Urban Effects on Participation and Wages: Are there Gender Differences?[1]

Euan Phimister[**]

Department of Economics and Arkleton Institute for Rural Development Research,

University of Aberdeen.

Centre for European Labour Market Research

Discussion Paper 2004-04

Editor: W David McCausland

www.abdn.ac.uk/business

[**] Postal address. Department Of Economics, Business School, University of Aberdeen, Edward Wright Building, Aberdeen AB24 3QY. United Kingdom. Email e.phimister@abdn.ac.uk. Fax 44 1224 272181.

ABSTRACT

This paper estimates participation and wage equations using panel data from the United Kingdom to explore differences in urban and rural wages and participation by gender. The results suggest a small but economically significant participation premium for urban women relative to rural female workers. Results from the wage estimations suggest that after controlling for sample selectivity, observed and unobserved heterogeneity, the wage premium received by urban women is larger than that obtained by men. Consistent with the hypothesis that poorer matching in less dense labour markets affects rural workers, there is also evidence of higher rural wage depreciation for both men and women, while returns to experience for rural men are also lower than for urban workers.

Keywords. Participation, wages, urban, rural, panel, sample selection.

Acknowledgements

The data used in this paper were made available through the ESRC Data Archive. The data were originally collected by the Institute for Social and Economic Research, University of Essex. Neither the original collectors of the data nor the Archive bear any responsibility for the analyses or interpretations presented here. I would also like to acknowledge financial support from the Leverhulme Trust (grant number F/152/Y), the help of Nick Buck, Alana Gilbert, Deborah Roberts and Esperanza Vera-Toscano, plus the comments and suggestions from the two anonymous referees. All usual caveats apply.

I. INTRODUCTION

Recent evidence suggests that various types of agglomeration externalities are important in explaining the observed urban wage premium. Wheaton and Lewis [31] find evidence of increasing urban wages consistent with increasing firm level returns and greater specialization of labor in denser urban markets. Glaeser and Maré [10] find a robust urban wage premium and show that rural-urban migrants experience higher wage growth consistent with greater learning spillovers in urban areas. Wheeler [32] provides empirical support for a model where larger labor market size increases job match quality, individual productivity and wages.

While the recent literature has explored how agglomeration influences the urban wage premium, there has been little consideration of whether these externalities impact differently on women and men. There are a number of reasons why such gender differences might exist. First, empirical evidence suggests that women are less spatially mobile than men, with job search conducted over a smaller area than men (Madden and Chiu [19]). In denser urban labor markets, improved urban job matching may counterbalance this reduced mobility and hence, one might expect that any urban wage premium would be larger for women than men. Second, women typically work fewer hours and years over their working lives and have a higher rate of job turnover than men (Altonji and Blank [2]). Such career interruptions are likely to reduce the effect of improved learning spillovers in urban markets and hence may reduce any urban wage growth effect. On the other hand, the more interrupted nature of female work histories may mean that women particularly benefit from improved urban job matching. As a result, the extent of wage depreciation suffered by women during periods out of employment may be less in denser urban markets.

By increasing the offered wage for any given individual, agglomeration effects on wages should also increase labor market participation. However, other density effects, especially on female participation, are also possible. In particular, population density allows greater provision of certain services, e.g. childcare provision, the availability of which are particularly likely to reduce the reservation wages of women. Indeed, female participation rates in urban areas in developed countries have been consistently higher than those observed in rural areas (Stabler [26]), while the lack of childcare facilities and access to transport are frequently cited as barriers to women accepting employment in rural areas (Porterfield [22], Lichter et al. [17]).

This paper explores gender differences in both urban wage and participation premiums using panel data from the United Kingdom. Specifically, using the British Household Panel Survey (BHPS), participation and wage equations are estimated for samples of urban and rural women and men and the extent of any urban premiums calculated. Differences in the structure of wages and participation are considered by testing for urban-rural differences in the impact of explanatory variables for women and men.

Higher urban wages and participation may also reflect living costs, differences in abilities, job characteristics or lower preferences for urban amenity (Roback [24], Hwang et al [14]) Therefore, when trying to identify agglomeration effects it is important to account for observed and unobserved differences in such factors. While many previous studies have been able to control for observed differences, their ability to control for unobserved differences in these factors has often been limited because the appropriate data was cross-sectional only (Wheeler [32], Wheaton and Lewis [31]). However, the availability of panel data does allow the estimation of models accounting for such unobserved differences or heterogeneity. For example Glaeser and Maré [10] used a fixed effects estimator to identify the male urban wage premium while controlling for unobserved heterogeneity.

This approach is less attractive if sample selection effects are important, as the fixed effects estimator is inconsistent if these effects do not vary with time (Verbeek and Nijman [30]). Clearly, the interrupted nature of female labor market participation means sample selection effects should be taken into account when considering the female urban wage premium. The fixed effect estimator also relies on urban-rural/rural-urban migrants for identification of the wage premium. This means that it is unsuitable if there are insufficient numbers of such migrants for credible estimation. For similar reasons, in this approach it is particularly difficult to identify urban-rural differences in parameters for explanatory variables such as education where there is little variation over time at the individual level.

Therefore, to account for sample selection and unobserved heterogeneity, while allowing urban-rural differences in the impact of explanatory variables to be identified, the panel sample selection model suggested by Vella and Verbeek [29] and Nijman and Verbeek [21] is estimated. In this model the correlation between unobserved components in the wage and participation equations controls for traditional sample selection effects and unobserved heterogeneity across all individuals. The identification of any urban premiums in this model does not rely on migrants but uses all available wage information.

The main results are as follows. There is evidence of a small but economically significant urban participation premium for women of around 3%. In contrast, there appears to be no economically significant participation premium for men. There is also evidence of gender differences in the structure of the urban wage premium. In particular, urban women experience lower wage depreciation associated with time out of the labor force, but their returns to experience are no higher. In contrast, there is evidence of both lower wage depreciation and higher returns to experience for urban men. The source of the female urban participation premium is difficult to determine. However, the urban wage premiums found are consistent with the hypothesis that those with lower spatial mobility are more disadvantaged in less dense labor markets, with for example, the wage premium substantially lower for single women than for those who are married or cohabiting.

The plan of the paper is as follows. Section II discusses in more detail the source and empirical implications of possible gender differences in agglomeration effects. Section III describes the data, the definitions used and provides some basic descriptive analysis. Section IV discusses the econometric specification, discussion of hypotheses and model estimation. Section V presents the results. Section VI concludes.

II. BACKGROUND

The various mechanisms through which agglomeration can increase wages (and participation) are likely to affect gender differences in a number of different ways. Arguably, agglomeration externalities that occur primarily at the firm level, are unlikely to induce large gender differentials. For example if urban firms are more productive because of lower set-up or transport costs, or through increased informational spillovers (Krugman [16], Glaeser [8]) there should be no gender differential as long as women and men are equally likely to be found in firms where such effects are important.

In contrast, agglomeration effects which accrue to the individual are more likely to be associated with gender differences. For example, because of the greater number of contacts between individuals, informal learning is likely to be greater in urban areas and individual productivity and wages in urban areas may grow faster as a result (Glaeser [9], Rauch [23]). However, women have historically had lower labor market attachment, higher job turnover, and have received less training than men (Altonji and Blank [2]). If learning spillovers increase the rate of human capital formation, the interrupted nature of the typical female work history and lower levels of training mean that urban women are likely to gain less from such spillovers than men and hence wage growth effects may be less evident for women.

