Equal pay for equal work? The effects of gender on the earnings of English general practitioners

Hugh Gravelle* **Arne Hole***

Abstract
Aims. To estimate the effects of gender on the earnings of English general practitioners and to compare alternative methods of quantifying these effects.

Methods. The traditional method of quantifying the effect of gender on pay is calculate an Oaxaca decomposition of the difference in the means of male and female wage rates, with the wage rate calculated as earnings divided by hours worked. This method is inappropriate for occupations in which earnings are not proportional to hours worked. We compare results from (a) a model of wage (earnings/hours), (b) a two stage least squares model of earnings which includes instrumented hours worked, (c) a reduced form earnings model which excluded the hours variable. We also test directly for discrimination by examining the effect of single handed working and the proportion of female GP on remuneration.

Data. The data are from the 2004 NPCRDC GP worklife survey which had usable responses from 1777 GPs.

Results. Earnings were higher in dispensing and PMS practices and in practices with more patients but was unaffected by the number of patients per GP. Male GPs earned £81K per annum, and worked 48 hours pw; female GPs earned £57K pa and worked 37 hours pw. Female GP earnings were more elastic with respect to hours (0.75) than were male earnings (0.45). Discrimination as measured by the share of the difference in log mean wages or earnings attributable to differences in coefficients in the log wage or log earnings regressions varies between 23% and 90% depending on the model and assumed counterfactual. Direct tests using the proportion of female GPs do not provide any support for the discrimination hypothesis. The positive effect of being single handed is greater for female GPs which is compatible with discrimination but the small number of female single handers means that the test is very low powered.

Keywords: Wage. Income. Earnings. Hours. Gender. Decomposition. Family physicians.
JEL nos: I11, J44, J7

WORK IN PROGRESS: NOT FOR QUOTATION

* National Primary Care Research and Development Centre, Centre for Health Economics, University of York; emails: hg8@york.ac.uk; ah522@york.ac.uk. NPCRDC receives core funding from the Department of Health. The views expressed are those of the authors and not necessarily those of the funders. The NPCRDC General Practitioner Worklife Survey was undertaken by Diane Whalley, Chris Bojke, Hugh Gravelle and Bonnie Sibbald. We are grateful to Mark Dusheiko, Paul Chambers-Dixon, and Andrew Wagner for providing additional data; and to Karen Mumford and Jo Swaffield for comments.

1 Introduction

In this paper we investigate the reasons for the difference in pay between male and female English general practitioners (GPs) using data from the spring of 2004.

Around two thirds of GPs were then paid under the General Medical Services contract which was negotiated at national level between the Department of Health and the British Medical Association. Under it GPs were paid by a mixture of capitation fees, lump sum payments, fee for service, and for meeting a small number of quality targets. Under the alternative Primary Medical Services contract the practice received a lump sum in return for providing all the services it provided under GMS plus additional services agreed with its local Primary Care Trust. The lump sum was typically based on the total payments the practice had previously received under GMS. Practices, under both GMS and PMS, had to meet the costs of providing services from their gross income.

There are two ways in which the remuneration system could lead to discrimination against female GPs. Almost all general practices operate as partnerships, pooling gross income from the GMS or PMS contract, deducting expenses and sharing out net income according to their partnership agreements. Hence, within a given practice, the partnership agreement could give female GPs a lower share of the practice profits than otherwise identical male GPs. We label this *within practice discrimination* Second, because of variations in the populations served and restrictions on the entry of practices, there may be variations in practice profits. Hence, if female GPs are discriminated against when seeking to join the more profitable practices, they will be disproportionately concentrated in the less profitable practices. We label this *entry discrimination*.

1.1 Quantifying discrimination

The standard method for quantifying the amount of discrimination is to estimate separate remuneration models for male and female workers and then to use the regression models to decompose the difference in mean wages (or more usually, differences in the mean log wage) into parts attributable to differences in the estimated coefficients and to differences in the means of the explanatory variables:

$$\overline{\ln w^M} - \overline{\ln w^F} = (\overline{x}^M - \overline{x}^F)\hat{\alpha}^M + \overline{x}^F(\hat{\alpha}^M - \alpha^F) \qquad (1)$$

where $\overline{\ln w}$, $\overline{x}$, is the mean of the log of wages, $\overline{x}$ is the mean of the vector of explanatory variables, $\hat{\beta}$ is the vector of estimated coefficients from the Mincer log wage regression (Mincer, 1970; 1974) and the superscripts indicate gender (Oaxaca, 1973; Blinder, 1973).

The first part of the decomposition, due to differences in the means of the explanatory variables, is the "explained" difference and the second part, due to differences in coefficients, is the "unexplained" difference. The percentage of the overall difference in means due to the unexplained difference is often used as a summary measure of discrimination. Note that the x vector includes the unit constant so that the differences in the regression constant terms are counted as part of discrimination.

Since the difference in explanatory variable means is not counted in the "unexplained" or discrimination part of the decomposition, the measure is at best a measure of within practice discrimination. It fails to capture entry discrimination.

The assumption justifying (1) is that in the absence of discrimination female workers would have their wage or earnings determined in the same way as male workers: the effect (β) of their characteristics (x) on their remuneration would be the same as for male workers and the second term in the decomposition would be zero. The assumption can be questioned: in general changing the remuneration function will change the labour supply of workers, leading to a change in the remuneration function to restore equilibrium in the labour market. A number of authors have suggested variants of the basic Oaxaca-Blinder which amount to changing the assumptions about what the counterfactual equilibrium remuneration function would be (Reimers, 1983; Cotton, 1988). In the absence of evidence on the appropriate counterfactual remuneration function we calculate decompositions with a number of different assumptions about the counterfactual.

