where $\hat{\theta}^{*}$ and $\hat{\theta}^{**}$ are parameter vectors estimated by maximising the component likelihoods. Under the null hypothesis of correct specification, this is distributed as χ^2 with degrees of freedom equal to the dimension of $\boldsymbol{\theta}$. Table 1 provides the estimated values of the above test statistic for the hourly and the weekly models and for different numbers of pay ranges. We reject the null hypothesis for all the bottom and top ranges.

Table 1. ML tests for impact of aggregating tail pay ranges.

Number of pay intervals aggregated	Weekly pay		Hourly wage	
	Bottom tail	Top tail	Bottom tail	Top tail
7	986.7	9058.4	915.1	8040.1
8	1131. 5	29641.4	1023.3	28014.0
9	1174.9	949.3	1311.4	44926.1
10	46571.9	60939.7	745.2	58105.9
Note: all statistics are $\chi^2(59)$; 10%, 5% and 1% critical values are 73.28, 77.93 and 87.17.				

The second test is based on the observation that the IR model is nested within the following ordered probit model:

$$w_{ih}^{*} = \boldsymbol{x}_{ih}\gamma + \upsilon_{ih} \tag{8}$$

where $w_{ih}^{*} = w_{ih} / \sigma$, $\upsilon_{ih} = v_{ih}/\sigma$ and $\gamma = \beta / \sigma$ and σ is the standard deviation of v_{ih}. Then:

$$\Pr(y_{ih} = j \mid \boldsymbol{x}_{ih}) = \Phi(C_j - \boldsymbol{x}_{ih}\gamma) - \Phi(C_{j-1} - \boldsymbol{x}_{ih}\gamma) \tag{9}$$

where C_0 ... C_m are fixed parameters normalised by $C_0 = -\infty$, $C_m = +\infty$. Then the following restrictions should be satisfied by the ordered probit model:

$$(W_j - \beta_0)/\sigma = C_j \tag{10}$$

These equalities can be tested with a likelihood ratio test (LR). For both the hourly and weekly pay definitions, they are highly significant and reject the IR model (Table 2).

Table 2. LR tests of ordered probit against interval regression.

Pay definition	$\chi^2(9)$ test statistic[1]
Hourly	567.9
Weekly	437.8
[1] Critical values 14.68, 16.92 and 21.67 at the 10%, 5% and 1% levels.	

Figures 1 and 2 plot the left- and right-hand sides of equation (10) and show the nature of departures from normality. The main problem is the bottom tail of the conditional pay distribution, where the ordered probit results show that the upper limit of the bottom pay range would need to be shifted rightwards to capture the relatively large number of low-pay individuals.

Figure 1. Comparison of interval regression and ordered probit thresholds (weekly model).

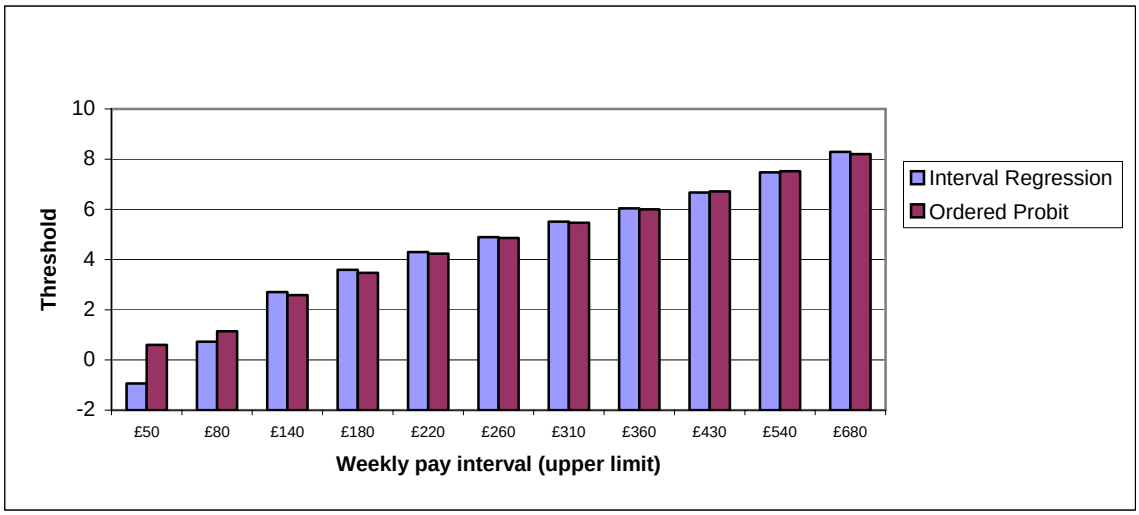
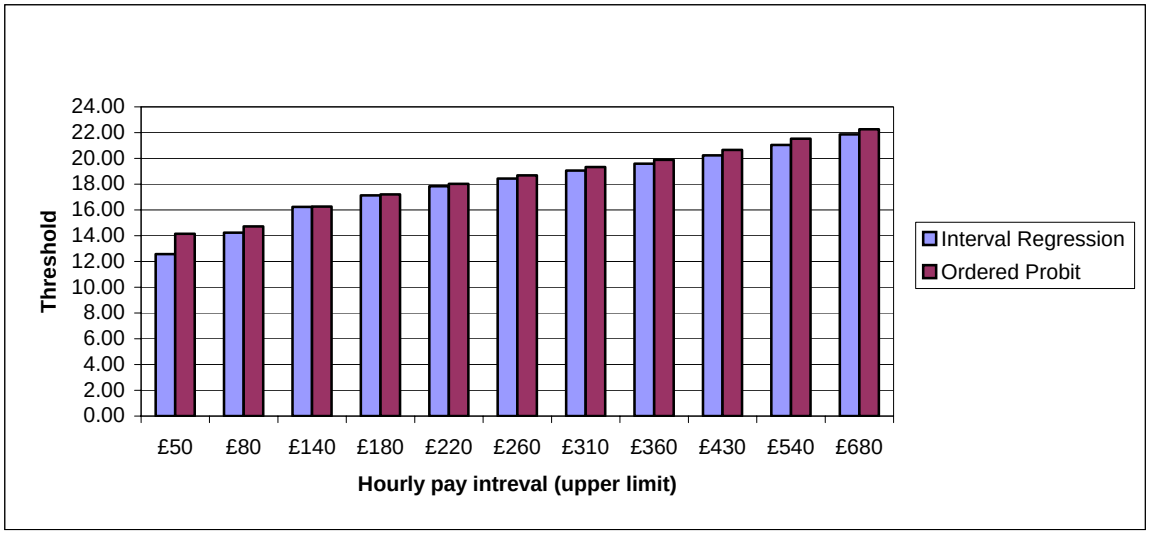


Figure 2. Comparison of interval regression and ordered probit thresholds (hourly model).