On the other hand, women are likely to gain significantly from better job matching in denser markets. Various models predict higher wages and growth if the quality of job matching improves in urban markets. For example, Hesley and Strange [13] illustrate how agglomeration itself can be driven by increasing expected urban match quality, while Wheeler [32] shows how decreasing search costs lead to more productive matches, increased sorting and wages in urban labor markets. Evidence suggests that women are less spatially mobile than men, with job search conducted over a smaller area than men (Madden and Chiu [19]). Hence, denser markets and better job matching may counterbalance this reduced mobility. Further, improved urban job matching should also reduce the wage depreciation associated with periods out of the labor force (Mincer and Ofek [20]). The more interrupted nature of female work patterns should mean that this type of effect will be particularly important for women.

Although the evidence is mixed, a number of authors have argued that returns to education will be higher in denser markets. Possible gender effects in such education differentials are perhaps rather ambiguous. For example, Rauch [23] argues that higher returns to education in urban areas arise because the transmission of ideas is likely to improve with higher levels of human capital. Hence, these spillovers may be larger in urban areas where average education levels are higher. Such effects, which are external to the individuals, are unlikely to induce gender differences. However, Frank [7] argues that overeducation effects are more likely in less dense markets and that such effects are more likely to affect the 'second' earner in the household, i.e. typically women. In this case, any increase in the return to education associated with denser markets might well be larger for women.

If market size increases wages it should also increase participation for identical individuals. Moreover, if the female urban wage premium is larger, any associated increased participation effect should also be higher for women. Finally, the effect of population density on service provision may also tend to exacerbate any gender difference in the urban participation premium. In particular, in areas of low population density, access to services such as transport, housing and childcare, may be more difficult. If they exist, such barriers to employment are likely to impact differentially on women, increasing their reservation wages and hence further reducing female participation in low density areas (Porterfield [22], Lichter et al [17]).

In summary, the previous discussion suggests a number of general hypotheses concerning the nature of possible gender differences in urban wage and participation premiums. First, if agglomeration effects are only felt at the firm level, the urban wage premium is likely to be similar for both sexes. In contrast, where job matching effects are important, the urban wage premium may be larger for women as denser markets counterbalance the effect of their smaller job search area. Further, the urban wage premium should, as a result, be larger for certain groups whose spatial mobility is likely to be particularly restricted, e.g. married women.

Learning spillover and job matching effects are both likely to be important in the male urban wage premium. However, the interrupted nature of their work histories means that job matching effects, e.g. through lower wage depreciation, are likely to be relatively more important than learning spillover effects for women. In contrast, it is more difficult to make predictions whether, if they exist, higher urban returns to education should vary by gender. Finally, any larger female urban wage premium should also imply a larger urban participation premium for women. This differential may also be exacerbated by higher service provision in denser urban areas.

III. DATA

The data was drawn from the first eight waves 1991-1998 of the British Household Panel Survey, a longitudinal survey following some 10,000 individuals representative of the British population. The labour market component of the survey has detailed information on individual earnings, hours worked, other individual characteristics and work histories from which standard measures of usual hourly earnings, highest education level attained, total experience and time out of the labour force can be calculated (see the footnote to Table 1 for details on definitions).

Unlike in the United States there is no single accepted definition of what constitutes a (dense) metropolitan area in the UK. However, additional information made available by UK Institute for Social and Economic Research, made it possible to split the sample into urban and rural residents based on place of residence consistent with those definitions used by policy makers. Specifically, in England, Local Authority Districts are classified into Remote Rural, Accessible Rural, Coalfield areas, Urban and Metropolitan (Cabinet Office [4], Tarling et al [27]). For Scotland and Wales, rural Local Authority Districts are identified using the Randall definition, where population density in the district is less than one person per hectare, and then rural districts are classified as remote or accessible rural depending on their proximity to urban centers (Scottish Office [25]).

While these definitions do not allow the examination of city size effects, it is argued that the contrast between urban and rural samples can still be exploited to identify agglomeration effects. To ensure those in the rural sample are resident in areas characterized by population scarcity and distance from urban centers (Cabinet Office [4]), it is important to exclude those living in rural areas but within commuting distance to urban centers. Hence, the rural sample consists only of those individuals resident in remoter rural districts only, while the urban sample includes only those resident in districts defined to be urban and metropolitan.

To estimate the sample selection model specified in the next section requires at least three (consecutive) observations for each individual in the sample. To maximize the use of the available data and in particular to ensure adequate rural observations, an unbalanced panel was constructed from the rural and urban samples of individuals interviewed in three or more consecutive waves. For each individual, only information from one set of consecutive interviews was used. Hence, if data was missing in a given wave, only the information from the longest set of consecutive of interviews was included for that individual. This procedure resulted in repeated observations on 2177 women, of which 1040 were observed in all eight waves, and 1626 repeated male observations, with 565 observed in all eight waves. Moves between urban and rural locations are extremely rare in the data, with 71 total moves for women and only 48 total moves for men in either direction.

Table 1 reports key summary statistics for the urban and rural samples by gender. As expected, the Table shows higher urban participation rates and wages for both sexes. The urban-rural differences in participation rates are rather similar for women and men, i.e. 4 and 3 percent respectively. However, the female urban wage premium at 15 percent is significantly larger than the male urban-rural difference at 8 percent. Apart from a few exceptions, e.g. a lower proportion of married or cohabiting urban women, a higher proportion of urban men with a degree, most of the characteristics appear similar across the urban and rural samples.

IV. MODEL

Urban premiums in wages and participation may arise from differences in observed and unobserved individual characteristics or in differences in the impact of given explanatory variables on an individual's wages and participation probability. The following model captures both types of effect. First, consider the offered wage equation

$$(1) \qquad w_{it} = x_{it}^{u}\beta^{u} + x_{it}^{r}\beta^{r} + \alpha_i + e_{it},\ i=1,..,N,\ t=1,..,T_i$$ [1]

where w_{it} represents the potential offered (log) wage of individual i in time t, x_{it}^{k} ($k=u,r$) is a vector of observed characteristics for the urban (u) and rural (r) samples ($x_{it}^{k}=0$ if individual i is not part of sample k), β^{k} are returns to these characteristics in the two samples, while α_i and e_{it} are random components. In the model estimated below x_{it}^{k} contains quadratic functions of each individual's total experience and time out of the labor force, education dummies plus time and regional dummies. Differences in the structure of returns to characteristics across the two samples can be considered in equation (1) by testing whether the parameters β^{k} are identical. Unobserved heterogeneity across individuals in terms of cost of living, productivity, preference for amenity and other job characteristics, e.g. industry and occupation, is controlled for by the random effect term α_i, while e_{it} accounts for other time varying

[1] In the empirical work, separate equations are estimated by gender. For brevity the specification described below does not explicitly distinguish between the sexes.

random shocks.

Assume, as is standard, that an individual's decision to work is determined by whether the offered wage is above their reservation wage, where both are influenced by observed and unobserved individual characteristics. This implies a standard reduced form model for participation as a function of all the variables affecting both the reservation and offer wage plus any individual unobserved effect. However, there is extensive evidence to suggest that even after accounting for individual heterogeneity, individuals also exhibit a considerable degree of persistence in their labor market state (Heckman [11], Hyslop [15]). This suggests that the reduced form model should also allow for state dependence as follows

(2) $$y_{it}^{*} = \delta_0^u y_{it-1}^u + \delta_0^r y_{it-1}^r + z_{it}^u \delta_1^u + z_{it}^r \delta_1^u + \theta_i + \eta_{it}$$

(3) $$y_{it} = I(y_{it}^{*} > 0)$$

where y_{it} is equal to one when individual participates in period *t,* y_{it}^{*} is the latent variable which when positive implies the individual will participate, z_{it}^k are vectors containing the variables assumed to influence both offered and reservation wages for the urban (*u)* and rural (*r)* samples. The vector δ_1^u captures the (net) effects of the variables in the reservation and offered wage functions by location. In the specification below, z_{it}^k contains all the explanatory variables used in x_{it}^k plus a set of demographic and other variables, including martial status, number and age of children and non-labor income. The extent of state dependence or persistence in participation is captured by the presence of the lagged participation variables y_{it-1}^k (equal to one if the individual was working in the previous period and resident in location *k*). Unobserved heterogeneity is captured by θ_i while η_{it} is pure random component.

