CENTER FOR LABOR ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY
WORKING PAPER NO. 11

Interpreting Movements in Aggregate Wages: The Role of Labor Market Participation

Richard Blundell,* Howard Reed† and Thomas Stoker‡

March 1999

Abstract

This paper analyzes the role of labor market participation in the relationship between aggregate wages and individual wages. A new and easily implementable framework for the empirical analysis of aggregation biases in this context is developed. Aggregate real wages are shown to contain three important bias terms: one associated with the dispersion of individual wages, a second reflecting the distribution of working hours, and a third deriving from compositional changes in the (selected) sample of workers. Noting the importance of these issues for recent experience in Britain, data on real wages and participation for British male workers over the period 1978-1995 are studied. A close correspondence between the estimated biases and the patterns of differences shown by aggregate wages is established. This is shown to have important implications for the interpretation of real wage growth over this period.

JEL: J21, J31.

Keywords: Aggregate Real Earnings, Participation, Wage Distribution.

*University of California at Berkeley, Institute for Fiscal Studies and Department of Economics, University College London, Gower Street, London, WC1E 6BT.

† Institute for Fiscal Studies and Department of Economics, University College London, Gower Street, London, WC1E 6BT.

‡ Sloan School of Management, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142 USA.

1. Real Wages and Participation[1]

Aggregate figures for real wage growth are used extensively in policy debate. They are deployed to reflect changes in the well being of workers over time and are also used for comparisons across education or cohort groups and for cross country or cross region comparisons. However, as pointed out in the original study by Bils (1985), if participation rates change differentially across the time periods or across the groups used in these comparisons, then aggregate real wages are likely to provide a misleading picture of changes in the structure of real wages facing individual workers. For example, if the overall distribution of skills in the workforce remains unchanged, aggregate wages will increase when relatively low wage individuals leave employment, but it is hard to argue that 'well being' has been improved in any meaningful way.

This paper develops a simple characterization of the relationship between participation and aggregate wages and derives the precise form of the bias in inferring the behavior of individual wages from the analysis of aggregate (average) hourly earnings, or aggregate wages. The bias is decomposed into three interpretable terms reflecting changes in the distribution of individual wages, changes in participation and changes in hours worked. This bias is then investigated using data for male wages from the British economy in the 1980s and 1990s. These data analyses point to significant deviations between aggregate and individual measures that imply important revisions in the interpretation of real wage growth

[1]Acknowledgements: We would like to thank Orazio Attanasio, Amanda Gosling, James Heckman, Costas Meghir, Julian McCrae and seminar participants at Chicago, MIT and LSE for comments on an earlier draft. The financial support of the ESRC Centre for the Micro-Economic Analysis of Fiscal Policy at IFS is gratefully acknowledged.

over this period.

We identify three reasons why the British labor market experience during this period is particularly attractive for this analysis. First, there have been strong secular and cyclical movements in male participation over this period. Second, there exists a long and representative time series of individual survey data, collected at the household level, that records detailed information on individual hourly wages as well as many other individual characteristics and income sources. Finally, over this period, there has been a systematic change in the level of real out-of-work income. The household survey data utilized in this study allows an accurate measure of this income variable which, in turn, acts as an informative instrument in controlling for participation in our analysis of wages.

Labor market behavior in Britain over the last twenty years serves to reinforce the importance of these issues. Indeed the relationship between wage growth and participation in Britain has often been the focus of headline news.[2] Figure 1.1 shows aggregate participation,[3] aggregate wages and average hours worked from 1978 to 1995. In 1978-9, over 90% of men aged between 18 and 59 were employed. The participation rate fell dramatically in the recession of the early 1980s and then recovered somewhat in the late 1980s (although not to its initial level). In the early 1990s there was another recession and another sharp decline. In contrast, log average wages show reasonably steady increase from 1978 through the 1990s, with slight decreases in 1982 and 1989 (and a slower rate of growth in the 1990s). Finally, over the same period, average weekly hours show very limited variation.

[2]For example, "Rise in Earnings and Jobless Sparks Concern", *Financial Times*, front page, June 18th, 1998.