4. Semi-parametric random-effects estimation

The simple IR model is clearly not tenable empirically. We now explore a generalised approach that allows for the interaction of non-normal workplace effects with individual workers' gender/ethnicity characteristics.

4.1 The finite mixture approach

We use a semi-parametric finite mixture (FM) random-effects estimator based on the model (3)-(4). This approximates the distribution of the random effects $\mathbf{u}_h$ by an arbitrary trivariate discrete distribution, where the location and magnitude of the probability mass points are treated as fixed parameters. Thus:

$$E_{\mathbf{u}}\left[P_{ih}\left(\mathbf{u}_h\right)\right] = \sum_{q=1}^{Q} \pi^q P_{ih}\left(\mathbf{u}^q\right) \tag{11}$$

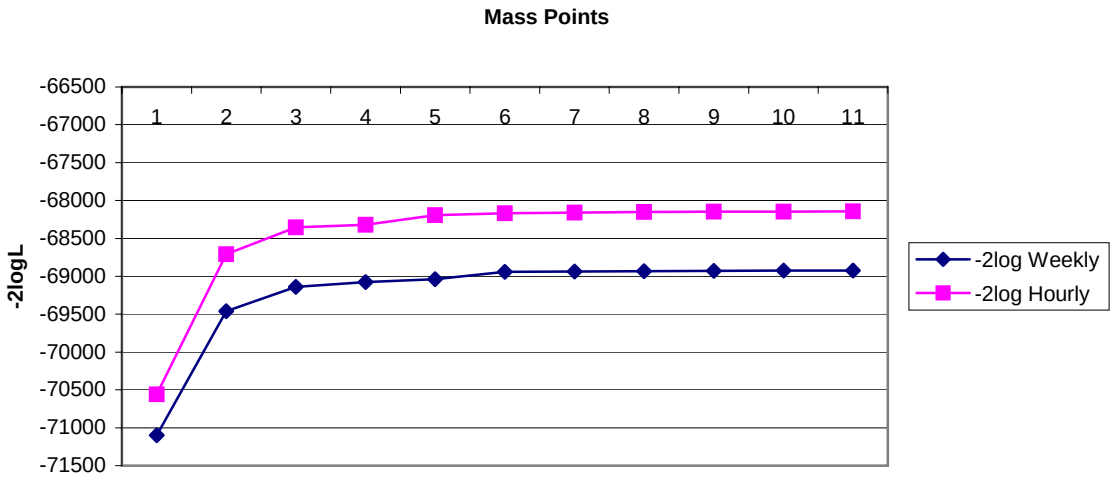
The mass points $\mathbf{u}^q$ are additional parameters. The probabilities π^q must be non-negative and sum to unity, so we parameterise them (without loss of generality) as a multinomial logit:

$$\pi^q = \frac{\exp(\rho^q)}{\sum_{r=1}^{Q} \exp(\rho^r)} \tag{12}$$

where ρ^1 is normalised to 0 and ρ^2 ... ρ^Q are free parameters. Note that distribution $\{\pi^q \; u^q\}$ is degenerate unless some elements of the vectors u^q are equal. However, since the parameters space is not constrained with respect to the u^q, degeneracy is not being imposed on the distribution. An alternative specification would be to have a fixed set of mass points for each u_0, u_1 and u_2 and then generate the u^q as their cartesian products. This would be a restricted version of our model and would have the disadvantage that for a given number of parameters it would generate many more terms in the sum of (11).

The log-likelihood function (2) is maximised numerically for a sequence of specifications with different numbers of mass points *Q* (Figure 3). It is important to repeat the computational algorithm from a number of alternative starting points (five in this case), since there are known to be multiple optima in this class of likelihoods (see Laird, 1978; Heckman and Singer, 1984).

Figure 3. Comparison of mass points of the weekly and hourly specifications.



4.2 Results

We report results for the basic IR and FM models in Appendix Tables A2 and A3. The impact of individual ethnicity and gender was specified as follows. We started with a ten-category breakdown formed from the interaction of the two gender groups with five ethnic groups: Black (Afro-Caribbean and African); Indian; Pakistani, Bangladeshi, other (white European and others). In a simple IR model, the intercepts for these groups could be represented adequately by three dummy variables: females; ethnic minority and ethnic minority women. The restrictions embodied in this specification were accepted at the 5% level of significance despite the large sample size. Since there was virtually no difference between the restricted and unrestricted models in the estimates of other coefficients, this specification was judged to be acceptable. It is also particularly convenient since it allows us to work with a single disadvantaged racial group despite the heterogeneity of educational qualifications and labour market outcomes between Britain's ethnic minority groups (Dustmann and Theodoropoulos, 2006).

Job characteristics include eight dummy variables that capture broad occupational classification[1] according to the 1991 Standard Occupational Classification Guides (manager, professional, associate professional and technical, clerical and secretarial, craft and skilled service, personal and protective service, sales, plant and machine operatives, omitted category "other occupation" i.e. cleaner, postal worker). We also allow for trade union membership, having a temporary job status and years of tenure in the current job (seniority).

[1] It has long been argued whether one should include controls for occupational status in a study of discrimination since occupational differences may be caused by discrimination (Blau and Ferber, 1987). However, occupations are affected differentially by compensating wage differentials and efficiency wages (Bell and Ritchie, 1998). Thus, excluding occupation might lead to omitted variable bias.

Other establishment attributes include: the size of the workforce (the log number of employees at the establishment), union density and its interaction with gender and ethnicity, the legal status of the establishment (public sector administration and its interaction with gender and private sector services; the omitted category is private sector manufacturing), being a multi-establishment (part of a larger organization, a 0/1 dummy), if the degree of competition in the market that the establishment operates is very high/high (a 0/1 dummy), and if the establishment supplies its goods and services to the local market (a 0/1 dummy). Additionally, we control for the region at which the establishment is located according to the standard statistical region classification by including four region specific dummy variables (London, Rest of the South East, West-Midlands and Scotland)[2] as well as two dummy variables coming from the management questionnaire and capturing different levels of the unemployment to vacancy rate by travel to work area.

Table A3 in the Appendix gives the results of the random effects wage equation for both the weekly and hourly specifications estimated using a likelihood based on the distribution of equation 3. The random effects model involves three establishment-specific unobservables: a general firm specific effect u_{0h}; a female firm specific effect u_{1h}; and an ethnic minority firm specific effect u_{2h}. The coefficients are in general consistent in sign, significance and magnitude across the two specifications. We present the results simultaneously from the weekly and hourly specifications.