The model is completed by assuming that the error terms are jointly normally distributed with zero means and constant variances. The sample selection problem induced by the potential correlation between unobserved components in the participation and wage equations can then be incorporated by allowing for non-zero covariances between the random effects in the α_i and θ_i, and between the two shocks e_{it} and η_{it}. All other covariances between elements of the error terms are assumed zero.[2]

Testing Urban-Rural Differences

The sources of possible gender differences in urban wage and participation premiums suggest a number of differences in the urban and rural coefficients in equations (1) and (2). First, as Glaeser and Maré [10] note, firm level agglomeration effects imply an urban wage level effect. Hence, any wage premium for both sexes should disappear once observed and unobserved heterogeneity is controlled for, i.e. $H_o : \beta^u = \beta^r$ should not be rejected in equation (1). Moreover, if an urban wage premium is observed it should not differ significantly by gender. In contrast, job matching effects should mean any overall urban wage premium is larger for women, and that this premium will be greater in samples where spatial mobility is thought to be lower.

Learning spillover effects should increase urban wage growth. Improved job matching effects may also increase wage growth but equally should decrease wage depreciation associated with time out of the labor force. Disentangling these effects is obviously difficult. First, the exact relationship between any learning spillover and job matching effects and the estimating equations is not clear-cut. Second, the small number of migrants in the dataset means it not possible to use this group to identify

[2] The sample selection model controls for labor market participation but not the potential endogeneity of location. However, informal tests using sub-samples of 'less mobile' individuals, e.g. those with lower education levels, indicate that the results are robust to this limitation.

the source of the urban premiums (Glaeser and Maré [10]). However, the lack of urban-rural mobility does suggest that that any urban premium in wage growth and/or reduced wage depreciation should be reasonably reflected in the different estimates of returns to experience and wage losses associated with time out of the labor force. For example, if improved job matching effects are more important for women than learning spillovers, any urban wage growth effect captured through increased returns to experience should be less important than any reduced urban wage depreciation effect.

There are no clearcut predictions as to whether there will be any difference in any urban educational premium by gender. However, any such effects will be captured by differences in the coefficients reflecting returns to education in the wage equation (1).

Any urban impacts on returns to experience, wage depreciation and education will also affect the coefficients in the participation equation. However, as these variables may also affect the reservation wage, differences in these coefficients are more difficult to interpret directly. However, after accounting for other factors, the impact of a higher female urban wage premium should mean a greater urban female participation effect.

The other hypothesized participation effects that can be captured by equation (3) are also likely to reinforce any higher urban female participation effect. In particular, for women, the presence of (young) children is likely to increase reservation wages, and therefore decrease the probability of labor market activity, by increasing the opportunity costs of working. If population density effects on childcare service provision are important, this reduction in reservation wages may be lower for urban women. Hence, the urban coefficients capturing the overall impact of children on participation on women should be less negative than the rural ones.

Econometric Implementation

Estimation of the wage and participation equations poses a number of econometric problems. As specified, the estimation model would need to assume that the random effect θ_i is independent of z_{it}. This assumption is not tenable for a number of the explanatory variables. This poses a problem as, if this assumption is violated, the estimated coefficients will not correctly identify the marginal impacts of the time varying independent variables. The standard approach to control for this is to model the random effect θ_i as a function of all the independent variables in all time periods (Chamberlain [5]) or, slightly more restrictively, as a simple function of the individual level means of the independent variables (Arulampalam et al [3]). This latter approach is followed here. Hence,

(4) $$\theta_i = \bar{z}_i^u \phi^u + \bar{z}_i^r \phi^r + \mu_i$$

where μ_i is assumed independent of z_{it}^k and the correlations between the original random effect and the regressors may vary by location.[3]

After substitution of (4) in (2), the estimation of the resulting participation and wage equations is undertaken using the two-step procedure suggested by Nijman and Verbeek [21]), and Vella and Verbeek [29]). In the first step, estimates of the parameters in the participation equation (2) are obtained from a dynamic random effects probit. One additional problem is that the presence of unobserved heterogeneity μ_i in conjunction with lagged participation y_{it-1} induces an initial

[3] One problem of this approach is that significant collinearity problems may be induced between the additional regressors and the original independent variables. Hence, in the empirical work only a subset of the independent variables are used, i.e. those with sufficient variation within the panel.

conditions problem, where the initial observation y_{it0} will be correlated with the unobserved random effect and hence maximum likelihood will produce inconsistent estimates. This is addressed using the suggestion of Heckman [12], where a reduced form equation is specified for the initial period, where the error term from this equation is assumed correlated to the unobserved random effects.

(5) $y_{it0} = z_{it0}\lambda + \varepsilon_{oi}$

where z_{it0} is a vector of strictly exogenous variables, ε_{oi} is the error term, where this error and the random effect μ_i are correlated.

The second step of the estimation procedure is an extension of the standard Heckman approach dealing with sample selection. Consider the conditional expectation of equation (1), conditional on the y_i, the vector of all participation states observed for individual *i*,

(6) $$E(w_{it} \mid y_i) = x_{it}^u \beta^u + x_{it}^r \beta^r + E(\alpha_i \mid y_i) + E(e_{it} \mid y_i).$$

As the errors are assumed to be drawn from a multivariate normal distribution, the conditional expectations $E(\alpha_i \mid y_i)$ and $E(e_{it} \mid y_i)$ are linear functions of the covariances $\sigma_{\alpha\mu}$ and $\sigma_{e\eta}$. Specifically, Verbeek and Nijman [30] show that

(7) $$E[\alpha_i \mid y_i] = \sigma_{\alpha\mu}\left[\frac{1}{\sigma_\eta^2 + T_i\sigma_\mu^2}\sum_{s=1}^{T} a_{is} E[\mu_i + \eta_{is} \mid y_i]\right]$$

(8) $$E[e_{it} \mid y_i] = \sigma_{e\eta}\left[\frac{1}{\sigma_\eta^2}\left(E[\mu_i + \eta_{it} \mid y_i] - \frac{1}{\sigma_\eta^2 + T_i\sigma_\mu^2}\sum_{s=1}^{T} a_{is} E[\mu_i + \eta_{is} \mid y_i]\right)\right]$$

where $T_i = \sum_{s=1}^{T} a_{is}$ is the number of periods an individual is observed ($T = \max_i(T_i)$, $a_{is} = 1$ if *i* observed in period *s*, *0* otherwise), and $\sigma_{\alpha\mu}$ $\sigma_{e\eta}$ are the covariances between the random effects and error terms respectively. It can be shown that the bracketed terms on the right hand side of (7) and (8) are functions of the parameters in the participation equation only. Hence, as in the standard Heckman case, once estimates of the participation parameters have been obtained, estimates of these correction terms can be obtained via numerical integration. Then equation (6) can be estimated including the two correction terms using OLS, where standard errors are adjusted to allow for the estimated nature of the correction terms. The coefficients provide estimates of $\sigma_{\alpha\mu}$ and $\sigma_{e\eta}$, and therefore the significance of these two coefficients provides a test for the importance of sample selection effects.