[3]We define the participation rate as the ratio of the number of employees over the number of [employees plus the unemployed plus men who are not active in the labour market] as measured in the UK Family Expenditure Survey each year. The self-employed are excluded from both the numerator and the denominator of this expression. So for us, 'participation rate' and 'employment rate' are the same thing.

Figure 1.2 gives the same statistics broken down by three education groups: men who left full-time education aged 16 or less, those who left full-time education aged 17 or 18, and those who left aged 19 or older. While labor market participation rates for all three education groups declined over this period, the fall was much larger for the lowest-educated group than it was for the other two groups. This gap widened in both recessions which took place during the period. In contrast, the aggregate wages of all three education groups have grown by a roughly similar amount between 1978 and 1995. Average hours do show some variation by group; in the late 1970s the least well-educated group worked about 4 hours more on average than the most educated, whereas by 1995, hours of work for the different education groups had more or less converged.

Another very relevant feature of recent British experience are the well documented changes in the real income which different groups of individuals receive (or would receive) whilst out of work. Figure 1.3 shows the variation of two measures of out-of-work income (which is mostly made up of income from the benefit system) for married men with non-working spouses.[4] Total out-of-work benefits show a broad increase in the early 1980s for this group, followed by a shallower increase in the late 1980s and early 1990s.[5] The housing benefit component of out-of-work income shows a smoother uniform increase over the same period. Although it is unlikely that variation in real value of benefit income can explain all of the variation in participation rates, we argue that changes in real benefits serve as an important "instrumental variable" for controlling for endogenous selection

[4]For the discussion in this introduction we present figures for the period 1984-1995. Although we have a measure of this variable for 1978/79 which we also use in the statistical model reported below we do not, at present, have an accurate measure for all individuals in our sample for the period 1980-1983.

[5]The simulated out-of-work income measure is computed by the Institute for Fiscal Studies' tax and benefit microsimulation model TAXBEN. See Section 3.1 for more information on this model and on the structure of the British tax and benefit system throughout the period.

in real wages.

Our framework begins with a basic model of human capital and skill price as developed in Heckman and Sedlacek (1985). The returns to human capital are allowed to be time varying in response to demand and supply shocks over time, or in particular, over the business cycle. Biases occur when trying to assess the cyclicality or trend behavior of wages or returns using aggregate wage measures. The biases result from the varying composition of those individuals in work at any point in time and from variation in hours that is correlated with hourly wages. In accordance with this, we show that there are three aggregation factors that need to be accounted for in examining the evolution of aggregate wages. The first factor describes the dispersion of wages and arises from aggregation over the standard log-linear model of individual wages. This term explicitly reflects the effect of increasing wage dispersion separately from the impact of participation. The second factor measures the adjustment for composition changes in hours and depends on the size of the covariance between wages and hours. The final factor highlights the importance of the participation decision, capturing the effects of composition changes within the selected sample of workers from which measured wages are recorded. As in the standard selection bias literature, this third factor depends on the covariance between participation and wages.

These aggregation biases are likely to be particularly important for the study of wages and returns in Europe where there have been dramatic and systematic changes in the variance of hourly wages, the distribution of hours of work and in participation rates. These have occurred both secularly and cyclically. Our application to real wages for men in Britain shows important impacts of heterogeneity and labor participation. To anticipate, we find that changes in dispersion of individual wages, attributable to both observable and unobservable factors, lead to a secular increase in the bias from using aggregate wage measures. In contrast we find that the changes in composition, induced by the pattern of labor market participation, induce a counter cyclical bias in the aggregate measure.

More specifically, following earlier specifications for British data (see Gosling et. al (1996) and Schmitt (1995)), our individual (log) wage equations allow for date-of-birth cohort, trend and cycle variation. The participation equation is identified by the systematic variation in the welfare benefit levels eligible to these men over time across individual types. This is especially the case for housing benefit variation which varied strongly across time, location and cohort group. The cohort variation occurs because individuals in lower educated older cohorts had a much higher chance of spending their lives in public housing. We take this variation to be exogenous to the individual participation decision conditional on the cohort, education, region, trend and cycle effects. The individual level wage equation results show a significant selection effect that varies systematically over the trend and cycle and differs across education groups.