Besides gender, ethnicity and their interaction, other individual attributes in the model include a quadratic in age, marital status, health status and educational attainment. The age

[2] Initial tests showed that the dummy variables for the other regions obtained an insignificant coefficient in both the weekly and hourly specifications and were excluded. Their exclusion did not have an impact on the coefficients of the other covariates.

profile has an inverse U-shape. Unmarried status and work-relevant health problems are both associated with a significant decrease in both weekly and hourly wages. Returns to educational attainment range from 6% for low education qualification (low CSE) to 22% for a university degree. A postgraduate degree has an incremental return of 26% for weekly wages and 17% for hourly wages. Vocational qualifications, such as a trade apprenticeship, NVQ, or a City and Guilds Certificate, are associated with a small but significant wage disadvantage of around 1%. There are large, significant occupational pay differentials, especially for the highest status jobs. To allow for the possibility that the return to educational qualifications might be reduced if the individual is denied access to an appropriate occupation, we include interactions between education and occupation. These are significant, particularly for degree and managerial/professional (15%), A-level or above and technical (7%) and high CSE or above and services (5%). Within the sales occupation, low educational attainment brings a significant disadvantage (-11%). We find that female skilled employees enjoy a higher skill premium than skilled men (see also Groshen, 1991).

Workforce composition variables capture the profile of the stock of workers employed by the establishment, in terms of gender, ethnicity and their interactions with the gender and ethnicity dummy variables respectively. We also control for the proportion of each occupational group at the establishment, the proportion of part-time employees, and proportions of staff over 50 and under 21 years of age. We find that the higher the percentage of female and ethnic minority employees at the establishment the lower is pay for all employees. Although the coefficient of the percentage of female employees is negative, its interaction with the female is positive and significant. Thus female employees experience less pay disadvantage in establishments with high densities of female employees and males

working in female-dominated establishments do less well than other men. One interpretation of this is that bargaining power for women is higher in establishments where women are overrepresented. Another interpretation is that women are attracted to employers who do not discriminate. In contrast, the corresponding interaction term for ethnic minority employees is negative, implying that ethnic minority employees face greater disadvantage in establishments with a high density of ethnic minority employees. The opposite results found for these individual-firm interactions for gender and ethnicity conflict with explanations based on bargaining power or supply-driven segregation. If a concentration of women in a workplace increases women's power to oppose discrimination or is a signal of a lack of discrimination, then why does not the same mechanism work for ethnic minorities? Our view is that these results reflect an important distinction: that women are a large, widely-dispersed group, whereas ethnic minorities are small, locally-concentrated and less integrated in wider society. For a member of an ethnic minority, working in a minority-dominated establishment may be a symptom of weak integration and poor access to the opportunities offered by wider society. The poor outside option counteracts the bargaining power that a large group of workers might otherwise have.

We find a significant relationship between the occupational profile of an establishment and its wage-setting behaviour: establishments employing a high proportion of managerial, professional and skilled staff tend to be high-wages employers, whereas workplaces employing high proportions of part-time, young, and old staff tend to pay relatively low wages.

Union density within the workplace has a small but significant positive effect for both members and non-members. Like Hildreth (1999), we also find an interaction with gender:

hourly wages for female employees are lower in establishments with high levels of union density. The interaction effect of ethnicity and union density is positive and significant in both specifications (see Blau and Kahn, 1996).

Firm size has a positive influence on wages. The business conditions faced by the employer also play a significant role. Establishments whose main product market is local rather than regional, national or international, are associated with generally lower levels of pay. Being part of a large firm (multi-establishment) increases both weekly and hourly pay by 4.6% and 3.4% correspondingly. There is also evidence of rent sharing, with highly competitive product market conditions implying slightly lower wage levels of about 1.6% for weekly wages and 2.6% for hourly wages. These are features that are shared with most other studies (see Blanchflower *et al*., 1996; Hellerstein *et al*., 2002; Troske, 1999).

Working in the public sector (administration) involves a significant pay disadvantage of 7.4% for weekly wages and 6.8% for hourly wages. However, the interaction dummy between a public sector establishment and being female is positive and significant. This implies that there is a positive return to working in the public sector for women, possibly reflecting more effective equal pay policies.

We find significant regional differences. Both weekly and hourly earnings are highest in London, the South East, West Midlands and (hourly earnings only) Scotland. Local labour market conditions are also significant, with high unemployment to vacancies ratios reducing expected earnings at the level of travel to work area.

5. The random effects distribution

Our final specification is a 9-point trivariate discrete distribution for the establishment effects.[3] The specification was determined by means of likelihood ratio criteria. We reject the specifications of 10 and 11 mass points in favour of 9 mass points for both models (see also Figure 3). The computed χ^2 for the weekly random effects model between 9 and 10 mass points is 3.45 (4 degrees of freedom) and between 9 and 11 mass points is 4.49 (8 degrees of freedom). Similarly, for the hourly specification, the χ^2 statistics were 2.42 and 4.83 respectively.

The implied means, standard deviations and correlations of the three establishment effects u_{0h}, u_{1h}, u_{2h} for the weekly and the hourly regressions are given in Table 3.

Table 3. Wage equations: semi-parametric random effects.

PARAMETER	WEEKLY REGRESSION	HOURLY REGRESSION
Mean of general effect (u_o)	4.150 (0.035)	0.407 (0.037)
Mean of gender effect (u_1)	-0.218 (0.013)	-0.133 (0.013)
Mean of race effect (u_2)	-0.159 (0.027)	-0.140 (0.026)
Std dev (u_o)	0.140 (0.005)	0.148 (0.005)
Std dev (u_1)	0.071 (0.008)	0.065 (0.008)
Std dev (u_2)	0.086 (0.045)	0.076 (0.022)
$\hat{\sigma}$	0.251 (0.001)	0.247 (0.001)
Correlations		
$\hat{\rho}_{01}$	-0.520 (0.061)	-0.576 (0.065)
$\hat{\rho}_{02}$	-0.041 (0.462)	-0.015 (0.250)
$\hat{\rho}_{12}$	-0.461 (0.607)	-0.283 (0.360)

Note: Standard errors in parentheses.

[3] Estimation was done in GAUSS using the MAXLIK procedure. Since WERS98 is a stratified two-stage probability sample we also carried out a weighted estimation of the above specification by including into the estimation the establishment weight. The changes in the magnitude of the coefficients were very moderate, and the standard errors very close or the same with those obtained from the semi-parametric random effects estimation.

The means for gender and race firm specific effects are different, especially for the weekly regression, where the means imply an average pay disadvantage of 22% for women and a 16% pay disadvantage for ethnic minorities. Our estimated gender pay gap is as high as that reported by Harkness (1996) and slightly higher than the 21% disadvantage reported by Mumford and Smith (2004) who use the same dataset as in this study.

However, the positive coefficient estimated for the interaction dummy variable of being female and member of an ethnic minority group (10%, see Table A3 in the Appendix) implies that the ethnic minority females face an average pay disadvantage of roughly 28% rather than the 38% that would otherwise be implied. In other words, pay differentials are on average 22% for ethnic minority men but only 6% for ethnic minority women.[4]

The variances of the firm effects in the weekly regression are highly significant. The variance of the general firm effects is roughly double the variance of the gender and ethnicity effects. The significant negative correlation $\hat{\rho}_{01}$ implies that high-wage firms tend to pay low female weekly wages. However, the correlations between the general firm effect and ethnicity wage premium ($\hat{\rho}_{02}$), and between the female wage premium and the ethnicity wage premium ($\hat{\rho}_{12}$) are insignificant.