In principle, the error assumptions identify all the parameters in the offered wage equation as the selection terms are non-linear functions of the exogenous variables. However as Vella [28] notes, the degree of the non-linearity in the selection terms may be limited given the actual range of values of the regressors. Hence, further exclusion restrictions are desirable. Here, a number of variables are excluded from the wage equation, namely, the demographic variables and non-labor income, plus the lagged participation value. It is often argued that household demographic variables should be included to capture unobserved motivational factors. Here such factors are controlled for by the unobserved heterogeneity. Even without these restrictions, the wage regression coefficients can be identified by the exclusion of the lagged participation variable. [4]

V. RESULTS

Participation

The results of the estimation of the participation equation for women and men are

[4] Experimentation shows that the results are robust to a variety of exclusion restrictions, including the case where lagged participation is the only explanatory variable in the participation equation excluded from the wage regression.

reported in Tables 2 and 3 respectively. For comparative purposes each Table also reports (in Columns 1 and 2) the results from a simple static Random effects probit with no adjustment for potential correlations between the regressors and the random effect. The urban and rural estimates from the full dynamic model (equations (2)-(4)) are reported in columns 3 and 4 respectively. Hence, comparing across the two specifications shows the impact on the participation estimates of allowing for dynamics and correlations between regressors and the random effects. For brevity in all cases, the estimated coefficients on the regional and wave dummies are not reported. In the dynamic participation equation, the initial conditions equation estimates and the adjustment for correlated errors given by equation (4) are also omitted. Unless otherwise indicated, the standard error is given in brackets below each estimated coefficient.

The second section of Tables 2 and 3 reports the value of the inter-temporal correlation coefficient for the participation equations. The correlation coefficient $\hat{\rho}$ is typically interpreted as the proportion of the variance unexplained by the regressors in the random effects probit and accounted for by variation between individuals (Arulampalam et al [3]). The associated standard error indicates whether taking account of unobserved heterogeneity in the participation equation is important. Consistent with previous studies, this correlation is significant for all specifications in both Tables 2 and 3.

In the final panel of the Tables, a number of estimation evaluation measures are reported. For all estimations, the hypothesis that all coefficients are zero is strongly rejected. Further, consistent with the existence of urban density effects, the hypotheses that all the urban and rural coefficients are identical ($H_o : \delta_0^u = \delta_0^r, \delta_1^u = \delta_1^r$) is rejected at 5 percent significance for all cases, except for the dynamic model in the female sample. Here, this hypothesis is rejected at 10 percent significance (p-value 0.056).

Finally, the urban participation premium is reported at bottom of each Table. This is derived for each of the estimations as follows. First, the urban and rural coefficient estimates are applied to the urban sample separately to generate two sets of individual participation probability predictions. The urban participation premium is the difference between the average of the predictions obtained using the urban coefficients and the average obtained using the rural coefficients. Formally, the average predicted probability from random effects estimation represents the (conditional) probability that a randomly chosen individual will be observed participating (Arulampalam, et al [3]). This premium can therefore be interpreted as the increase in the probability that a randomly chosen individual will participate if located in an urban area rather than a rural one. As predicted, the implied urban participation premium for women is positive (0.028-0.031) and larger than for men. Indeed, for men the 'premium' is either negative (-0.0007) or extremely small (0.008).

Turning to the coefficients for women presented in Table 2. For the urban women estimates (columns 1 and 3), most are broadly in line with prior expectations. In both the static and dynamic specifications, the participation probability is strongly positively associated with education level, and negatively associated with time out of the labor force, pre-school children and non-labor income. In the dynamic specification, the coefficient on the lagged participation variable is positive and significant. The inclusion of this variable substantially decreases the proportion of the unexplained variance attributed to individual heterogeneity. A number of the estimated coefficients also change substantially between the static and dynamic specification. For example, the marginal impact of total actual experience is now apparently negative (although insignificant), the education coefficients now exhibit a more distinct pattern, with increasing education associated more clearly with increasing participation probability, while the marginal negative impact of young children is much reduced. This latter result is consistent with the probability of women having children being correlated with preferences for not working. The majority of the rural coefficients are of same sign as the urban ones, with the education coefficients being a notable exception. However, unsurprisingly given the smaller underlying sample many of these coefficients are less well determined.

The coefficients on education, experience and time out of the labor force capture the effects of these variables on both offered and reservation wages. Hence, clearcut conclusions in terms of urban density effects are difficult to draw from any urban-rural differences for these variables. However, the main source of the rejection of the hypothesis that all the urban and rural coefficients are identical does appear to arise from the education variables. Two of the education variables are individually significantly different in both specifications, while the joint Wald test of equality of all education coefficients across the urban and rural samples is rejected in both the static and dynamic specifications. (p-values 0.004 and 0.028 respectively). Overall the urban-rural differences in the experience variables and time out of the labor force do not exhibit clear differences across the urban and rural sample. While the impact of all experience and time out of the labor force variables is statistically different across the two samples (p-value 0.020), this difference is not apparent in the dynamic specification.

Because these variables appear in the participation model only, stronger interpretations are possible if urban-rural differences in the impact of children are observed. For example, the marginal coefficients on young children are less negative for urban women in both specifications (consistent with lower urban reservation wages from better childcare provision). However, in neither the static or dynamic specification are these differences statistically significant, while the urban-rural differences in the coefficients on older children do not conform to this hypothesis.

The Table 3 estimates for the male sample follow a rather similar pattern to the female sample. Both urban and rural estimates generally follow prior expectations. Relative to the static model, the full dynamic specification with correlated regressors reduces the correlation coefficient $\hat{\rho}$, while none of the variables reflecting number of children is statistically significant in this latter specification. Both urban and rural coefficients on lagged participation variable are positive. In this case the urban coefficient is less than the rural one but this difference is not statistically significant.

Although difficult to interpret in terms of urban density effects because they are net effects, the overall rejection of equality of the urban and rural coefficients does appear to arise from differing impacts of the experience and out of the labor force variables. In the static specification, the equality of urban-rural coefficients of both the experience variables and the out of the labor force variables is rejected (p-values 0.08 and 0.001 respectively). In the dynamic specification, although there are no individually significant differences, the test that both coefficients on the out of the labor force variables are equal across the samples is still rejected at 10 percent (p-value 0.08). In contrast to women, there is little evidence of urban education effects. Although the coefficient on one of the education dummies is significantly different in the dynamic case, the joint test that all education coefficients are identical cannot be rejected in either specification.

Wages

Tables 4 and 5 present the Sample Selection model wage equation estimates for women and men. To provide an indication of the sensitivity of the results to the underlying assumptions used to identify this model, OLS and Fixed Effects wage equation results are also reported in these Tables. In particular, the OLS results provide a simple way to judge the effects of allowing for sample selection and unobserved heterogeneity using the Sample Selection model. In contrast, the Fixed Effects model accounts for unobserved heterogeneity but not all types of sample selection and relies on migrants to identify the urban effects.

There are two additional coefficients in the Sample Selection model associated with the two selection terms generated using the dynamic random effects probit results in Tables 2 and 3. These coefficients provide estimates of the covariance between the random effects, $\sigma_{\alpha\mu}$ and the covariance between the random shocks $\sigma_{e\eta}$.[5] The results indicate that sample selection is important for men and women. For both samples the estimate of $\sigma_{\alpha\mu}$ is positive, although it is only significant for women, while the estimates of $\sigma_{e\eta}$ are negative and significant at 5% for both sexes.[6]

In the bottom panel of the Tables, a number of estimation evaluation measures are reported. The tests that all coefficients are zero are strongly rejected in all specifications for both women and men. The test of equality between all the urban

[5] The first selection term also accounts for unobserved heterogeneity in the wage equation.