In most decomposition studies the measure of remuneration is the log hourly wage rate, with hourly pay calculated as total earnings divided by hours worked. If workers are indeed on fixed hourly wages, so that pay is proportional to hours, then the procedure is sensible. But if pay is not proportional to hours, a remuneration function with the calculated wage as the dependent variable is misspecified. The coefficients in the estimated wage equation will reflect the market reward for characteristics such as experience, the effect of the characteristics on hours worked, and the effect of unobserved characteristics which influence hours worked. Hence it may not be sensible to test for the existence of discrimination (by comparing coefficients on particular characteristics across genders) and to quantify it by a decomposition.

If pay is not proportional to hours worked an earnings model which includes hours worked can be estimated. The decomposition can then be applied to the estimated earnings equations:

$$\overline{\ln y^M} - \overline{\ln y^F} = (\bar{x}^M - \bar{x}^F)\hat{\beta}^{xM} + (\overline{\ln h^M} - \overline{\ln h^F})\hat{\beta}^{hM} + \bar{x}^F(\hat{\beta}^{xM} - \hat{\beta}^{xF}) + \overline{\ln h^F}(\hat{\beta}^{hM} - \hat{\beta}^{hF}) \quad (2)$$

where *y* is earnings and *h* is hours. See for example studies on gender discrimination in the legal profession (McNabb and Wass, 2006; Wood, Corcoran and Courant, 1993). But using the sum of the last two terms as measure of discrimination raises two further problems. First, hours worked is an endogenous variable jointly determined with earnings and so the estimated coefficient on hours worked is likely to be biased. Since hours worked may be an important determinant of earnings the bias in estimating the coefficients could have serious consequences for the decomposition of the earnings difference. Hence in this paper, unlike McNabb and Wass (2006) and Wood, Corcoran and Courant (1993), we use two stage least squares to estimate the earnings equation, with marital status and number of children under 18 as instruments.

The second possible difficulty with using an earnings equation (either 2SLS or OLS) to quantify discrimination is that the decomposition of the difference in earnings will attribute differences in earnings due to differences in hours worked to the "explained" non-discriminatory part of the decomposition. But hours worked are determined, in part by the earnings function, and so the "unexplained" part of the earnings difference will underestimate effect of differences in the coefficients in the earnings function. Whether this matters in constructing a summary measure of discrimination depends on whether one believes that all the differences in the coefficients in the hours models are due to discrimination. It seems plausible that at least part of the difference in coefficients in hours models is due to differences in preferences. For example, it is seems odd to include differential responses to family circumstances in a measure of gender discrimination at work.

If one does wish to include differences in coefficients in the hours models as part of discrimination, a way to do so is to use the first stage hours equation to decompose the difference in mean hours into a part due to differences in variables ("explained") and a part due to differences in the coefficients in the hours model ("unexplained"):

$$\overline{\ln h^M} - \overline{\ln h^F} = (\bar{x}^M - \bar{x}^F)\hat{\gamma}^{xM} + (\bar{z}^M - \bar{z}^F)\hat{\beta}^{zM} + \bar{x}^F(\hat{\gamma}^{xM} - \hat{\gamma}^{xF}) + \bar{z}^F(\hat{\gamma}^{zM} - \hat{\gamma}^{zF}) \quad (3)$$

where *z* is a vector of instruments. We can then substitute the log hours decomposition (3) into the log earnings decomposition (2) to get

$$\overline{\ln y^M} - \overline{\ln y^F} = (\bar{x}^M - \bar{x}^F)\hat{\delta}^{xM} + (\bar{z}^M - \bar{z}^F)\hat{\delta}^{zM} + \bar{x}^F(\hat{\delta}^{xM} - \hat{\delta}^{xF}) + \bar{z}^F(\hat{\delta}^{zM} - \hat{\delta}^{zF}) \quad (4)$$

where $\hat{\delta}^{xM} = \hat{\beta}^{xM} + \hat{\gamma}^{xM}\hat{\beta}^{hM}$, $\hat{\delta}^{zM} = \hat{\gamma}^{zM}\hat{\beta}^{hM}$, $\hat{\delta}^{xF} = \hat{\beta}^{xF} + \hat{\gamma}^{xF}\hat{\beta}^{hM}$ and $\hat{\delta}^{zF} = \hat{\gamma}^{zF}\hat{\beta}^{hM}$. Discrimination is then measured as the sum of the last two terms in (4).

We illustrate this approach by adopting a simpler but equivalent method: estimate a single reduced form earnings model for each gender on all the exogenous variables (the union of the sets of *x* and *z* variables) and then calculate the standard two part decomposition for this enlarged set of variables.

1.2 Testing for within practice discrimination

One possible form of discrimination is in the way practice earnings are shared amongst the partnership's GPs: female GPs may get a smaller share of practice income given their characteristics than male GPs. We test for within practice discrimination in two ways:

(a) There is likely to be less discrimination against women in practices where women are a higher proportion of the partners. We therefore include the proportion of GPs who are female as an explanatory variable in the wage and earnings regressions. If there is discrimination against female GPs we expect to find that coefficient on the proportion of female GPs is positive for female GPs. If female GPs also discriminate against male GPs we would expect also that the coefficient on the female GP proportion variable is negative in the male hours and earnings regressions.