On an hourly pay basis, ethnic pay differentials are on average 13% for ethnic minority men and 6% for ethnic minority women. The finding for males closely mirrors the 11% male wage differential found in the UK in the 1990s by Blackaby *et al*. (2002). The estimated variances and correlations of the firm effects are similar to those in the weekly earnings model.

[4] Bronars and Famulari (1997) using US matched employer-employee data (Bureau of Labor Statistics White Collar Pay Survey, 1989-1990) find a monthly wage gap of 27.5% between white and black full-time private sector workers.

6. Firm-specific indicators of disadvantage

The posterior distribution of **u** conditional on the observed variables relevant to establishment *h* is given by:

$$dF(\mathbf{u}_h \mid \mathbf{X}_h, \Xi_h) = \frac{\Pr(\mathbf{y}_h \mid \mathbf{X}_h, \Xi_h, \mathbf{u}_h)\, dF(\mathbf{u}_h)}{\Pr(\mathbf{y}_h \mid \mathbf{X}_h, \Xi_h)} \tag{15}$$

where $\mathbf{y}_h = \{y_{1h}\ y_{2h} \ldots\}$, $\mathbf{X}_h = \{\mathbf{x}_{1h}\ \mathbf{x}_{2h} \ldots\}$, $\Xi_h = \{\xi_{11h}\ \xi_{21h}\ ,\ \xi_{12h}\ \xi_{22h}\ ,\ \ldots\}$. The mean of this distribution for workplace *h*, using the finite mixture assumption is given by:

$$\hat{\mathbf{u}}_h = \frac{\sum_{q=1}^{Q} \mathbf{u}^q \Pr(\mathbf{y}_h \mid \mathbf{X}_h, \Xi_h, \mathbf{u}^q)\pi^q}{\sum_{q=1}^{Q} \Pr(\mathbf{y}_h \mid \mathbf{X}_h, \Xi_h, \mathbf{u}^q)\pi^q} \tag{16}$$

Our aim here is to uncover indirect indicators of disadvantage at the workplace by investigating the empirical relationship between $\hat{u}_{0h}$, $\hat{u}_{1h}$ and $\hat{u}_{2h}$, and a range of variables relating to employees' perceptions about managers and working conditions (Table 4) and employer's policies/practices and the establishment's performance (Table 5). These variables were not included in the econometric specification as they are potentially endogenous. However, they can be used post-estimation to shed light on the firm-specific unobservables $\hat{u}_{1h}$ and $\hat{u}_{2h}$. If these are clearly related to other observable indicators of the firm's attitude towards equal opportunities, this would provide some support for an interpretation in terms of employers' tastes for discrimination.

We restrict this analysis to establishments in which there was at least one employee recorded and a positive percentage of women and/or ethnic minority employees who responded to the survey, giving a subsample of 1,715 firms.

Table 4. Correlations between the estimated means of the three firm specific effects and variables of interest from the employee questionnaire.

		WEEKLY						HOURLY					
		General		Gender		Ethnic		General		Gender		Ethnic	
Variables	Binary Outcome	No. of Responses	$\bar{\hat{u}}_{0h}$	No. of Responses	$\bar{\hat{u}}_{1h}$	No. of responses	$\bar{\hat{u}}_{2h}$	No. of responses	$\bar{\hat{u}}_{0h}$	No. of responses	$\bar{\hat{u}}_{1h}$	No. of Responses	$\bar{\hat{u}}_{2h}$
Employees who say that managers are poor/very poor in treating them fairly	y=0 y=1	1338 377	4.150 4.149	1403 144	-0.218 -0.214	320 15	-0.163 -0.167	1338 377	0.405 0.402	1403 144	-0.132 -0.132	320 15	-0.148 -0.155
Employees who say that managers are poor/very poor in dealing with work problems that employees may have	y=0 y=1	1297 418	4.150 4.150	1388 159	-0.218 -0.212	319 16	-0.163 -0.157	1297 418	0.406 0.400	1371 176	-0.133 -0.130	319 16	-0.148 -0.157
Employees who disagree/strongly disagree that they share many of the values of their organization	y=0 y=1	1459 256	4.152 4.141	1442 105	-0.218 -0.207	327 8	-0.163 -0.164	1459 256	0.405 0.401	1442 105	-0.133 -0.120	327 8	-0.148 -0.149
Employees who disagree/strongly disagree that their job is secure in the workplace	y=0 y=1	1335 380	4.152 4.143	1401 146	-0.218 -0.212	320 15	-0.163 -0.164	1335 380	0.406 0.398	1401 146	-0.133 -0.128	320 15	-0.149 -0.139
Employees who think that the job at the workplace that personally do is done only/mainly by men	y=0 y=1	1099 616	4.146 4.158	1500 47	-0.218 -0.215	325 10	-0.163 -0.164	1099 616	0.399 0.414	1500 47	-0.132 -0.129	325 10	-0.148 -0.150
Employees who think that the job at the workplace that personally do is done equally by men and women	y=0 y=1	1154 561	4.152 4.146	1279 268	-0.218 -0.215	311 24	-0.162 -0.172	1154 561	0.406 0.401	1279 268	-0.132 -0.133	311 24	-0.147 -0.157
Employees who think that the job at the workplace that personally do is done only/mainly by women	y=0 y=1	1177 538	4.152 4.145	1063 484	-0.218 -0.217	321 14	-0.163 -0.153	1177 538	0.407 0.398	1063 484	-0.134 -0.129	321 14	-0.149 -0.136

Table 5. Correlations between the estimated means of the three firm specific effects and variables of interest from the management questionnaire.