[6] As the covariance between the two time-varying components of the participation and wage equation must be zero for the fixed effects estimator to be consistent, the significance of the second time varying selection term provides some evidence, conditional on the sample selection model's identifying assumptions, to suggest that the fixed effects estimator may not be appropriate in this case.

and rural coefficients ($H_o : \beta^u = \beta^r$) is also clearly rejected in the OLS and Sample Selection models in both Table 4 and 5. However, the results for the Fixed Effects estimator are more ambiguous with little evidence to support the rejection of this hypothesis for men (p-value 0.111), while the test is rejected only at the 10 percent level only for women.

Finally, the implied urban wage premium is given for all specifications. These are calculated by applying the urban and rural coefficient estimates separately to the urban sample to generate two sets of individual predicted offer wages. The urban wage premium reported is the difference between the average of the offer wage predictions obtained using the urban estimates and the average obtained using the rural estimates.

As with previous evidence by Glaeser and Maré [10], a positive urban wage premium remains for both women and men after controlling for both observed and unobserved heterogeneity. Although firm level agglomeration effects imply an urban premium in wage levels, the unobserved heterogeneity, which allows each individual's intercept to differ, should purge the observed premium of such effects. Hence, as an urban premium remains in both the Fixed Effects and Sample Selection models, this suggests that the premium is not simply a firm level effect. Further, in all specifications the female urban wage premium is significantly larger than that for men, consistent with the hypothesis that improved job matching in denser urban areas is important in explaining higher urban wages.

Turning to the coefficient estimates in Table 4 and 5. The OLS and Sample Selection models provide similar results that are consistent with prior expectations. For both men and women, the results for these models indicate that wages increase with experience and education level but decline with time out of the labor force.

Consistent with previous evidence (Light and Ureta [18]), they also suggest that returns to experience and that wage depreciation are lower for women. Returns to education below degree level also appear lower for women. In contrast, the estimated coefficients in for the Fixed Effects model are often poorly determined, particularly for variables such as education level where there is little time variation within individuals. For example, the coefficients on the education variables do not follow the expected pattern for either men or women, while only four of the sixteen education coefficients estimated are significant. In addition, although the estimated coefficients have the expected signs neither time out of the labor force or its square are found to have a significant effect on wages for women in this model.

In the OLS and Sample Selection results, the source of urban-rural differences in wages are well determined and – apart from results for the education variables - consistent with prior hypotheses on potential gender differences. For women, the estimates in Table 4 are similar across OLS and Sample Selection models. They both indicate significant reduced wage depreciation in the urban sample consistent with improved job matching with the urban coefficients on time out of the labor force and time out of the labor force squared around half the rural estimates. Further, the joint test of urban-rural equality of the two out of the labor force coefficients is rejected in both specifications (p-values <0.001). On the other hand there is no evidence of higher urban returns to experience or to education. Indeed, with respect to returns to education, there is some evidence that the returns to lower range qualifications are in fact higher in the rural sample. In contrast, the results for the Fixed Effects model suggests that the lack of migrants in the data mean it is difficult to identify the source of any urban wage premium using this approach. For example, only one of the differences between an individual urban and rural estimate is statistically significant, i.e. the effect of A-Levels, and in this case neither individual estimate is significant.

The OLS and Sample Selection results for men in Table 5 indicate that urban wage depreciation associated with time out of the labor force is lower consistent with improved urban job matching. However, urban returns to experience for men are also significantly higher consistent with the existence of improved informal spillover effects. Joint tests of the experience variables and the out of the labor force variables strongly reject the hypotheses that these coefficients are identical across urban and rural samples in either the OLS or Sample Selection specification (p-values 0.024 and 0.013 respectively). As for women, there is no evidence of higher urban returns to education. The results for the Fixed Effects model also indicate higher urban returns to experience for men (despite the fact that the hypothesis that all urban and rural coefficients are equal cannot be rejected). However in this case, no significant urban-rural differences are found in the effect of time out of the labor force.

Model Evaluation

The ability of the Sample Selection model to identify the urban wage premium and control for unobserved heterogeneity depends on a number of assumptions, e.g. joint normality of errors, independence between the error components and regressors in the wage equation, non-zero correlations between unobserved heterogeneity in participation and wage equations etc. As model misspecification induced by violations of these assumptions is likely to be reflected in the wage residuals, we use these as the basis of an informal test of overall model validity.

The three estimation approaches used, i.e. OLS, Fixed Effects, and Sample Selection, also provide specific predictions about the behavior of the wage residuals. In particular, because OLS does not take account of unobserved heterogeneity, the unadjusted residuals from the OLS wage regressions should be strongly correlated for individuals. In contrast, if the time varying sample selection effects are not important, Fixed Effects should control effectively for unobserved heterogeneity and the residuals in this model should be uncorrelated at the individual level. Similarly, if the Sample Selection model does control effectively for unobserved heterogeneity through the correlation between the unobserved effects in the participation and wage equations, the unadjusted residuals from the Sample Selection wage regressions should also be uncorrelated at the individual level.

The residuals ($\hat{u}_{it}$) from each of the six wage estimations in Tables 4 and 5 are analyzed using techniques applied when examining the covariance of earnings (Dickens [6], Abowd and Card [1]). First, within individual residual covariances are calculated, and an estimate of the residual covariance matrix obtained. Second, from these matrices the average covariances by lag length are calculated. These are reported in Table 6 for each wage regression. Zero correlation in the residuals would mean that each of these covariances should not be statistically significant.

Overall the results from Table 6 do not suggest serious underlying misspecification in the Sample Selection model. Furthermore, they indicate that this model is more effective in eliminating correlation in the residuals in both the female and male wage estimations than either OLS or Fixed Effects approaches. The reported OLS residual covariances are large and all strongly statistically significant for both sexes. Although somewhat smaller than for the OLS estimation, all residual covariances beyond one lag remain statistically significant for both sexes in the Fixed Effects results. In contrast, the covariances from the Sample Selection model are generally smaller, with only one individual covariance remaining statistically significant in each case.

Spatial Mobility and Urban Participation-Wage Premiums

The results reported in Table 3-5 provide evidence that urban participation and wage premiums are larger for women and that the structure of urban-rural differences in returns to experience and time out of the labor force differ by gender. If the source of these differences is improved urban job matching effects, we would expect that the urban premium will be greater for groups where spatial mobility is thought particularly restricted.

To explore this Table 7 reports the urban participation and wage premiums for a sub-sample thought a priori likely to be less spatially mobile, i.e. married and cohabiting individuals, and a sub-sample thought to be more mobile, i.e. unmarried individuals. The results for the urban characteristics (columns 1 and 3) are calculated in the same way to the premiums presented in Tables 3-5, i.e. the urban and rural estimates are applied to the urban sample to provide two sets of predictions, with the urban premium equal to the average predicted value using the urban coefficients minus the average using the rural coefficients. In addition, the urban premiums obtained when the urban and rural coefficients are applied to the rural samples of men and women are also reported. Each set of results is based on the estimation of a separate dynamic Random Effects participation probit and Sample Selection model wage regression for the appropriate sub-sample. To provide a general indication as to the robustness of the urban premiums reported, the result of the joint test that all urban and rural coefficients are identical in the model used to generate the predictions is also reported for each case.