We consider several different ways to compare individual and aggregate wages. In this paper our focus is on comparison between the evolution of the log aggregate hourly wage and the mean from the individual level model. We show a large discrepancy in the level and growth between the aggregate and individual wage paths. This discrepancy is shown to be almost completely captured by the aggregation factors described above, validating our model specification and providing a detailed interpretation of the aggregation biases involved. This is associated with an important upward bias in the aggregate trend of real wages and a reduction in the degree of procyclicality.

The layout of the paper is as follows. Section 2 lays out the modeling framework that will underlie the empirical work. We derive some new results on aggregating over lognormal distributions, and then we apply the results to spell out the empirical implications of our model to individual and aggregate level wage data. Section 3 covers the data, our main results, and results relevant for the validation of our modeling framework. Section 4 draws some conclusions.

2. Aggregation and Selection

2.1. A Model for Real Wages

We begin by introducing our approach for modeling individual wages. We follow Roy (1951) in basing wages on human capital or skill levels, assuming that any two workers with the same human capital level are paid the same wage. Thus we assume that there is no comparative advantage, and no sectoral differences in wages for workers with the same human capital level.[6] We assume that the mapping of skills to human capital is time invariant, and that the price or return to human capital is not a function of human capital endowments. In particular, we begin with a framework consistent with the proportionality hypothesis of Heckman and Sedlacek (1990).

2.1.1. One Homogeneous Sector and Nonparticipation

The simplest version of the framework assumes that each worker i possesses a human capital (skill) level of H_i. Human capital is nondifferentiated, in that it commands a single price r_t in each time period t. In this case the wage paid to worker i at time t is

$$w_{ti} = r_t H_i \tag{2.1}$$

[6]Heckman and Sedlacek (1985) provide an important generalization of this framework to multiple sectors. We plan on examining a multisectoral model as part of future research. In addition, the importance of normality assumptions in such a generalization is explored further in Heckman and Honore (1993).

Human capital H_i is assumed log-normally distributed[7], with mean

$$E(\ln H_i) = \delta_{js}$$

and variance σ^2, where δ_{js} is a level that varies with the cohort j to which i belongs and the education level s of worker i. In other words, the log wage equation has the additive form

$$\ln w_{it} = \ln r_t + \delta_{js} + \epsilon_{it} \tag{2.2}$$

where ϵ_{it} is $\mathcal{N}(0, \sigma^2)$.[8] In this model growth in returns is constant across all individuals. Below we extend this model to allow education returns to differ over time.

Reservation wages w_{it}^* are also assumed to be lognormal, with

$$\ln w_{it}^* = \alpha \ln b_{it} + \eta_{js} + \zeta_{it} \tag{2.3}$$

where ζ_{it} is $\mathcal{N}\left(0, \sigma_\zeta^2\right)$ and where b_{it} can be interpreted as an exogenous benefit level that varies with individual characteristics and time. Participation occurs if $w_{it} \geq w_{it}^*$, or with

$$\ln r_t - \alpha \ln b_{it} + \delta_{js} - \eta_{js} + \epsilon_{it} - \zeta_{it} > 0 \tag{2.4}$$

We represent the participation decision by the indicator $I_i = 1\left[w_{it} \geq w_{it}^*\right]$.

For examining hours, we will make one of two assumptions in our empirical work. One is to assume that the distribution of hours is fixed. The other is to assume that *desired* hours h_{it} are chosen by utility maximization, where reservation wages are defined as $h_{it}(w^*) = h_0$ and h_0 is the minimum number of hours available for full-time work.[9] We assume $h_{it}(w)$ is normal for each w, and approximate

[7]Although we utilize lognormality assumptions extensively in this section, their reliability is assessed in the empirical analysis that follows.

[8]Clearly, there is an indeterminacy in the scaling of r_t and H_t. Therefore, to study r_t, we will normalize r_t for some year $t = 0$ (say to $r_0 = 1$). We could equivalently set one of the δ's to zero.

[9]This allows for a simple characterisation of fixed costs, see Cogan (1981).

desired hours by

$$h_{it} = h_0 + \gamma\left(\ln w_{it} - \ln w_{it}^*\right) \tag{2.5}$$

$$= h_0 + \gamma\left(\ln r_t - \alpha \ln b_{it} + \delta_{js} - \eta_{js} + \epsilon_{it} - \zeta_{it}\right)$$

For our derivations of aggregation formulae below, we retain the second assumption (since we can easily specialize to the first assumption).