		Weekly						Hourly					
		General		Gender		Ethnic		General		Gender		Ethnic	
Variables	Binary outcome	No. of responses	$\bar{\hat{u}}_{0h}$	No. of responses	$\bar{\hat{u}}_{1h}$	No. of responses	$\bar{\hat{u}}_{2h}$	No. of responses	$\bar{\hat{u}}_{0h}$	No. of responses	$\bar{\hat{u}}_{1h}$	No. of responses	$\bar{\hat{u}}_{2h}$
If there is a formal written policy on equal opportunities or managing diversity policy have you tried to measure the effects of equal opportunities policies on the workplace or on the employees at the establishment	y=0 y=1	1358 357	4.151 4.147	1358 357	-0.218 -0.216	1358 357	-0.158 -0.161	1358 357	0.406 0.399	1358 357	-0.133 -0.129	1358 357	-0.145 -0.146
Managers who assess that labour productivity at the establishment is a lot better/better than average compared with other establishments in the same industry	y=0 y=1	999 716	4.147 4.154	999 716	-0.217 -0.218	999 716	-0.159 -0.159	999 716	0.401 0.410	999 716	-0.131 -0.134	999 716	-0.145 -0.144
If tribunal application in last year how many complaints were made in the last year	y=0 y=1	1363 352	4.149 4.156	1363 352	-0.218 -0.216	1363 352	-0.158 -0.161	1363 352	0.403 0.409	1363 352	-0.132 -0.132	1363 352	-0.144 -0.149
If tribunal application in last year were the grounds on sex discrimination	y=0 y=1	1660 55	4.150 4.165	1660 55	-0.217 -0.217	1660 55	-0.159 -0.157	1660 55	0.404 0.416	1660 55	-0.132 -0.136	1660 55	-0.145 -0.147
If tribunal application in last year were the grounds on race discrimination	y=0 y=1	1668 47	4.150 4.157	1668 47	-0.217 -0.221	1668 47	-0.159 -0.160	1668 47	0.404 0.413	1668 47	-0.132 -0.137	1668 47	-0.145 -0.147
Grievance on sex or race discrimination raised in the past year through a procedure or not	y=0 y=1	1601 114	4.149 4.169	1601 114	-0.217 -0.220	1601 114	-0.159 -0.156	1601 114	0.403 0.429	1601 114	-0.132 -0.138	1601 114	-0.145 -0.146

Tables 4 and 5 both show very small differences between the mean values of $\hat{u}_{1h}$ and $\hat{u}_{2h}$ between groups of firms categorised by variables representing employees' perceptions, or workplace conditions, practices and performance. The differences turn out to be so small that there is no need to construct formal hypothesis tests to assess their statistical significance. The finding of no association between estimated firm-specific gender and ethnic pay differentials is striking. It makes it difficult to sustain an interpretation of the econometric results as a reflection of explicit discriminatory practices and suggests instead more subtle explanations, such as those based on differences in unmeasured human/social capital.

7. Conclusions

This is the first British study that uses matched employer employee data and examines the role of firm specific effects in shaping gender and ethnicity pay differentials. We find robust evidence in support of significant pay differentials between men and women and between white and non-white employees. The empirical estimation reveals a 22% weekly gender pay gap and a 28% weekly race pay gap. The corresponding hourly estimates are 13% and 19%. We also show that the inclusion of unobserved establishment heterogeneity in a conventional human capital based earnings function adds an important feature to the determinants of wages. For instance, we find strong evidence that high wage firms tend to pay low female wages.

Another contribution of the paper in the applied econometrics literature is that we propose two alternative tests and show that the maintained hypotheses of normality in the widely-used Interval Regression model is not tenable empirically. Thus, using a semi-parametric finite mixture random effects estimator we address non-normality and heteroskedasticity

problems by allowing for the interaction of non-normal workplace effects with individual workers' gender/ethnicity characteristics.

Calculating the posterior distribution of the firm specific unobservable variables and investigating their relationship with other variables that may act as indirect indicators of discrimination, we find no significant relationship. Thus, it is very difficult to reach any clear conclusions about the source of inter-firm variations in pay differentials or the corresponding policy implications.

References

Abowd, J., Kramarz, F. and Margolis, D. (1999). 'High wage workers and high wage firms', *Econometrica*, Vol. 67, pp. 251-334.

Altonji, J. and Blank, R. (1999). Race and gender in the labor market. In Ashenfelter, O. and Card, D. (eds.). Handbook of Labor Economics, Vol. 3, North-Holland, Amsterdam, 3143-3259.

Bell, D. and Ritchie, F. (1998). 'Female earnings and gender differentials in Great Britain 1977-1994', *Labour Economics*, Vol. 5, pp. 331-357.

Blackaby, D., Leslie, D., Murhpy, P. and O'Leary, N. (2002). 'White/ethnic minority earnings and employment differentials in Britain: evidence from the LFS', *Oxford Economic Papers*, Vol. 54, pp. 270-297.

Blanchflower, D., Oswald, A. and Sanfey, P. (1996). 'Wages, profits and rent sharing', *Quarterly Journal of Economics*, Vol. 111, pp. 227-252.

Blau, F. and Ferber, M. (1987). 'Discrimination: empirical evidence from the Unites States', *American Economic Review (Papers and Proceedings)*, Vol. 77, pp. 316-320.

Blau, F. and Kahn, L. (1996). 'Wage structure and gender earnings differentials: an international comparison', *Economica*, Vol. 63, pp. S29-S62.

Blau, F. and Kahn, L. (2003). 'Understanding international differences in the gender pay gap', *Journal of Labor Economics*, Vol. 21, pp. 106-144.

Blau, F. and Kahn, L. (2006). 'The US gender pay gap in the 1990s: slowing convergence', IZA Discussion Paper No. 2176.

Borooah, V. and Lee, K. (1988). 'The effect of changes in Britain's industrial structure on female relative pay and employment', *Economic Journal*, Vol. 98, pp. 818-832.

Bowles, S., Gintis, H. and Osborne, M. (2001). 'The determinants of earnings: a behavioral approach', *Journal of Economic Literature*, Vol. 39, pp. 1137-1176.

Bronars, S. and Famulari, M. (1997). 'Wage, tenure, and wage growth variation within and across establishments', *Journal of Labor Economics*, Vol. 15, pp. 285-317.

Card, D. and Krueger, A. (1992). 'School quality and black-white relative earnings: a direct assessment', *Quarterly Journal of Economics*, Vol. 107, pp. 151-200.

Cardoso, A. (2000). 'Wage differentials across firms: an application of multilevel modelling', *Journal of Applied Econometrics*, Vol. 15, pp. 343-354.

Chandra, A. (2003). Is the convergence of the racial wage gap illusory? NBER Working Paper No. 9476.

Chay, K. and Honore, B. (1998). 'Estimation of semiparametric censored regression models: an application to changes in black-white earnings inequality during the 1960s', *Journal of Human Resources*, Vol. 33, pp. 4-38.

Cully, M., Woodland, S. O'Reilly, A and Dix, G. (1999). Britain at work: as depicted by the 1998 Workplace Employee Relations Survey, Routledge.

Darity, W. and Nembhard, G. (2000). 'Racial and ethnic economic inequality: the international record', *American Economic Review (Papers and Proceedings)*, Vol. 90, pp. 308-311.

Donohue, J. and Heckman, J. (1991). 'Continuous versus episodic change: the impact of civil rights policy on the economic status of blacks', *Journal of Economic Literature*, Vol. 29, pp. 1603-1643.

Dustmman, C. and Theodoropoulos, N. (2006). Ethnic minority immigrants and their children in Britain, *CReAM* Discussion Paper, Department of Economics, University College London.

Goldin, C. (2004). 'The long road to the fast track: career and family', NBER Working Paper No. 10331.