The participation results provide little support for the hypothesis that the observed urban premium arises from differences in spatial mobility. For women, the hypothesis that there are no urban-rural differences in the coefficients used to generate the results in Table 3 cannot be rejected for either sub-sample, while the calculated premiums are larger for those thought to be less spatially constrained, i.e. unmarried women. For men, the urban participation premiums are larger for the samples of married men but remain small.

In contrast, the urban wage premium results do provide further evidence that the urban-rural gender differences observed in Tables 4 and 5 are driven by differences in spatial mobility. First, the urban-rural differences in estimated coefficients underlying the calculated premiums appear robust, with the hypothesis that the urban and rural coefficients are identical rejected (at 5 % significance) in all cases. Second, the urban wage premium varies as predicted. So although the urban wage premium does not disappear for single women, it is substantially lower than for the married/cohabiting sample. For example, when the characteristics of rural single women are used the premium falls to 0.011 but rises to 0.082 for the sample of rural married women. Similarly, the wage premium is larger for married than single men.

VI. SUMMARY AND CONCLUSIONS

This paper has considered the extent of gender differences in both the urban wage and participation premiums using panel data from the United Kingdom. Specifically, participation and wage equations were estimated for urban and rural women and men using a panel sample selection model, which controlled for observed and unobserved heterogeneity. As the identification of the wage component of the model requires a number of relatively strong assumptions, e.g. joint normality of errors, the wage equation results were also compared with both OLS and Fixed Effects estimations. These comparisons showed that the Sample Selection estimator does provide an effective way in which to control for unobserved heterogeneity and sample selection in cases where identification problems reduce the usefulness of Fixed Effects estimator.

From the results, there is evidence of a small but economically significant urban participation premium for women. In contrast, there appears to be no economically significant participation premium for men. However, for both women and men there is evidence that participation structure differs in urban areas, although specific urban density effects are difficult to identify.

In contrast, the wage regressions do suggest that urban density effects induce gender differences in wages. Even after controlling for observed and unobserved heterogeneity, the urban premium is larger for women. Further, consistent with hypothesis that higher urban market density counteracts the effects of lower spatial mobility, the urban wage premium for women was substantially larger for those who were married or cohabiting relative to those who were single.

Finally, while there is no evidence of higher urban returns to experience, wage depreciation for women is appreciably lower in the urban sample. In contrast, both higher returns to experience and lower wage depreciation help explain the male urban wage premium. While not conclusive, these results do suggest that improved urban job matching effects are relatively more important in the female urban wage premium than learning spillover effects.

The results indicate a number of possible questions for further research. The urban-rural categorization used here is necessarily rather broad. Are there different effects if a finer scale is available, e.g. are there city size effects? Also, the tests of the impact of higher service provision on participation applied are rather indirect. Can measures be found which would allow such effects to be tested more fully, e.g. effects of differences in public transport provision?

REFERENCES

[1] J. Abowd, D. Card, The Covariance Structure of Earnings and Hours Changes, Econometrica 57 (1989) 411-445.

[2] J.G. Altonji, R. Blank, Race and Gender in the Labor Market, in: O.C. Ashenfelter and D. Card (Eds.) Handbook of Labor Economics, Vol. 3, North Holland, Amsterdam, 1999, pp. 3144-3260.

[3] W. Arulampalam, A. Booth, M. Taylor, Unemployment persistence, Oxford Economic Papers 52 (2000) 24-54.

[4] Cabinet Office, Rural Economies. Report of the Performance and Innovation Unit. HMSO, London 1999.

[5] G. Chamberlain, Panel Data, in: Z.Griliches and Intriligator (Eds.), M. Handbook of Econometrics Vol. 2, North Holland, Amsterdam, 1984, pp.1247-1318.

[6] R. Dickens, The Evolution of Individual Male Earnings in Great Britain: 1975-95, Economic Journal 110 (2000) 27-49.

[7] R. Frank, Why Women earn less: The Theory and Estimation of Differential Overqualification, American Economic Review 68 (1978) 360-373.

[8] E. Glaeser, Are Cities Dying, Journal of Economic Perspectives 12 (1998) 139-160.

[9] E. Glaeser, Learning in Cities, Journal of Urban Economics 46 (1999) 254-277.

[10] E. Glaeser, D Maré, D Cities and Skills, Journal of Labor Economics 19 (2001) 316-342.

[11] J. Heckman, Heterogeneity and state dependence, in: S.Rosen (Ed.) Studies in Labor Markets, Chicago Press Chicago, 1981, pp.91-139.

[12] J. Heckman, The incidental parameters problem and the problem of initial conditions in estimating a discrete-time-discrete-data stochastic process, in: C.F.

Manski, D., McFadden (Eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge MA, 1981, pp.114-78.

[13] R. Hesley, W. Strange, Matching and agglomeration economies in a system of cities, Regional Science and Urban Economics 20 (1990) 189-212.

[14] H. Hwang, W. Reed, C. Hubbard, Compensating Wage Differentials and Unobserved Productivity, Journal of Political Economy 100 (1992) 835-858.

[15] D. Hyslop, State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women, Econometrica 67 (1999) 1255-1294.

[16] P. Krugman, Geography and Trade, MIT Press, Cambridge MA, 1991.

[17] D. Lichter, G. Johnson, D. McLaughlin, Changing linkages between work and poverty in Rural America, Rural Sociology 59 (1994) 395-415.

[18] A. Light, M. Ureta, Early-Career Work Experience and Gender Wage Differentials, Journal of Labor Economics 13 (1995) 121-154.

[19] J. Madden, L. Chiu, The Wage Effects of Residential Location and Commuting Constraints on Employed Married-women, Urban Studies 27 (1990) 353-369.

[20] J. Mincer, H. Ofek, Interrupted work careers: Depreciation and Restoration of Human Capital, Journal of Human Resources 17 (1982) 1-24.

[21] T. Nijman, M. Verbeek, Nonresponse in panel data: The impact on estimates of a life cycle consumption function, Journal of Applied Econometrics 7 (1992) 243-157.

[22] S. Porterfield, On the precipice of reform: welfare spell durations for rural female headed families, American Journal of Agricultural Economics 80 (1998) 994-999.

[23] J. Rauch, Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities, Journal of Urban Economics 34 (1993) 380-400.

[24] J. Roback, Wages, Rents, and the Quality of Life, Journal of Political Economy 90 (1982) 1257-78.

[25] Scottish Office, Towards a development strategy for Rural Scotland. A discussion paper. Scottish Office, Edinburgh 1997.

[26] J. Stabler, Rural America: a challenge to regional scientists, Annals of Regional Science 33 (1999) 1-14.

[27] R. Tarling, J. Rhodes, J. North, G. Broom, The Economy and Rural England, Rural Development Commission, London 1993.

[28] F. Vella, Estimating Models with Sample Selection Bias: A Survey, Journal of Human Resources 33 (1998) 127-169.

[29] F. Vella, M Verbeek, Two-step estimation of panel data models with censored endogenous variables and selection bias, Journal of Econometrics, 90 (1999) 239-263.

[30] M. Verbeek, T. Nijman, Testing for selectivity bias in panel data models, International Economic Review 33 (1992) 681-703.

[31] W.C. Wheaton, M.J. Lewis, Urban wages and labor market agglomeration, Journal of Urban Economics 51 (2002) 542-562.

[32] C.H. Wheeler, Search, Sorting and Urban Agglomeration, Journal of Labor Economics 19 (2001) 879-899.