This is our base level specification that maintains the proportionality hypothesis. There are no trend or cycle interactions with cohort or education level in either equation. Such interactions arise with simple extensions of the framework, as follows.

2.1.2. Two Useful Extensions

We consider two extensions of this basic framework, made necessary by our empirical findings. First, suppose that education produces a differentiated type of human capital. That is, a high education worker i has human capital (skill) level of H_i^H and is paid the wage $r_t^H H_i^H$. A low education worker i has human capital (skill) level of $r_t^L H_i^L$ and is paid the wage $r_t^L H_i^L$. Again similar workers with a particular skill level are paid the same in all sectors. If D_i is the high education dummy, the log wage equation has the form

$$\ln w_{it} = D_i \ln r_t^H + D_i \delta_{js}^H + (1 - D_i) \ln r_t^L + (1 - D_i)\, \delta_{js}^L + \epsilon_{it}. \tag{2.6}$$

Here, education can have a time varying impact on wages.

The second extension is to allow the different stock of labor market experience that is associated with each cohort at any specific calender time to have an impact on returns. This generalizes the basic model to allow log wages to display different trend behavior for each date-of-birth cohort group.

2.2. Aggregate Wages

The aggregate wage is measured by

$$\overline{w}_t = \frac{\sum_{i\in(I=1)} e_{it}}{\sum_{i\in(I=1)} h_{it}} = \sum_{i\in(I=1)} \mu_{it} w_{it} \tag{2.7}$$

where $i \in (I = 1)$ denotes a labor market participant and where $e_{it} = h_{it}w_{it}$ is the earnings of individual i in period t, and where μ_{it} are the hours weights

$$\mu_{it} = \frac{h_{it}}{\sum_{i\in(I=1)} h_{it}}.$$

We take the population of participating workers as sufficiently large so that we can ignore sampling variation in average earnings and average hours; modeling the aggregate wage as

$$\overline{w}_t \cong \frac{E\left[h_{it}w_{it}|I_{it} = 1\right]}{E\left[h_{it}|I_{it} = 1\right]}$$

where $E\left[\cdot\right]$ refers to the mean across the population.

2.2.1. Various Micro-Macro Comparisons

Measured wages at the individual level are represented by an entire distribution. Therefore, there are many ways to pose the question of whether aggregate wage movements adequately reflect movements in individual wages. We consider various alternatives here, each of which could be adopted.

The basic framework suggests an economically sensible answer to how to compare individual and aggregate wages. From (2.1), the natural question is whether aggregate wage movements accurately reflect movements in the skill price r_t, or the price of human capital. For example, if aggregate production in the economy has total human capital $(\sum_i H_i)$ as an input, then the appropriate price for that input is r_t. Therefore, the economic comparison to the relevant (quality adjusted) price of labor is

$$r_t \text{ versus } \overline{w}_t.$$

Other interpretable comparisons arise on statistical grounds. Following the tradition of measuring "returns" from coefficients in log wage equations, one could focus on the behavior of the mean log wage. This refers to the comparison

$$E(\ln w_{it}) \text{ versus } \ln \overline{w}_t.$$

This approach is adopted in the work of Solon, Barsky and Parker (1994), as well as in our empirical work. Note that if the log mean of H_i is constant over time in our basic framework, then the mean log wage comparison matches the original "r_t versus $\bar{w}_t$" comparison (in log form). We have listed these comparisons separately because one might be interested in the log wage comparison even without a framework tracing wages to human capital. For completeness, note that one could compare aggregate wages with many other individual concepts, such as the mean log wage for participating workers, as in

$$E(\ln w_{it} | I = 1) \text{ versus } \ln \overline{w}_t.$$

2.3. Some Aggregation Results

Our formulations of aggregate wages are based on results on aggregation of nonlinear relationships over normal and lognormal distributions. We make use of several standard formulae familiar from the analysis of selection bias collected in Appendix A, as well as some further results presented in the following Lemma 2.1. While these further results are rather basic, we could not find specific references to them in the literature, and so we have included a proof as part of Appendix A.