(b) There can be no discrimination in profit sharing against female GPs within single handed practices. We include a dummy variable taking the value 1 if the practice is single handed, and zero otherwise in the log wage/earnings models. If there is discrimination against female GPs in multipartner practices the coefficient on the single hander dummy will be larger in the female regressions than in the male regressions. Incentive theory suggests that there will be dilution of effort as the number of partners increases (Gaynor and Pauly, 1990). Hence we expect that even in the male wage/earnings regressions, the coefficient on the single hander dummy will be positive. If there is discrimination it will be smaller than the coefficient in the female GP models.

The next section briefly describes the data and has more detail on the methods.

2 Data and Methods

2.1 Data

The data are from the 2004 National Primary Care Research and Development Centre's General Practitioner worklife survey (Whalley et al, 2005). The questionnaire was posted in February 2004 to 4208 salaried and principal GPs in England, of whom 2166 (52%) responded. Usable responses were received from a total of 598 female and 1179 male GPs. The respondents were asked to report their earnings, defined as total annual income from their

practice before taxes but after deducting expenses from their practice (in bands)[1]; usual hours worked per week; and personal and practice characteristics.

We were also able to add further information the GPs' practices from the Attribution Data Set which has detailed information on variables which determine the practice's capitation income and from the General Medical Services data set maintained by the Department of Health.

2.2 Explanatory variables

We divide the variables potentially affecting GP remuneration and hours worked into

personal. As is standard we include a measure of experience (decades since qualification) and its square to allow for the usual positive but declining effect of experience on earnings. We include ethnicity (white/non-white) to allow for any possible discrimination based on race. We also know whether the GP qualified in the UK or abroad. This will be correlated, though not perfectly with ethnicity, and because we also include ethnicity the coefficient on non-UK qualified may reflect the impact of less relevant initial experience on earnings. We also include the GPs salaried status and expect that it will have a negative effect on hours worked and on earnings.

family. We use a set of family characteristics as instruments for hours worked in the 2SLS models of earning. We know the GP's marital status and whether their partner works. We collapse this information into a single dummy which takes the value 1 if the GP has a partner who does not work and 0 otherwise (no partner or partner works). We expect that GPs with a non working partner will work longer hours to boost family income and because a non-working partner will be able to provide child care. We also include dummy variables for numbers of children under 18.

practice. In addition to whether the GP works in a single handed practice and the proportion of female GPs in the practice which we use to test directly for discrimination in the wage and earnings equations, we also have information on other practice characteristics which are likely to affect total practice income. Under the GMS contract the per patient capitation income of the practice depended on three adjustments. The age/sex adjustment is a weighted sum of the age and sex mix of the practice population, with the weights intended to reflect national age and sex specific consultation rates. Broadly, the practice receives more gross income the

[1] The income bands are: less than £25.000, £25.000-£49.999, £50.000-£69.999, £70.000-£84.999, £85.000-£99.999, £120.000-£149.999 and £150.000 or more.

older its population. The nursing home adjustment is based on the proportion of the population in nursing homes, with higher capitation paid the higher the proportion. The additional needs adjustment is based on the morbidity of population as measured by the standardised limiting long term illness ratio and the standardised mortality ratio. We have information on these three adjustments for each practice from the Attribution Data Set. Note that the nursing home adjustment and the additional need adjustment are attributed to practices using the proportion of their patients living in different areas

We include the number of practice patients per whole time equivalent GP and expect GPs in practices with larger lists per GP to have higher earningss. We also include the total practice population. The total list is highly correlated with number of GPs and the incentive literature suggests that practices with more GPs will have lower earnings per GP. On the other hand, there may be economies of scale in delivering services, so that earnings per GP may increase with total list size.

Practices can earn additional income from patients by providing them with various fee per item services, such a flu vaccinations, and by meeting targets for vaccination of children and cervical screening. It may be harder to generate this type of income from less educated or poorer populations. We include a measure of the education of the local population and the Low Income Scheme Index score. The LISI score is the cost of prescriptions dispensed without charge to the patient on grounds of low income as a proportion of the total cost of prescriptions dispensed for practice patients (Lloyd, Harris and Clucas, 1995).

Finally, we include indicators of whether the practice is in PMS and whether it is permitted to dispense as well prescribe pharmaceuticals to its patients. Given the relative information and incentives of GPs and PCT managers we expect PMS practices to have negotiated a contract with their PCT which generates higher total practice profits than they would earn under the nationally negotiated GMS contract. Dispensing practices are reimbursed for their purchases of medicines at fixed prices which generally exceed those at which they actually buy.

region. We use 8 Government Office Region dummies to capture geographical differences in the cost of practice inputs.