Goldin, C. and Katz, L. (2000). 'Changes in fertility patterns', *American Economic Review (Papers and Proceedings)*, Vol. 90, pp. 461-465.

Goslin, A., Machin, S. and Meghir, C. (2000). 'The changing distribution of male wages in the UK', *Review of Economic Studies*, Vol. 67, pp. 635-666.

Groshen, E. (1991). 'The structure of the female/male wage differential: is it who you are, what you do, or where you work'? *Journal of Human Resources*, Vol. 26, pp. 457-472.

Hamermesh, D. (1999). 'LEEping into the future of labor economics: the research potential of linking employer and employee data', *Labour Economics*, Vol. 6, pp. 25-41.

Harkness, S. (1996). 'The gender earnings gap: evidence from the UK', *Fiscal Studies*, Vol. 17, pp. 1-36.

Heckman, J. and Rubinstein, Y. (2001). 'The importance of non-cognitive skills: lessons from the GED testing program', *American Economic Review*, Vol. 91, pp. 145-149.

Heckman, J. and Singer, B. (1984). 'A method for minimizing the impact of distributional assumptions in econometric models for duration data', *Econometrica*, Vol. 52, pp. 271-320.

Hellerstein, J., Neumark, D. (2005). Using matched employer-emlpoyee data to study labor market discrimination, IZA Discussion Paper, No. 1555.

Hellerstein, J., Neumark, D. and Troske, K. (1999). 'Wages, productivity and worker characteristics. Evidence from plant-level production functions and wage equations', *Journal of Labor Economics*, Vol. 17, pp. 409-446.

Hellerstein, J., Neumark, D. and Troske, K. (2002). 'Market forces and sex discrimination', *Journal of Human Resources*, Vol. 37, pp. 353-380.

Hildreth, A. (1999). 'What has happened to the union wage differential in Britain in the 1990s?' *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp. 5-31.

Laird, N. (1978). 'Nonparametric maximum likelihood estimation of a mixing distribution', *Journal of the American Statistical Association*, Vol. 73, pp. 805-811.

Lang, K. and Manove, M. (2006). Education and labor-market discrimination, NBER Working Paper No. 12257.

Meng, X. (2004). 'Gender earnings gap: the role of firm specific effects', *Labour Economics*, Vol. 11, pp. 555-573.

Mumford, K. and Smith, P. (2004). The gender earnings gap in Britain, IZA Discussion Paper, No. 1109.

Neal, D. and Johnson, W. (1996). 'The role of premarket factors in black-white wage Differentials', *Journal of Political Economy*, Vol. 104, pp. 869-895.

O'Neil, J. and Polachek, S. (1993). 'Why the gender gap narrowed in the 1980's', *Journal of Labor Economics*, Vol. 11, pp. 205-228.

Pudney, S. (2000). Pay differentials, discrimination and worker grievances, Public Sector Economics Research Centre Discussion Paper, Department of Economics, University of Leicester.

Pudney, S. and Shields, M. (2000). 'Gender, race, pay and promotion in the British nursing profession: estimation of a generalised ordered probit model', *Journal of Applied Econometrics*, Vol. 15, pp. 367-399.

Ruud, P. (1984). 'Tests of specification in econometrics', *Econometric Reviews*, Vol. 3, pp. 211-242.

Smith, J. and Welch, F. (1989). 'Black economic progress after Mydral', *Journal of Economic Literature*, Vol. 27, pp. 519-564.

Troske, K. (1999). 'Evidence on the employer-size wage premium from worker establishment matched data', *Review of Economics and Statistics*, Vol. 81, pp. 15-26.

Zabalza, A. and Tzannatos, Z. (1985). 'The effect of Britain's anti-discriminatory legislation on relative pay and employment', *Economic Journal*, Vol. 95, pp. 679-699.

APPENDIX

Table A1. Sample properties of variables (Weighted N=20,345).

VARIABLE	MEAN	VARIABLE	MEAN	VARIABLE	MEAN
Job tenure (years)	7.128	Clerical	0.154	Female percentage	48.242
	(0.010)		(0.006)		(1.174)
Temporary job	0.019	Craft	0.133	Female percentage & female	21.777
	(0.001)		(0.007)		(0.681)
Working hours per week	42.677	Service occupation	0.059	Trade union density	0.264
	(0.137)		(0.007)		(0.013)
Trade union member	0.443	Sales	0.047	Minority & union density	0.012
	(0.013)		(0.003)		(0.001)
Age (years)	39.508	Operatives	0.164	Female & union density	0.122
	(0.016)		(0.010)		(0.004)
Female	0.366	Skilled female	0.192	Public sector	0.264
	(0.008)		(0.006)	Administration	(0.016)
Minority	0.028	Sale & CSE	0.010	Private sector (services)	0.545
	(0.002)		(0.001)		(0.018)
Ethnic minority woman	0.012	Service & (GCSE or	0.040	Female & public sector	0.134
	(0.002)	Alevel or Degree)	(0.005)		(0.006)
Ethnic minority man	0.016	Technical & (Alevel or	0.054	Competitive firm	0.584
	(0.001)	Degree)	(0.003)		(0.017)
Ethnic majority man	0.619	Professional staff or	0.112	Ln (Size of firm)	3.675
	(0.009)	manager & Degree	(0.005)		(0.025)
Ethnic majority woman	0.353	Proportion part-time	0.192	Part of a large firm	0.691
	(0.008)		(0.008)		(0.023)
Low school qualification (CSE)	0.122	Proportion managerial	0.112	Market local	0.308
	(0.004)		(0.005)		(0.019)
High school qualification (GCSE)	0.248	Proportion professional	0.145	London	0.112
	(0.005)		(0.009)		(0.016)
A-level	0.147	Proportion technical	0.074	Rest of the South East	0.226
	(0.004)		(0.008)		(0.021)
Degree	0.175	Proportion clerical	0.194	West Midlands	0.010
	(0.004)		(0.009)		(0.016)
Postgraduate degree	0.062	Proportion craft	0.114	Scotland	0.076
	(0.003)		(0.009)		(0.009)
Vocational qualification	0.404	Proportion sales staff	0.081	Banded total	0.667
	(0.006)		(0.006)	unemployment/vacancy rate by travel to work area (3%-6%)	(0.023)
Health problem	0.061	Proportion services staff	0.075	Banded total unemployment/	0.070
	(0.003)		(0.007)	vacancy rate by travel to work area (more than 7%)	(0.010)
Unmarried	0.227	Proportion over 50	0.152	Ln(lower level of wage per	5.519
	(0.005)		(0.006)	week)	(0.013)
Managerial	0.113	Proportion under 21	0.051	Ln(upper level of wage per	5.804
	(0.004)		(0.003)	week)	(0.013)
Professional	0.144	Minority percentage	4.136	Ln(lower level of wage per	1.778
	(0.005)		(0.288)	hour)	(0.013)
Technical	0.094	Minority percentage	0.598	Ln(upper level of wage per	2.063
	(0.004)	& minority	(0.474)	hour)	(0.013)

Note: Standard errors in parentheses.