Lemma 2.1. *Suppose that* $(\mathcal{U},\mathcal{V})$ *are jointly normal random variables: namely*

$$\begin{pmatrix} \mathcal{U} \\ \mathcal{V} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_{\mathcal{U}} \\ \mu_{\mathcal{V}} \end{pmatrix}, \begin{pmatrix} \sigma^2_{\mathcal{U}} & \sigma_{\mathcal{U}\mathcal{V}} \\ \sigma_{\mathcal{U}\mathcal{V}} & \sigma^2_{\mathcal{V}} \end{pmatrix}\right).$$

Suppose also that

$$\ln \mathcal{W} = \mathcal{U} \text{ and } I = 1\left[\mathcal{V} < 0\right].$$

Then:

A.

$$E\left[\mathcal{W}|I=1\right] = e^{\mu_{\mathcal{U}}+\frac{1}{2}\sigma^2_{\mathcal{U}}} \cdot \frac{\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]}{\Phi\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]}. \tag{2.8}$$

B.

$$E\left[\mathcal{V}\mathcal{W}|I=1\right] = e^{\mu_{\mathcal{U}}+\frac{1}{2}\sigma^2_{\mathcal{U}}} \cdot \left\{\mu_{\mathcal{V}} + \sigma_{\mathcal{U}\mathcal{V}} - \sigma_{\mathcal{V}}\lambda\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]\right\} \cdot \frac{\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]}{\Phi\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]}. \tag{2.9}$$

Proof. *See Appendix A.*

The formulations (2.8)-(2.9) can be rewritten in terms of the unconditional mean of $\mathcal{W}$, since

$$E\left(\mathcal{W}\right) = e^{\mu_{\mathcal{U}}+\frac{1}{2}\sigma^2_{\mathcal{U}}}.$$

For instance, (2.8) can be rewritten as an adjustment to the unconditional mean as

$$E\left[\mathcal{W}|I=1\right] = E\left(\mathcal{W}\right) \cdot \frac{\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]}{\Phi\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]}$$

and the other equations can be similarly recast.

2.4. Micro Regressions

We now use these aggregation results to build a micro-macro model of market wages, that accounts for nonparticipation in work. In summary, the underlying model is comprised of the following log-wage equation, an hours equation and a selection equation

$$\begin{aligned} \ln w &= \beta_0 + \beta' x + \epsilon, \\ h &= h_0 + \gamma \cdot \left(\alpha_0 + \alpha' z + \nu\right), \\ I &= 1\left[\alpha_0 + \alpha' z + \nu > 0\right]. \end{aligned} \tag{2.10}$$

where x refers to predictors in the log-wage equation, such as human capital variables that would represent δ_{js} in (2.2), or the predictors in the extended versions of the model.

To derive the implications of the behavioral model on individual level data (at time t), we require

Micro Assumption: (ϵ, v) is a joint normal random variable: namely

$$\begin{pmatrix} \epsilon \\ \nu \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\epsilon^2 & \sigma_{\epsilon\nu} \\ \sigma_{\epsilon\nu} & \sigma_\nu^2 \end{pmatrix} \right)$$

To use the results, such as Lemma 2.1, we apply the correspondence

$$\begin{aligned} \mathcal{U} &= \beta_0 + \beta' x + \epsilon, \\ \mathcal{V} &= -\alpha_0 - \alpha' z - \nu, \end{aligned} \tag{2.11}$$

and consider the population distributions conditional on the values of x and z. This gives

$$\begin{aligned} \mu_{\mathcal{U}} &= \beta_0 + \beta' x \\ \mu_{\mathcal{V}} &= -\alpha_0 - \alpha' z \\ \sigma_{\mathcal{U}}^2 &= \sigma_\epsilon^2 \\ \sigma_{\mathcal{U}\mathcal{V}} &= -\sigma_{\epsilon\nu} \\ \sigma_{\mathcal{V}}^2 &= \sigma_\nu^2 \end{aligned} \tag{2.12}$$

Using this correspondence, we can derive several familiar results. The micro participation regression has a standard probit form

$$E\left[I|x,z\right]=\Phi\left[\frac{\alpha_0+\alpha^{'}z}{\sigma_\nu^2}\right].$$