2.3 Decomposition methods

The general decomposition, using the difference in the log of male and female wages as an example, is

$$\overline{\ln w^M} - \overline{\ln w^F} = (\bar{x}^M - \bar{x}^F)(\lambda\hat{\alpha}^M + (1-\lambda)\hat{\alpha}^F) + ((1-\lambda)\bar{x}^M + \lambda\bar{x}^F)(\hat{\alpha}^M - \hat{\alpha}^F) \quad (5)$$

where $\lambda \in [0,1]$. It is assumed that if there was no discrimination the coefficients on characteristics would be a weighted combination of the estimated coefficients. The original, and still most common assumption, which we used in section 1, is that $\lambda = 1$: in the absence of discrimination women would face the male wage equation. Alternative weighting schemes considered in the literature include $\lambda = 0.5$ (Reimers, 1983) and $\lambda = N^M/(N^M + N^F)$ (Cotton, 1988), where N^M and N^F are the number of males and females in the sample. In the absence of any theoretical model to fix λ, we investigate the robustness of the decomposition, and hence the measure of discrimination, by setting $\lambda = 1$ and $\lambda = 0$.

Oaxaca and Ransom (1999) show that a detailed decomposition of the unobserved component $((1-\lambda)\bar{x}^M + \lambda\bar{x}^F)(\hat{\beta}^M - \hat{\beta}^F)$ in (5) suffers from an identification problem if x includes one or more dummy variables: the attribution of the component to specific variables is not invariant to the choice of reference category for the dummy variables. We adopt the solution proposed by Yun (2005) which is equivalent to calculating the decomposition with all possible regressions with different reference categories and averaging the results.

3 Results

3.1 Descriptives

The sample has broadly similar characteristics to the GP population (Department of Health, 2004) in terms of gender (34% female against 38% in the GP population), PMS status (33% against 37%), and working in a dispensing practice (18% against 16%).

Table 1 has descriptive statistics for the estimation sample. The mean annual gross earnings for male respondents is £81,313 while the mean earnings among female respondents is £57,312 or about 70% of mean male earnings.[2] Male GPs the work 48.4 hours per week and female GPs 36.7 hours per week on average, or about 76% of male hours. Female GPs have a hourly wage which is about 93% of the male wage.

Male GPs have slightly more experience than female GPs (22.9 vs. 20.1 years) and are more likely to be non-white, qualified outside the UK and on a salaried contract. They are also more likely to be married with a non-working partner but there is no substantial difference in the likelihood of not having children.

Male GPs are more likely to be single handed and they work in practices with a smaller proportion of female GPs. Their practices have a larger total list and a larger list per GP. While male GPs tend to work in practices with a higher share of patients living in nursing homes, male and female GPs work in similar practices in terms of the patients' age/sex mix and degree of morbidity as measured by the additional needs adjustment. Male GPs tend work in practices with somewhat smaller LISI scores but more likely to work in a practice located in a ward with a high share of people with no qualifications. They are also more likely to work in a dispensing and PMS practice.

3.2 Model results: hours

Tables 2 to 5 report the regression models. All models have heteroscedasticity robust standard errors. Table 3 presents the first stage of the two-stage least squares model in which the log of weekly hours worked is regressed against the full set of exogenous variables and the family circumstance instruments. The instrumental variables are jointly significant in both of the hours regressions, indicating that they satisfy one of the requirements for validity (Baum et al., 2003), although the variables have larger coefficients and are more significant in the model for female hours worked. Both of the models (see the second stage log earnings results in Table 4) pass the robust Sargan test for overidentifying restrictions (Baum et al., 2003), suggesting that the instruments are uncorrelated with the error term. They also pass the Pesaran and Taylor Reset functional form test (Pesaran and Taylor, 1999). Also reported are the Pesaran and Smith R^2 criterion for instrumental variable models (Pesaran and Smith, 1994).

The pattern of effects of the exogenous variables on log hours are similar for male and female GPs. The most obvious, and puzzling, difference is that the number of patients per GP has a much larger positive and highly significant effect on the hours of female GPs.

The family circumstance instruments have the same qualitative and plausible effects on male and female GPs but the magnitude of the effects is greater for female GPs. Having a non-working partner leads female GPs to increase their hours by proportionately much more than male GPs. Similarly the negative effect of having dependent children is much larger for female GPs.

[2] These figures are based on the midpoint of the reported income bracket. The 0.7% of the respondents earning £150.000 or more are classified as earning £175.000.

3.3 Model results: wages and earnings

Tables 2, 4 and 5 presents the estimation results for the log wage and log earnings models. All the models pass the RESET functional form test, except the female reduced form log earnings model which has a p value of 0.04.

The results for coefficients which are common across the log wage, 2SLS log earnings and the reduced form log earnings models are qualitatively similar in terms of signs and significance.[3] GPs in practices with larger lists earn more: the effects of economies of scale seem to more than offset the attenuation of incentives. The number of patients per GP has a positive but insignificant effect. A higher LISI score, indicating a more deprived population, is associated with lower earnings. The contractual age/sex adjustment has a negative effect. Since the contract gives practices a higher average capitation payment the higher its age/sex adjustment, these results suggest that the higher revenue is more than offset by higher costs associated with more demanding populations. The nursing home and additional needs adjustments have positive but insignificant effects. Being in a dispensing practice or a PMS practice has large positive and significant effects on log wages and earnings. Being salaried reduces GPs wages/earnings substantially, although the effect is only significant in the reduced form model and only for female GPs.

Other results are less easy to interpret. In none of the models is experience or its square significant and in some the relationship is U shaped and in others reverse U shaped. It may be that years since qualification is a poor measure of relevant experience, especially for GPs with children, or GPs from overseas. Being non-white has a very small negative and insignificant effect on male GP wages/earnings but larger positive effect on female GP earnings which is significant at 5% in the 2SLS earnings model and at 10% in the reduced form.