Table A2. Non-heterogeneous interval regression results (robust standard errors).

WEEKLY PAY			HOURLY PAY		
COVARIATE	COEFFICIENT	STD ERROR	COVARIATE	COEFFICIENT	STD ERROR
Tenure/10	0.063***	0.0051	Tenure/10	0.070***	0.0051
Temporary Job	-0.091***	0.0176	Temporary Job	-0.073***	0.0173
Age	0.503***	0.0174	Age	0.466***	0.0167
(Age/10)^2	-0.055***	0.0020	(Age/10)^2	-0.050***	0.0019
Unmarried	-0.086***	0.0056	Unmarried	-0.079***	0.0056
Trade union member	0.065***	0.0069	Trade union member	0.054***	0.0068
Low school qualification (Low CSE)	0.053***	0.0088	Low school qualification (Low CSE)	0.054***	0.0087
High school qualification (High CSE)	0.119***	0.0077	High school qualification (High CSE)	0.116***	0.0076
Intermediate school qualification (A level)	0.153***	0.0088	Intermediate school qualification (A level)	0.151***	0.0087
University degree	0.212***	0.0130	University degree	0.202***	0.0127
Postgraduate degree	0.404***	0.0136	Postgraduate degree	0.344***	0.0134
Vocational qualification	-0.016***	0.0049	Vocational qualification	-0.014***	0.0049
Health problem	-0.040***	0.0098	Health problem	-0.030***	0.0093
Female	-0.226***	0.0161	Female	-0.129***	0.0158
Minority	-0.161***	0.0275	Minority	-0.132***	0.0272
Ethnic minority woman	0.091***	0.0260	Ethnic minority woman	0.087***	0.0244
Managerial	0.581***	0.0135	Managerial	0.512***	0.0131
Professional	0.429***	0.0145	Professional	0.407***	0.0141
Technical	0.265***	0.0152	Technical	0.284***	0.0151
Clerical	0.096***	0.0143	Clerical	0.132***	0.0137
Craft (skilled manual)	0.162***	0.0130	Craft (skilled manual)	0.150***	0.0128
Service occupation	0.165***	0.0280	Service occupation	0.157***	0.0321
Sales	0.191***	0.0221	Sales	0.192***	0.0202
Operative	0.036***	0.0130	Operative	0.015	0.0130
Professional or managerial & degree	0.145***	0.0143	Professional or managerial & degree	0.112***	0.0142
Technical & (alevel or degree)	0.078***	0.0152	Technical & (alevel or degree)	0.074***	0.0153
Service & (highcse or, alevel or, degree)	0.102***	0.0293	Service & (highcse or, alevel or, degree)	0.084***	0.0316
Sales & low CSE	-0.142***	0.0329	Sales & low CSE	-0.091***	0.0308
Female and skilled	0.038***	0.0112	Female and skilled	0.036***	0.0110
Percentage of female & female	0.0009***	0.0003	Percentage of female & female	0.0006**	0.0002
Percentage of females	-0.0024***	0.0003	Percentage of females	-0.0014***	0.0003
Percentage of ethnic minority & minority	-0.0017**	0.0079	Percentage of ethnic minority & minority	-0.0023***	0.0007
Percentage of minorities	-0.0009*	0.0005	Percentage of minorities	-0.0005	0.0005
Proportion part-time staff	-0.237***	0.0321	Proportion part-time staff	-0.219***	0.0310
Proportion managerial staff	0.195***	0.0421	Proportion managerial staff	0.256***	0.0411
Proportion professional staff	0.186***	0.0264	Proportion professional staff	0.224***	0.0267
Proportion technical staff	0.098***	0.0253	Proportion technical staff	0.192***	0.0267
Proportion clerical staff	0.225***	0.0249	Proportion clerical staff	0.310***	0.0253
Proportion craft (skilled) staff	0.037	0.0241	Proportion craft (skilled) staff	0.091***	0.0258
Proportion service staff	0.103***	0.0299	Proportion service staff	0.141***	0.0322
Proportion sales staff	0.216***	0.0326	Proportion sales staff	0.267***	0.3300
Proportion staff over 50	-0.227***	0.0416	Proportion staff over 50	-0.171***	0.0420
Proportion staff under 21	-0.374***	0.0669	Proportion staff under 21	-0.395***	0.0595
Ln(employment)	0.030***	0.0032	Ln(employment)	0.0336***	0.0033

Continued

Table A2. Non-heterogeneous interval regression results (robust standard errors).

Continued

WEEKLY PAY			HOURLY PAY		
COVARIATE	COEFFICIENT	STD ERROR	COVARIATE	COEFFICIENT	STD ERROR
Union density	0.045**	0.0178	Union density	0.110***	0.0188
Female & trade union density	-0.008	0.0185	Female & trade union density	-0.053***	0.0184
Minority & trade union density	0.093**	0.0400	Minority & trade union density	0.071*	0.0384
Public sector (administration)	-0.097***	0.0167	Public sector (administration)	-0.106***	0.0166
Female & public sector	0.037***	0.0122	Female & public sector	0.018	0.0124
Private sector (service)	-0.029**	0.0132	Private sector (service)	-0.051***	0.0135
Part of large firm	0.025	0.0119	Part of large firm	0.026**	0.0122
Local product market	-0.029***	0.0096	Local product market	-0.023**	0.0098
Highly competitive market	-0.020**	0.0091	Highly competitive market	-0.033***	0.0092
London	0.219***	0.0148	London	0.217***	0.0152
Rest of the South East	0.089***	0.0105	Rest of the South East	0.086***	0.0112
West-Midlands	0.040***	0.0139	West-Midlands	0.036***	0.0141
Scotland	0.186	0.0131	Scotland	0.040***	0.0125
Banded total unemployment vacancy rate by travel to work area (3%-6%)	-0.052***	0.0101	Banded total unemployment vacancy rate by travel to work area (3%-6%)	-0.050***	0.0102
Banded total unemployment vacancy rate by travel to work area (more than 7%)	-0.078***	0.0161	Banded total unemployment vacancy rate by travel to work area (more than 7%)	-0.065***	0.0171
Constant	4.176***	0.0484	Constant	0.380***	0.0453
Sigma	-1.262***	0.0094	Sigma	-1.271***	0.0097
Llog=-35549.694	Wald chi2(59)=16970.47		Llog=-35279.004	Wald chi2(59)=13570.85	
N=20345	Prob>chi2=0.0000		N=20345	Prob>chi2=0.0000	

*Note: *,**,***, 10%, 5% and 1% level of significance correspondingly. The results are obtained by estimating equation 2,* $w_{ih} = \beta_0 + \boldsymbol{x}_{ih}\boldsymbol{\beta} + u_{0h} + \xi_{1ih}u_{1h} + \xi_{2ih}u_{2h} + \varepsilon_{ih}$ *as a standard IR model. The dummy variables* ξ_{1ih} *and* ξ_{2ih} *are absorbed into the vector of observable characteristics* $\boldsymbol{x}_{ih}$ and we *treat the composite error term* v_{ih} *as a residual.*

Table A3. Semi-parametric random effects results (robust standard errors).