The micro log-wage regression for participants is

$$E\left[\ln w|I,x,z\right]=\beta_0+\beta^{'}x+\frac{\sigma_{\epsilon\nu}}{\sigma_\nu}\lambda\left[\frac{\alpha_0+\alpha^{'}z}{\sigma_\nu^2}\right], \tag{2.13}$$

familiar from Heckman (1979), among many others.[10] For purposes of later comparison, the log mean wage is

$$\ln E\left[w|I,x,z\right]=\beta_0+\beta^{'}x+\frac{1}{2}\sigma_\epsilon^2+\ln\left[\frac{\Phi_{\sigma_{\epsilon\nu}}}{\Phi}\right]$$

(from (2.8)), where

$$\left[\frac{\Phi_{\sigma_{\epsilon\nu}}}{\Phi}\right]=\left[\frac{\Phi\left[\frac{1}{\sigma_\nu}\left(\alpha_0+\alpha^{'}z+\sigma_{\epsilon\nu}\right)\right]}{\Phi\left[\frac{1}{\sigma_\nu}\left(\alpha_0+\alpha^{'}z\right)\right]}\right].$$

When hours are observed, we can likewise compute average hours and weighted average wages. Noting that $h=h_0-\gamma\mathcal{V}$, we have

$$E\left[h|I,x,z\right]=h_0+\gamma\alpha_0+\gamma\alpha^{'}z+\gamma\sigma_v\lambda\left[\frac{\alpha_0+\alpha^{'}z}{\sigma_\nu^2}\right].$$

Applying (2.9) gives

$$E\left[hw|I,x,z\right]=e^{\beta_0+\beta^{'}x+\frac{1}{2}\sigma_\epsilon^2}\cdot$$

$$\left(h_0+\gamma\alpha_0+\gamma\alpha^{'}z+\gamma\sigma_{\epsilon\nu}+\gamma\sigma_v\lambda\left[\frac{\alpha_0+\alpha^{'}z+\sigma_{\epsilon\nu}}{\sigma_\nu^2}\right]\right)\left[\frac{\Phi_{\sigma_{\epsilon\nu}}}{\Phi}\right]$$

[10]Here $\lambda\left[\cdot\right]=\phi\left[\cdot\right]/\Phi\left[\cdot\right]$ is the inverse Mills ratio, where ϕ and Φ are the standard normal density and c.d.f. respectively.

Therefore, we have

$$\frac{E\left[hw|I,x,z\right]}{E\left[h|I,x,z\right]}=e^{\beta_0+\beta' x+\frac{1}{2}\sigma_\epsilon^2}\cdot\left[\frac{\Lambda_{\sigma_{\epsilon\nu}}}{\Lambda}\right]\cdot\left[\frac{\Phi_{\sigma_{\epsilon\nu}}}{\Phi}\right]$$

where

$$\left[\frac{\Lambda_{\sigma_{\epsilon\nu}}}{\Lambda}\right]=\frac{h_0+\gamma\alpha_0+\gamma\alpha' z+\gamma\sigma_{\epsilon\nu}+\gamma\sigma_v\lambda\left[\frac{\alpha_0+\alpha' z+\sigma_{\epsilon\nu}}{\sigma_\nu^2}\right]}{h_0+\gamma\alpha_0+\gamma\alpha' z+\gamma\sigma_v\lambda\left[\frac{\alpha_0+\alpha' z}{\sigma_\nu^2}\right]}.$$

Taking logs, we have

$$\ln\left(\frac{E\left[hw|I,x,z\right]}{E\left[h|I,x,z\right]}\right)=\beta_0+\beta' x+\frac{1}{2}\sigma_\epsilon^2+\ln\left(\frac{\Lambda_{\sigma_{\epsilon\nu}}}{\Lambda}\right)+\ln\left(\frac{\Phi_{\sigma_{\epsilon\nu}}}{\Phi}\right).$$

These equations represent the implications of the micro model on individual level regression relationships.