The coefficients for log hours in the second stage log earnings regressions in Table 4 are less than one in both cases, suggesting that earnings are not proportional to hours worked. A 1% increase in hours worked leads to a 0.45% increase in earnings for male GPs and a 0.75%. increase for female GPs. Because of the imprecise estimate of the hours coefficient in the male earnings model, the hypothesis of proportional returns to hours worked cannot be

[3] They are also broadly similar to those in a recent study (Morris et al, 2006) which uses the Inland Revenue tax records of all GPs in England and which therefore has a much larger sample (yielding more significant coefficients) and more accurate data on earnings. Morris et al (2006) estimate log wage models with the wage calculated using predicted hours derived from a model estimated on the 2004 NPCRDC GP Worklife Survey sample used in the present study.

rejected at conventional significance levels. In the female GP model the hypothesis of proportional returns can be rejected at the 10% level of significance. Models which included the log of hours and its square (both instrumented by family circumstances) had insignificant coefficients on both terms in both models, though the effect of log hours was positive and declining for female GPs and negative and increasing but very small for male GPs over the range of observed hours.

These results suggest that the log wage models are perhaps less well specified than the log earnings models because of their implicit assumption that earnings are proportional to hours.

3.4 Model results: tests for within practice discrimination

We suggested in section 1 that it would be possible to test for within practice discrimination via the effects of the single handed dummy and the proportion of female GPs. The effect of the proportion of female GPs female wages/earnings is negative in all but one model (male GPs log wage) and always insignificant. Since the proportion of female GPs may have negative effects on total practice income, for example if female GPs are less entrepreneurial or devote a smaller share of their hours to generating income, a better test for discrimination is to compare the coefficients of the variable in the male and female models. Allowing for the effect of gender mix on total practice profits we expect that if there is discrimination against female GPs in income sharing, the coefficient on the proportion of female GPs will be less negative than for male GPs. The coefficient on the proportion of female GPs is in fact more negative for female GPs in all the wage/earnings models. Thus the coefficients on the proportion of female GPs lends no support to the hypothesis of discrimination against female GPs.

We argued in section 1 that the coefficients on single handedness would be positive, reflecting incentive effects (remember that we have also included the total practice population to allow for economies of scale), but that if there was discrimination against women, the single hander coefficient would be larger for female GPs. The coefficient on the single hander dummy is positive in all models for both males and females. It is also always larger for the female GPs models. In no cases are the coefficients significantly different from zero. However, there are a very small number of single handed female GPs (19 out of 598) so that the test is not highly powered.

We also estimated a 2SLS log earnings model (not reported) in which we pooled the male and female GPs and included, in addition the full set of explanatory variables, a female dummy and the interaction of the female dummy with the single hander dummy. The coefficient on

the female dummy variable was unsurprisingly negative (- 0.191) and highly significant (t = -4.03). But the coefficient on the interaction of the female GP and single hander dummies was large, positive (0.227) and had a p value of 0.07. These findings provide some limited support for the discrimination hypothesis.

3.5 Decomposition results

Table 6 reports the results of the decompositions of the differences in mean log wages and earnings based on the three models and using two weights to reflect alternative assumptions about the appropriate counterfactual no discrimination wage or earnings functions.

The difference in wages is 0.084 log-points and highly significant, as is the difference in earnings of 0.384 log-points. The decompositions are sensitive both to the model and to the assumption about the appropriate counterfactual. The log wage decomposition, which we argue is most likely to be misspecified, gives the measure of discrimination as 89% with a male reference group and 61% with a female reference group. The 2SLS model gives the lowest estimate (60% or 23%). The reduced form log earnings model, which attributes some of the variation in hours in the 2SLS model to differences in family circumstance coefficients not included in the 2SLS model, yields discrimination measures of (90% and 43%).

4 Tentative conclusion

This is work in progress we finish with some very tentative conclusions and a description of possible further work.

We suspect that the 2SLS model should be the preferred model for constructing a summary discrimination measure since the log wage model is misspecified in its treatment of hours and the reduced form model goes too far in allowing differences in coefficients in the hours model to enter the discrimination measure.

But using the proportion of the difference in mean log wages or earnings due to differences in regression coefficients as a summary measure of within practice discrimination does not seem useful:

- it is very sensitive to the choice of regression model and to the assumed counterfactual no discrimination wage/earnings equation.
- the contribution of some of the variables to the discrimination measure has no obvious theoretical justification
- even our preferred 2SLS model suggests that 3/5ths of the 40% excess of male over female earnings is due to discrimination. In the face of rather small and generally insignificant differences in the coefficients on most of the variables this seems too large.

In further work we plan to draw on theoretical models of partnership sharing agreements to suggest alternative specifications of the remuneration function. We hope that this will produce some guidance on which coefficient differences should be included in the discrimination measure and suggest some further direct tests for discrimination.

References

Baum, C.F., Schaffer, M.E., Stillman, S. Instrumental variables and GMM: Estimation and testing. The Stata Journal 3, 2003: 1-31.

Blinder A. Wage discrimination: Reduced form and structural estimates. Journal of Human Resources 8, 1973: 436-455.

Cotton J. On the decomposition of wage differentials. Review of Economics and Statistics 70, 1988: 236-243.

Department of Health. Statistics for General Medical Practitioners in England 1994-2004.www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsStatistics/PublicationsStatisticsArticle/fs/en?CONTENT_ID=4106726&chk=nKjsm8. Accessed 20 December 2005.