Table 1: Summary Statistics: Means

	Women Urban	Women Rural	Men Urban	Men Rural
Working	0.71	0.66	0.86	0.83
$Log_e(wage)$*	1.58	1.43	1.83	1.75
Work Experience				
Total actual experience (years)	14.40	14.43	18.18	19.55
Total time out of labor force (years)	7.03	7.93	2.15	2.17
Highest Education Level Attained				
O-levels or equivalent	0.37	0.38	0.28	0.28
A-levels or equivalents	0.10	0.12	0.14	0.11
Nursing or other higher qualifications	0.17	0.17	0.23	0.30
Degree plus	0.13	0.11	0.17	0.11
Children				
Number Children < 5 years of age	0.23	0.24	0.19	0.21
Number Children 5 –11 years of age	0.38	0.40	0.31	0.34
Number Children 12–16 years of age	0.21	0.19	0.17	0.21
Other				
Married or cohabiting	0.73	0.81	0.71	0.74
Non-labor income £000	3.14	3.35	2.73	2.95
Total Individuals	1918	259	1430	196
Total Observations	11968	1717	8533	1213

* *Wages* are usual monthly labor earnings divided by usual hours worked with an adjustment factor of 1.5 for overtime hours. *Total Experience* is years in full or part-time employment constructed from the work history files. *Out of the labor force* is years unemployed or out of the labor force constructed from the work history files. *O-levels* (and equivalent) are nationally examinations taken in up to 10 subjects by students normally at the end of compulsory schooling at 16. *A-levels* are national examinations typically taken in up to 3 subjects by students aged 18. Results from both types of examinations are graded and are used as the basis for acceptance at College or University.

Table 2: Female Participation Estimates

	Static		Dynamic	
	Urban	**Rural**	**Urban**	**Rural**
Constant	0.608	0.941	0.088	0.418
	(0.172)	(0.396)	(0.224)	(0.437)
Lagged working			1.150	1.139
			(0.066)	(0.146)
Experience	0.236	0.229	-0.004	-0.056
	(0.014)	(0.029)	(0.044)	(0.071)
Experience squared	-0.0052	-0.0060	-0.002	-0.003
	(0.0004)	(0.0008)	(0.0003)	(0.001)
Out of labor force	-0.207	-0.155	-0.200	-0.263
	(0.015)	(0.042)	(0.049)	(0.085)
Out of labor force squared	0.0024	-0.0003	0.001	-0.001
	(0.0006)	(0.0019)	(0.0005)	(0.002)
O-level	** 0.270	-0.618	** 0.394	-0.322
	(0.096)	(0.267)	(0.092)	(0.227)
A-level	0.773	0.213	0.531	0.187
	(0.134)	(0.357)	(0.132)	(0.324)
Nursing etc	** 0.776	-0.155	** 0.673	-0.006
	(0.115)	(0.294)	(0.115)	(0.283)
Degree plus	0.700	0.483	0.712	0.672
	(0.131)	(0.36)	(0.132)	(0.391)
No. Children < 5 years old	-1.293	-1.352	-0.555	-0.609
	(0.053)	(0.131)	(0.064)	(0.17)
No. Children 5-<12 years old	-0.174	-0.172	0.115	-0.008
	(0.041)	(0.093)	(0.064)	(0.16)
No. Children 12-16 years old	0.117	0.299	0.105	0.303
	(0.063)	(0.141)	(0.089)	(0.205)
Married	0.024	0.216	0.060	0.290
	(0.067)	(0.18)	(0.069)	(0.184)
Non labor income	-0.029	-0.034	-0.009	-0.014
	(0.001)	(0.008)	(0.003)	(0.01)
$\hat{\rho}$	0.729		0.486	
	(0.012)		(0.037)	
Ho: All Coefficients zero χ^2 (p-value)	5414.3	(< 0.001)	3199.6	(< 0.001)
Log Likelihood	-3466.8		-4054.7	
Ho: Urban=rural coefficients χ^2 (p-value)	36.9	(< 0.001)	22.0	(0.056)
Urban Participation Premium	0.031		0.028	

Equations estimated by Random Effects. Standard Errors in brackets. Starred coefficients represent significant urban-rural differences in individual coefficients (** 5%,*10%) . All regressions included regional and wave dummies. For the Dynamic equation a separate equation adjusts for initial conditions while average experience, time out of the labor force, number of children by age and non-labor income are included to control for potential correlations between random effects and regressors. The initial conditions equation included, own and spouse education level, own and spouse age, number and age of children, plus social class of mother and father as regressors.

Table 3: Male Participation Estimates

	Static		Dynamic	
	Urban	**Rural**	**Urban**	**Rural**
Constant	2.570	1.915	0.070	0.403
	(0.251)	(0.631)	(0.262)	(0.613)
Lagged working			1.194	1.521
			(0.107)	(0.266)
Experience	0.094	0.192	-0.155	-0.194
	(0.017)	(0.047)	(0.056)	(0.173)
Experience squared	-0.0026	-0.0044	-0.0012	-0.0020
	(0.0004)	(0.0011)	(0.0003)	(0.0011)
Out of labor force	-0.617	-0.690	-0.485	-0.263
	(0.035)	(0.101)	(0.072)	(0.177)
Out of labor force squared	0.0226	0.0329	0.0082	0.0135
	(0.0016)	(0.0054)	(0.0013)	(0.0052)
O-level	** 0.676	0.495	0.526	0.078
	(0.146)	(0.494)	(0.135)	(0.426)
A-level	1.128	1.089	0.731	0.144
	(0.185)	(0.59)	(0.166)	(0.47)
Nursing etc	** 1.157	1.409	0.734	0.257
	(0.172)	(0.543)	(0.148)	(0.462)
Degree plus	1.522	0.851	* 1.150	0.045
	(0.198)	(0.687)	(0.172)	(0.556)
No. Children < 5 years old	-0.198	-0.439	-0.113	-0.144
	(0.106)	(0.305)	(0.147)	(0.528)
No. Children 5-<12 years old	-0.069	-0.200	0.005	-0.083
	(0.076)	(0.2)	(0.107)	(0.24)
No. Children 12-16 years old	-0.184	-0.445	-0.021	-0.302
	(0.099)	(0.345)	(0.121)	(0.322)
Married	-0.078	0.134	0.144	0.262
	(0.113)	(0.423)	(0.104)	(0.354)
Non labor income	* -0.061	-0.134	-0.010	-0.047
	(0.003)	(0.018)	(0.006)	(0.024)
$\hat{\rho}$	0.787		0.433	
	(0.016)		(0.057)	
Ho: All Coefficients zero χ^2 (p-value)	4277.7	(< 0.001)	2576.1	(< 0.001)
Log Likelihood	-1650.1		-2095.9	
Ho: Urban=rural coefficients χ^2 (p-value)	51.5	(< 0.001)	23.0	(0.042)
Urban Participation Premium	-0.0007		0.008	

Equations estimated by Random Effects. Standard Errors in brackets. Starred coefficients represent significant urban-rural differences in individual coefficients (** 5%,*10%) . All regressions included regional and wave dummies. For the Dynamic equation a separate equation adjusts for initial conditions while average experience, time out of the labor force, number of children by age and non-labor income are included to control for potential correlations between random effects and regressors. The initial conditions equation included, own and spouse education level, own and spouse age, number and age of children, plus social class of mother and father as regressors.