2.5. Macroeconomic Equations

Because we have extensive individual level data on wages, we can model aggregate wages by "adding up" the respective terms; namely microsimulation. However, it is useful to derive specific representations of the impact of participation and hours heterogeneity, and for this we need an assumption on the distribution of the micro variables x and z in the population for a given time period t. We make the following distributional assumption, which is not only convenient but (as we show) reasonably accurate in our applications.[11]

Distributional Restriction: The indexes determining log wages and participation are joint normally distributed: namely

$$\begin{pmatrix}\beta_0+\beta' x\\ \alpha_0+\alpha' z\end{pmatrix}\sim\mathcal{N}\left(\begin{pmatrix}\beta_0+\beta' E\left(x\right)\\ \alpha_0+\alpha' E\left(z\right)\end{pmatrix},\begin{pmatrix}\beta'\Sigma_{xx}\beta & \alpha'\Sigma_{xz}'\beta\\ \beta'\Sigma_{xz}\alpha & \alpha'\Sigma_{zz}\alpha\end{pmatrix}\right)$$

[11]Since we utilize many discrete regressors in our application (cohort and education indicators), it is important that the normal distribution assumption is on the indexes $\beta_0+\beta' x$, $\alpha_0+\alpha' z$. If this assumption only applies within different population segments, then our equations could be applied segment by segment, and aggregated across segments to form the final specification of aggregate wages.

To use our aggregation results, including Lemma 2.1, we apply the same correspondence (2.11), slightly rewritten as

$$\begin{array}{l} \mathcal{U} = \beta_0 + \beta' E(x) + \beta' (x - E(x)) + \epsilon \\ \mathcal{V} = -\alpha_0 - \alpha' E(z) - \alpha' (z - E(z)) - \nu \end{array} \tag{2.14}$$

and consider (unconditional) expectations over the joint distribution of x and z and the disturbances ϵ and ν. This gives

$$\begin{array}{l} \mu_{\mathcal{U}} = \beta_0 + \beta' E(x) \\ \mu_{\mathcal{V}} = -\alpha_0 - \alpha' E(z) \\ \sigma_{\mathcal{U}}^2 = \beta' \Sigma_{xx} \beta + \sigma_\epsilon^2 \\ \sigma_{\mathcal{U}\mathcal{V}} = -\beta' \Sigma_{xz} \alpha - \sigma_{\epsilon\nu} \\ \sigma_{\mathcal{V}}^2 = \alpha' \Sigma_{zz} \alpha + \sigma_\nu^2 \end{array} \tag{2.15}$$

for use in the aggregation results.

We derive the macroeconomic participation equation as

$$E[I] = \Phi\left[\frac{\alpha_0 + \alpha' E(z)}{\sqrt{\alpha' \Sigma_{zz} \alpha + \sigma_\nu^2}}\right],$$

which is in the same form as the micro participation equation with z replaced by $E(z)$ and the spread parameter σ_ν replaced by the larger value $\sqrt{\alpha' \Sigma_{zz} \alpha + \sigma_\nu^2}$, that reflects the influence of heterogeneity in the predictors in the selection criteria.[12] Because[13]

$$E\left[\beta' x | I = 1\right] = \beta' E(x) + \frac{\beta' \Sigma_{xz} \alpha}{\sqrt{\alpha' \Sigma_{zz} \alpha + \sigma_\nu^2}} \lambda \left[\frac{\alpha_0 + \alpha' E(z)}{\sqrt{\alpha' \Sigma_{zz} \alpha + \sigma_\nu^2}}\right],$$

we can get an interesting formula

$$E[\ln w | I = 1] = \beta_0 + \beta' E(x | I = 1) + \frac{\sigma_{\epsilon\nu}}{\sqrt{\alpha' \Sigma_{zz} \alpha + \sigma_\nu^2}} \lambda \left[\frac{\alpha_0 + \alpha' E(z)}{\sqrt{\alpha' \Sigma_{zz} \alpha + \sigma_\nu^2}}\right],$$

[12] This formula was first derived by McFadden and Reid (1975)

[13] A formula of this form was originally derived by McCurdy (1987).

which has the same form as the micro equation (2.13), with the spread parameter σ_ν changed to $\sqrt{\alpha'\Sigma_{zz}\alpha + \sigma_\nu^2}$.