Gaynor, M and Pauly, M.V. Compensation and productive efficiency in partnerships: evidence from medical groups practice. Journal of Political Economy, 98, 1990, 544-573.

Lloyd, D.C.E.F., Harris, C.M., Clucas, D.W. Low income scheme index: a new deprivation scale based on prescribing in general practice. British Medical Journal, 310, 1995: 165-169.

McNabb, R. and Wass, V. Male-female earnings differentials among lawyers in Britain: a legacy of the law or current practice. Labour Economics 13, 2006: 219-235.

Mincer J. The distribution of labour incomes: A survey with special reference to the human capital approach. Journal of Economic Literature 8, 1970: 1-26.

Mincer J. Schooling, experience and earnings. National Bureau of Economic Research. Columbia University Press, 1974.

Morris, S., Sutton, M., Elliott, R., Gravelle, H., Hole, A., Ma, A., Sibbald, B. and Skåtun, D. Determinants of general practitioner wages. Mimeo, June 2006.

Oaxaca R. Male-female wage differentials in urban labour markets. International Economic Review 14, 1973: 693-709.

Oaxaca, R.L., and Ransom M.R. Identification in detailed wage decompositions. Review of Economics and Statistics 81, 1999: 154–57.

Pesaran M.H., Taylor L.W. Diagnostics for IV regressions. Oxford Bulletin of Economics and Statistics 61, 1999: 255-281.

Pesaran H.M. and Smith R.J. A generalized R^2 criterion for regression models estimated by the instrumental variables method. Econometrica 62, 1994: 705-710.

Reimers C. Labor Market discrimination against hispanic and black men. Review of Economics and Statistics 65, 1983: 570-579.

Whalley D, Bojke C, Gravelle H, Sibbald B. 2004 National Survey of General Practitioner Job Satisfaction. Final Report. December 2004. www.npcrdc.man.ac.uk/Publications/GP%20Satisfaction%20REPORT.pdf?CFID=792808&CFTOKEN=37289229

Wood, R.G., Corcoran, M.E. and Courant, P.N. Pay differences among the highly paid: the male-female earnings gap in lawyers' salaries. Journal of Labour Economics, 11, 1993: 417-441.

Yun, M-S. A simple solution to the identification problem in detailed wage compositions. Economic Inquiry 43, 2005: 766-772.

Table 1. Descriptive statistics

Variable	Male GPs Mean	SD	Female GPs Mean	SD	
Annual earnings in £`000s	81.313	22.675	57.312	22.244	**
Weekly hours worked	48.409	11.169	36.664	11.938	**
Decades of experience	2.293	0.697	2.010	0.707	**
Non-white	0.158		0.112		*
Non-UK qualified	0.120		0.089		*
Salaried	0.028		0.020		
Married – partner does not work	0.229		0.067		**
No children	0.326		0.321		**
One child under 18	0.159		0.135		
Two children under 18	0.295		0.356		
Three or more children under 18	0.221		0.187		
Single-handed GP	0.059		0.032		
Proportion of female GPs in practice	0.308	0.189	0.500	0.191	**
Low Income Scheme Index	9.288	5.888	9.453	6.179	
Proportion ward with no qualifications	0.286	0.094	0.275	0.095	
Practice list size in `000s	8.948	4.365	8.643	4.051	
Practice list size per WTE GP in `000s	2.024	0.471	1.953	0.378	*
Age/sex adjustment	22.420	0.914	22.281	0.879	**
Nursing home adjustment	24.889	24.036	15.203	15.623	**
Additional needs adjustment	96.549	10.723	96.558	10.808	
Dispensing practice	0.192		0.169		
PMS practice	0.342		0.321		
Government Office Region:					
North East	0.042		0.048		
North West	0.133		0.132		
Yorkshire	0.092		0.114		
East Midlands	0.092		0.062		
West Midlands	0.094		0.080		
East	0.115		0.095		
London	0.120		0.159		
South East	0.171		0.192		
South West	0.142		0.117		
N	1179		598		

Significance levels for differences in means or proportions: *p<0.05, **p< 0.01; t tests except Chi square for numbers of children and GOR

Table 2. Wage regressions. Dependent variable is log hourly wage.

Variable	Male GPs Coefficient	t	Female GPs Coefficient	t
Decades of experience	-0.019	-0.21	-0.174	-1.24
Decades of experience squared	0.010	0.53	0.046	1.37
Non-white	-0.039	-1.13	0.110*	2.22
Non-UK qualified	-0.024	-0.66	-0.015	-0.24
Salaried	-0.068	-1.38	-0.091	-0.87
Single-handed GP	0.021	0.31	0.087	0.60
Log proportion female GPs in practice	0.002	0.38	-0.013	-0.31
Log LISI score	-0.083**	-2.68	-0.047	-1.11
Log proportion ward with no qualifications	-0.012	-0.33	-0.069	-1.25
Log practice list size	0.144**	5.48	0.102**	2.63
Log practice list size per WTE GP	0.022	0.36	-0.010	-0.12
Log age/sex adjustment	-1.160**	-3.71	-0.873	-1.76
Log nursing home adjustment	0.007	0.69	-0.003	-0.25
Log additional needs adjustment	0.257	1.29	0.289	0.93
Dispensing practice	0.141**	5.53	0.148**	3.43
PMS practice	0.097**	4.49	0.087*	2.51
North West	0.035	0.73	0.138	1.73
Yorkshire	0.062	1.26	0.193*	2.32
East Midlands	0.068	1.27	0.094	1.03
West Midlands	0.084	1.72	0.163	1.75
East	0.087	1.64	0.119	1.22
London	0.079	1.43	0.140	1.40
South East	0.029	0.57	0.125	1.44
South West	0.069	1.25	0.115	1.35
Constant	5.700**	4.21	4.619	1.94
N	1185		602	
R-squared	0.121		0.096	
RESET test (p-value)	0.28 (0.84)		1.07 (0.36)	
F-test for the joint significance of the GOR Dummies (p-value)	0.88 (0.53)		0.95 (0.47)	