WEEKLY PAY			HOURLY PAY		
COVARIATE	COEFFICIENT	STD ERROR	COVARIATE	COEFFICIENT	STD ERROR
Tenure/10	0.063***	0.0041	Tenure/10	0.065***	0.0041
Temporary Job	-0.085***	0.0123	Temporary Job	-0.072***	0.0124
Age	0.491***	0.0115	Age	0.445***	0.0116
(Age/10)^2	-0.053***	0.0014	(Age/10)^2	-0.048***	0.0014
Unmarried	-0.072***	0.0056	Unmarried	-0.066***	0.0057
Trade union member	0.059***	0.0048	Trade union member	0.046***	0.0049
Low school qualification (Low CSE)	0.056***	0.0081	Low school qualification (Low CSE)	0.057***	0.0079
High school qualification (High CSE)	0.113***	0.0066	High school qualification (High CSE)	0.110***	0.0067
Intermediate school qualification (Alevel)	0.151***	0.0081	Intermediate school qualification (Alevel)	0.146***	0.0081
University degree	0.196***	0.0104	University degree	0.185***	0.0102
Postgraduate degree	0.392***	0.0099	Postgraduate degree	0.330***	0.0099
Vocational qualification	-0.012***	0.0041	Vocational qualification	-0.013***	0.0040
Health problem	-0.031***	0.0092	Health problem	-0.027***	0.0090
Ethnic minority woman	0.096***	0.0261	Ethnic minority woman	0.080***	0.0247
Managerial	0.569***	0.0087	Managerial	0.498***	0.0087
Professional	0.417***	0.0090	Professional	0.394***	0.0090
Technical	0.247***	0.0116	Technical	0.261***	0.0113
Clerical	0.079***	0.0100	Clerical	0.112***	0.0099
Craft (skilled manual)	0.158***	0.0094	Craft (skilled manual)	0.145***	0.0095
Service occupation	0.147***	0.0153	Service occupation	0.147***	0.0142
Sales	0.177***	0.0113	Sales	0.175***	0.0118
Operative	0.040***	0.0093	Operative	0.012	0.0092
Professional or managerial & degree	0.145***	0.0113	Professional or managerial & degree	0.111***	0.0112
Technical & (alevel or degree)	0.066***	0.0131	Technical & (alevel or degree)	0.060***	0.0127
Service & (highcse or, alevel or, degree)	0.051***	0.0172	Service & (highcse or, alevel or, degree)	0.039**	0.0165
Sales & low CSE	-0.107***	0.0216	Sales & low CSE	-0.063***	0.0224
Female & skilled	0.039***	0.0086	Female & skilled	0.035***	0.0086
Percentage of female & female	0.0007***	0.0002	Percentage of female & female	0.0004**	0.0002
Percentage of females	-0.0021***	0.0002	Percentage of females	-0.0014***	0.0002
Percentage of ethnic minority & minority	-0.0020***	0.0007	Percentage of ethnic minority & minority	-0.0016**	0.0006
Percentage of minorities	-0.0015***	0.0005	Percentage of minorities	-0.0013***	0.0004
Proportion part-time staff	-0.233***	0.0211	Proportion part-time staff	-0.207***	0.0232
Proportion managerial staff	0.231***	0.0356	Proportion managerial staff	0.250***	0.0401
Proportion professional staff	0.202***	0.0219	Proportion professional staff	0.200***	0.0231
Proportion technical staff	0.138***	0.0265	Proportion technical staff	0.210***	0.0285
Proportion clerical staff	0.223***	0.0207	Proportion clerical staff	0.314***	0.0228
Proportion craft (skilled) staff	0.028	0.0253	Proportion craft (skilled) staff	0.058**	0.0237
Proportion service staff	0.091***	0.0206	Proportion service staff	0.111***	0.0224
Proportion sales staff	0.152***	0.0235	Proportion sales staff	0.205***	0.0253
Proportion staff over 50	-0.215***	0.0285	Proportion staff over 50	-0.191***	0.0324
Proportion staff under 21	-0.254***	0.0361	Proportion staff under 21	-0.317***	0.0437
Ln(employment)	0.031***	0.0032	Ln(employment)	0.037***	0.0033
Union density	0.044***	0.0155	Union density	0.105***	0.0157

Continued

Table A3. Semi-parametric random effects results (robust standard errors).

Continued

WEEKLY PAY			HOURLY PAY		
COVARIATE	COEFFICIENT	STD ERROR	COVARIATE	COEFFICIENT	STD ERROR
Female & trade union density	0.005	0.0164	Female & trade union density	-0.029*	0.0164
Minority & trade union density	0.122***	0.0427	Minority & trade union density	0.088**	0.0434
Public sector (administration)	-0.077***	0.0166	Public sector (administration)	-0.070***	0.0184
Female & public sector	0.035***	0.0127	Female & public sector	0.022*	0.0122
Private sector (services)	-0.014	0.0122	Private sector (services)	-0.032**	0.0139
Part of large firm	0.045***	0.0091	Part of large firm	0.033***	0.0103
Local product market	-0.030***	0.0089	Local product market	-0.020**	0.0090
Highly competitive market	-0.016*	0.0085	Highly competitive market	-0.026***	0.0091
London	0.220***	0.0127	London	0.227***	0.0134
Rest of the South East	0.097***	0.0098	Rest of the South East	0.085***	0.0105
West-Midlands	0.032**	0.0132	West-Midlands	0.041***	0.0148
Scotland	0.011	0.0124	Scotland	0.034**	0.0134
Banded total unemployment/vacancy rate by travel to work area (3%-6%)	-0.046***	0.0088	Banded total unemployment/vacancy rate by travel to work area (3%-6%)	-0.051***	0.0096
Banded total unemployment/vacancy rate by travel to work area (more than 7%)	-0.078***	0.0157	Banded total unemployment/vacancy rate by travel to work area (more than 7%)	-0.071***	0.0159
Sigma	0.251***	0.0009	Sigma	0.247***	0.0009
-2Llog=68930.260 N=20345	AIC=40.021		-2Llog=68148.0 N=20345	AIC=39.568	

*Note: *,**,***, 10%, 5% and 1% level of significance correspondingly. The results are estimated by approximating equation 3,* $\Pr(w_{ih} \in R_{ih} \mid \mathbf{x}_{ih}, \xi_{1ih}, \xi_{2ih}) = E_{\mathbf{u}}[P_{ih}(\mathbf{u}_h)]$ *through a semi-parametric random effects finite-mixture estimator.*