Table 4: Female Wage Equation Estimates

	Women					
	OLS		Fixed Effects		Sample Selection	
	Urban	Rural	Urban	Rural	Urban	Rural
Constant	1.228	1.141			1.240	1.168
	(0.031)	(0.077)			(0.034)	(0.078)
Experience	0.035	0.034	0.063	0.072	0.034	0.032
	(0.002)	(0.006)	(0.029)	(0.03)	(0.002)	(0.006)
Experience squared	-0.0007	-0.0007	-0.001	-0.001	-0.0007	-0.0007
	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Out of labor force	** -0.027	-0.054	-0.012	-0.029	** -0.027	-0.055
	(0.003)	(0.008)	(0.035)	(0.041)	(0.003)	(0.008)
Out of labor force squared	** 0.0004	0.0024	0.0002	0.001	** 0.0004	0.0024
	(0.0001)	(0.0004)	(0.001)	(0.002)	(0.0001)	(0.0004)
O-level	** 0.137	0.247	-0.008	-0.066	* 0.139	0.243
	(0.017)	(0.05)	(0.044)	(0.098)	(0.017)	(0.051)
A-level	* 0.250	0.387	* -0.011	-0.181	0.256	0.377
	(0.022)	(0.072)	(0.049)	(0.096)	(0.023)	(0.073)
Nursing etc	0.347	0.346	0.027	-0.090	0.352	0.344
	(0.021)	(0.066)	(0.043)	(0.085)	(0.021)	(0.066)
Degree plus	0.817	0.866	0.062	0.135	0.821	0.868
	(0.022)	(0.073)	(0.069)	(0.123)	(0.023)	(0.074)
$\hat{\sigma}_{\alpha\theta}$					0.038	
					(0.009)	
$\hat{\sigma}_{e\eta}$					-0.054	
					(0.016)	
Ho: All coefficients zero χ^2 (p-value)	90294.4	(< 0.001)	296.67	(< 0.001)	89095.79	(< 0.001)
Ho: Zero autocorrelation χ^2 (p-value)	501.5	(< 0.001)	338.8617	(< 0.001)	87.03	(< 0.001)
Ho: Urban= Rural coefficients χ^2 (p-value)	41.5	(< 0.001)	15.44	(0.051)	42.4	(< 0.001)
Urban Wage Premium	0.063		0.071		0.064	

Standard Errors (in brackets) are robust to autocorrelation and heteroskedasticity. In the sample selection model they are adjusted for the two-step estimation process. Starred coefficients represent significant urban-rural differences in individual coefficients (** 5%,*10%). All regressions include a common set of regional and wave dummies.

Table 5: Male Wage Equation Estimates

		Men							
		OLS			**Fixed Effects**			**Sample Selection**	
		Urban	**Rural**		**Urban**	**Rural**		**Urban**	**Rural**
Constant		1.251	1.360					1.273	1.373
		(0.036)	(0.08)					(0.039)	(0.08)
Experience	**	0.040	0.023	**	0.076	0.055	**	0.039	0.022
		(0.003)	(0.006)		(0.023)	(0.023)		(0.003)	(0.006)
Experience squared	**	-0.0007	-0.0003	**	-0.001	-0.0003	**	-0.0007	-0.0003
		(0.0001)	(0.0001)		(0.0001)	(0.0002)		(0.0001)	(0.0001)
Out of labor force	**	-0.056	-0.100		-0.107	-0.150	**	-0.050	-0.097
		(0.006)	(0.015)		(0.034)	(0.051)		(0.007)	(0.015)
Out of labor force squared	**	0.0026	0.0049		0.009	0.012	**	0.0023	0.0050
		(0.0005)	(0.0009)		(0.003)	(0.005)		(0.0005)	(0.0009)
O-level	**	0.224	0.198		-0.041	0.042		0.219	0.201
		(0.021)	(0.054)		(0.047)	(0.090)		(0.022)	(0.054)
A-level	*	0.321	0.438	*	-0.046	0.167		0.314	0.432
		(0.025)	(0.073)		(0.051)	(0.086)		(0.026)	(0.073)
Nursing etc		0.449	0.477		0.013	0.138		0.443	0.475
		(0.023)	(0.056)		(0.044)	(0.075)		(0.024)	(0.056)
Degree plus		0.828	0.822		0.167	0.284		0.817	0.818
		(0.026)	(0.077)		(0.074)	(0.110)		(0.027)	(0.077)
$\hat{\sigma}_{\alpha\theta}$								0.015	
								(0.013)	
$\hat{\sigma}_{e\eta}$								-0.104	
								(0.033)	
Ho: All coefficients zero χ^2 (p-value)		91063.0	(< 0.001)		412.96	(< 0.001)		91730.3	(< 0.001)
Ho: Zero autocorrelation χ^2 (p-value)		533.4	(< 0.001)		362.8	(< 0.001)		30.1	(0.090)
Ho: Urban= Rural coefficients χ^2 (p-value)		24.2	(0.004)		13.0	(0.111)		24.7	(0.003)
Urban Wage Premium		0.037			0.043			0.038	

Standard Errors (in brackets) are robust to autocorrelation and heteroskedasticity. In the sample selection model they are adjusted for the two-step estimation process. Starred coefficients represent significant urban-rural differences in individual coefficients (** 5%,*10%). All regressions include a common set of regional and wave dummies.

Table 6: Wage Equation Residual Covariances

Covariance	Women			Men		
	OLS	Fixed Effects	Sample Selection	OLS	Fixed Effects	Sample Selection
$\text{cov}(\hat{u}_{it}, \hat{u}_{it-1})$	0.124	-0.001	0.005	0.136	-0.003	0.006
	(0.009)	(0.003)	(0.004)	(0.009)	(0.003)	(0.005)
$\text{cov}(\hat{u}_{it}, \hat{u}_{it-2})$	0.112	-0.008	0.002	0.128	-0.006	-0.001
	(0.008)	(0.003)	(0.004)	(0.009)	(0.002)	(0.005)
$\text{cov}(\hat{u}_{it}, \hat{u}_{it-3})$	0.104	-0.014	-0.002	0.121	-0.011	0.002
	(0.008)	(0.003)	(0.005)	(0.009)	(0.003)	(0.006)
$\text{cov}(\hat{u}_{it}, \hat{u}_{it-4})$	0.098	-0.017	0.008	0.114	-0.014	0.007
	(0.008)	(0.003)	(0.006)	(0.009)	(0.002)	(0.007)
$\text{cov}(\hat{u}_{it}, \hat{u}_{it-5})$	0.089	-0.023	0.005	0.113	-0.014	0.018
	(0.008)	(0.004)	(0.006)	(0.009)	(0.002)	(0.007)
$\text{cov}(\hat{u}_{it}, \hat{u}_{it-6})$	0.094	-0.014	0.025	0.112	-0.012	0.004
	(0.008)	(0.003)	(0.008)	(0.01)	(0.002)	(0.009)

Standard Errors in brackets

Table 7: Predicted Urban Participation and Wage premiums

Sample		Participation			Wages	
		Characteristics				
		Urban	***Rural***		***Urban***	***Rural***
				Women		
All	**	0.028	0.028	**	0.064	0.067
Married/Cohabiting		0.021	0.019	**	0.088	0.082
Single		0.025	0.056	**	0.036	0.011
				Men		
All	*	0.008	0.010	**	0.038	0.024
Married/Cohabiting	**	0.010	0.010	**	0.043	0.046
Single		-0.058	-0.038	**	0.013	-0.028

Participation results for All women and Men based on dynamic RE Probit estimates presented in Tables 2 and 3. Other participation results based on identically structured RE Probit estimations for these sub-samples. The participation premium is calculated as the difference in the average participation prediction using the urban and rural coefficients for urban or rural characteristics. The wage premium estimates for women and men based on Sample Selection model results from Tables 4 and 5, with other estimates obtained from identically structured wage estimations. The wage premium is calculated as the difference in the average offer wage prediction using the urban and rural coefficients for urban or rural characteristics. Starred values indicate that the joint hypothesis that all urban and rural coefficients are identical is rejected in the model used to generate the predictions (** 5%,*10%).