*p<0.05, **p< 0.01 Robust SEs. Omitted GOR: North East.
Hourly wage: annual earnings/47*usual hours.

Table 3. Hours regressions. Dependent variable is log weekly hours.

Variable	Male GPs Coefficient	t	Female GPs Coefficient	t
Decades of experience	0.206**	2.71	0.127	1.08
Decades of experience squared	-0.055**	-3.20	-0.033	-1.18
Non-white	0.032	1.18	-0.017	-0.31
Non-UK qualified	0.042	1.46	-0.034	-0.51
Salaried	-0.033	-0.82	-0.146	-1.32
Married – partner does not work	0.029	1.55	0.179**	3.47
One child under 18	-0.062*	-2.44	-0.082	-1.71
Two children under 18	-0.045*	-1.98	-0.183**	-4.84
Three or more children under 18	-0.054*	-2.01	-0.259**	-6.14
Single-handed GP	0.076	1.50	0.057	0.49
Log proportion female GPs in practice	-0.007**	-2.64	-0.077	-1.88
Log LISI score	0.001	0.06	0.005	0.15
Log proportion ward with no qualifications	0.031	0.89	0.026	0.62
Log practice list size	0.003	0.12	-0.056	-1.68
Log practice list size per WTE GP	0.032	0.65	0.321**	4.63
Log age/sex adjustment	0.451	1.86	-0.169	-0.41
Log nursing home adjustment	0.019*	2.06	0.039**	3.05
Log additional needs adjustment	-0.250	-1.66	0.145	0.59
Dispensing practice	0.019	0.99	0.028	0.69
PMS practice	-0.007	-0.41	-0.058	-1.95
North West	-0.023	-0.49	0.045	0.76
Yorkshire	-0.007	-0.15	0.057	0.89
East Midlands	-0.013	-0.29	0.123	1.76
West Midlands	-0.015	-0.32	0.003	0.04
East	-0.009	-0.18	0.123	1.70
London	-0.017	-0.33	0.118	1.60
South East	0.033	0.69	0.111	1.79
South West	-0.011	-0.21	0.050	0.81
Constant	3.390**	3.17	3.143	1.63
N	1179		598	
R-squared	0.065		0.183	
F-test for joint significance of excluded instruments (p-value)	2.39 (0.05)*		12.47 (0.00)**	

*p<0.05, **p< 0.01 Robust SEs. Omitted GOR: North East

Table 4. Earnings regressions – 2SLS with instrumented hours. Dependent variable is log annual earnings

	Male GPs		Female GPs	
Variable	**Coefficient**	**t**	**Coefficient**	**t**
Log of hours worked	0.425	1.21	0.742**	5.04
Decades of experience	0.079	0.86	-0.181	-1.39
Decades of experience squared	-0.015	-0.69	0.051	1.59
Non-white	-0.022	-0.72	0.118**	2.65
Non-UK qualified	-0.003	-0.08	-0.018	-0.33
Salaried	-0.083	-1.57	-0.121	-1.39
Single-handed GP	0.068	1.14	0.113	0.90
Log proportion female GPs in practice	-0.003	-0.68	-0.029	-0.68
Log LISI score	-0.083**	-3.40	-0.045	-1.12
Log proportion ward with no qualifications	0.008	0.22	-0.062	-1.19
Log practice list size	0.147**	6.86	0.093*	2.49
Log practice list size per WTE GP	0.038	0.76	0.066	0.74
Log age/sex adjustment	-0.889**	-2.89	-0.821	-1.81
Log nursing home adjustment	0.017	1.72	0.007	0.51
Log additional needs adjustment	0.114	0.59	0.332	1.15
Dispensing practice	0.151**	6.96	0.157**	3.72
PMS practice	0.093**	5.27	0.067*	1.97
North West	0.019	0.43	0.155	1.92
Yorkshire	0.056	1.30	0.213*	2.47
East Midlands	0.059	1.27	0.125	1.34
West Midlands	0.072	1.69	0.167	1.80
East	0.080	1.78	0.151	1.54
London	0.071	1.42	0.176	1.81
South East	0.045	0.97	0.149	1.70
South West	0.058	1.27	0.126	1.46
Constant	4.549**	2.62	2.035	0.94
N	1179		598	
Pesaran and Smith R-squared	0.179		0.186	
Pesaran and Taylor RESET test (p-value)	0.35 (0.56)		0.02 (0.88)	
F-test for joint significance of GOR dummies (p-value)	0.93 (0.49)		1.00 (0.43)	
Robust Sargan test (p-value)	1.07 (0.78)		0.56 (0.91)	

*p<0.05, **p< 0.01 Robust SEs. Omitted GOR: North East

Table 5. Earnings regressions – reduced form. Dependent variable is log annual earnings.

Variable	Male GPs Coefficient	t	Female GPs Coefficient	t
Decades of experience	0.171	1.72	-0.086	-0.67
Decades of experience squared	-0.038	-1.65	0.027	0.85
Non-white	-0.006	-0.19	0.105	1.95
Non-UK qualified	0.017	0.54	-0.043	-0.69
Salaried	-0.097	-1.52	-0.228**	-3.06
Married – partner does not work	0.008	0.38	0.116	1.92
One child under 18	-0.040	-1.44	-0.044	-0.87
Two children under 18	-0.013	-0.54	-0.125**	-2.88
Three or more children under 18	-0.009	-0.33	-0.204**	-3.92
Single-handed GP	0.099	1.90	0.162	1.52
Log proportion female GPs in practice	-0.006	-1.87	-0.085	-1.75
Log LISI score	-0.082**	-3.54	-0.041	-0.91
Log proportion ward with no qualifications	0.019	0.47	-0.041	-0.77
Log practice list size	0.148**	6.75	0.055	1.26
Log practice list size per WTE GP	0.051	1.04	0.305**	3.63
Log age/sex adjustment	-0.692*	-2.52	-0.927*	-2.00
Log nursing home adjustment	0.025**	2.79	0.036**	2.48
Log additional needs adjustment	0.011	0.06	0.439	1.51
Dispensing practice	0.159**	7.37	0.178**	3.46
PMS practice	0.089**	5.12	0.023	0.69
North West	0.008	0.17	0.188*	2.08
Yorkshire	0.053	1.11	0.256*	2.50
East Midlands	0.053	1.08	0.216	2.02
West Midlands	0.065	1.38	0.168	1.60
East	0.077	1.57	0.240*	2.24
London	0.062	1.07	0.262*	2.63
South East	0.059	1.11	0.232*	2.37
South West	0.052	1.03	0.163	1.65
Constant	5.948**	4.71	4.296*	2.03
N	1179		598	
R-squared	0.180		0.187	
RESET test (p-value)	0.45 (0.72)		2.83 (0.04)	
F-test for joint significance of the GOR dummies (p-value)	0.81 (0.59)		1.22 (0.28)	

*p<0.05, **p< 0.01 Robust SEs. Omitted GOR: North East

Table 6. Oaxaca decompositions based on wage and earnings regressions

	Log wage		Log earnings – 2SLS		Log earnings – reduced form	
	Mean	**SE**	**Mean**	**SE**	**Mean**	**SE**
Total difference	0.0843	0.0173	0.3840	0.0156	0.3840	0.0171
Weight = 1 (male ref group)						
Due to diff in variables						
Hours worked			0.1279	0.1057		
Personal characteristics	0.0039	0.0048	0.0031	0.0040	0.0019	0.0044
Family characteristics					0.0009	0.0035
Single-handed GP	0.0006	0.0019	0.0019	0.0016	0.0027	0.0014
Practice characteristics	0.0033	0.0089	0.0204	0.0125	0.0333	0.0083
Regional dummies	0.0016	0.0023	0.0008	0.0019	0.0004	0.0018
Total	0.0094	0.0095	0.1540	0.0970	0.0392	0.0092
Due to diff in coefficients						
Hours worked			-1.1226	1.3516		
Personal characteristics	0.1992	0.1652	0.2556	0.1498	0.1775	0.1606
Family characteristics					0.0460	0.0288
Single-handed GP	0.0309	0.0756	0.0211	0.0649	0.0294	0.0556
Practice characteristics	-1.0631	2.7222	-1.2841	2.4522	-1.4569	2.4529
Regional dummies	-0.0092	0.0083	-0.0098	0.0080	-0.0148	0.0089
Constant	0.9172	2.7251	2.3698	2.7772	1.5635	2.4483
Total	0.0750	0.0194	0.2299	0.0983	0.3447	0.0194
Unexplained share (discrimination)	89%		60%		90%	
Weight = 0 (female ref group)						
Due to diff in variables						
Hours worked			0.2230	0.0443		
Personal characteristics	0.0100	0.0080	0.0137	0.0080	0.0094	0.0081
Family characteristics					0.0187	0.0102
Single-handed GP	0.0024	0.0040	0.0031	0.0035	0.0045	0.0029
Practice characteristics	0.0224	0.0805	0.0598	0.0802	0.1880	0.0904
Regional dummies	-0.0019	0.0032	-0.0026	0.0031	-0.0033	0.0034
Total	0.0329	0.0805	0.2971	0.0781	0.2173	0.0914
Due to diff in coefficients						
Hours worked			-1.2178	1.4662		
Personal characteristics	0.1931	0.1669	0.2449	0.1506	0.1701	0.1615
Family characteristics					0.0282	0.0182
Single-handed GP	0.0291	0.0712	0.0198	0.0610	0.0276	0.0523
Practice characteristics	-1.0822	2.7342	-1.3235	2.4660	-1.6116	2.4663
Regional dummies	-0.0057	0.0073	-0.0065	0.0072	-0.0111	0.0081
Constant	0.9172	2.7251	2.3698	2.7772	1.5635	2.4483
Total	0.0514	0.0801	0.0868	0.0772	0.1667	0.0915
Unexplained share (discrimination)	61%		23%		43%	