If there were no variation in hours (i.e. if hours weights were equal across individuals), by Lemma 2.1, the appropriate macroeconomic wage equation would be

$$E\left[w|I=1\right] = e^{\beta_0+\beta' E(x)+\frac{1}{2}\left[\beta'\Sigma_{xx}\beta+\sigma_\epsilon^2\right]}\left[\frac{\Phi^a_{\sigma_{\epsilon\nu}}}{\Phi^a}\right]$$

$$\text{with } \left[\frac{\Phi^a_{\sigma_{\epsilon\nu}}}{\Phi^a}\right] = \left[\frac{\Phi\left[\frac{\alpha_0+\alpha' E(z)+\left(\beta'\Sigma_{xz}\alpha+\sigma_{\epsilon\nu}\right)}{\sqrt{\alpha'\Sigma_{zz}\alpha+\sigma_\nu^2}}\right]}{\Phi\left[\frac{\alpha_0+\alpha' E(z)}{\sqrt{\alpha'\Sigma_{zz}\alpha+\sigma_\nu^2}}\right]}\right]$$

For later comparison, we can write the log of mean wage as

$$\ln E\left[w|I=1\right] = \beta_0 + \beta' E\left(x\right) + \frac{1}{2}\left[\beta'\Sigma_{xx}\beta + \sigma_\epsilon^2\right] + \ln\left[\frac{\Phi^a_{\sigma_{\epsilon\nu}}}{\Phi^a}\right]$$

Turning to hours, we again have $h = h_0 - \gamma\mathcal{V}$ in (2.14), so that average hours are

$$E\left[h|I=1\right] = h_0 + \gamma\alpha_0 + \gamma\alpha' E\left(z\right) + \gamma\sqrt{\alpha'\Sigma_{zz}\alpha + \sigma_\nu^2}\cdot\lambda^a$$

in which

$$\lambda^a = \lambda\left[\frac{\alpha_0 + \alpha' E\left(z\right)}{\sqrt{\alpha'\Sigma_{zz}\alpha + \sigma_\nu^2}}\right].$$

To allow for variations in hours, we begin with the implication of (2.9) of Lemma 2.1:

$$E\left[hw|I=1\right] = e^{\beta_0+\beta' E(x)+\frac{1}{2}\left[\beta'\Sigma_{xx}\beta+\sigma_\epsilon^2\right]}$$

$$\cdot\left\{h_0 + \gamma\alpha_0 + \gamma\alpha' E\left(z\right) + \gamma\beta'\Sigma_{xz}\alpha + \gamma\sigma_{\epsilon\nu} + \gamma\sqrt{\alpha'\Sigma_{zz}\alpha + \sigma_\nu^2}\lambda^a_{\sigma_{\epsilon\nu}}\right\}\left[\frac{\Phi^a_{\sigma_{\epsilon\nu}}}{\Phi^a}\right]$$

in which

$$\lambda^a_{\sigma_{\epsilon\nu}} = \lambda\left[\frac{\alpha_0 + \alpha' E(z) + \beta'\Sigma_{xz}\alpha + \sigma_{\epsilon\nu}}{\sqrt{\alpha'\Sigma_{zz}\alpha + \sigma^2_\nu}}\right].$$

Drawing these results together we have that

$$\frac{E[hw|I=1]}{E[h|I=1]} = e^{\beta_0 + \beta' E(x) + \frac{1}{2}\left[\beta'\Sigma_{xx}\beta + \sigma^2_\epsilon\right]}\left[\frac{\Lambda^a_{\sigma_{\epsilon\nu}}}{\Lambda^a}\right]\left[\frac{\Phi^a_{\sigma_{\epsilon\nu}}}{\Phi^a}\right]$$

where we have defined the hours adjustment term

$$\frac{\Lambda^a_{\sigma_{\epsilon\nu}}}{\Lambda^a} \equiv \frac{h_0 + \gamma\alpha_0 + \gamma\alpha' E(z) + \gamma\beta'\Sigma_{xz}\alpha + \gamma\sigma_{\epsilon\nu} + \gamma\sqrt{\alpha'\Sigma_{zz}\alpha + \sigma^2_\nu}\cdot\lambda^a_{\sigma_{\epsilon\nu}}}{h_0 + \gamma\alpha_0 + \gamma\alpha' E(z) + \gamma\sqrt{\alpha'\Sigma_{zz}\alpha + \sigma^2_\nu}\cdot\lambda^a}.$$

In conclusion, log aggregate wages are given as

$$\ln\frac{E[hw|I]}{E[h|I]} = \beta_0 + \beta' E(x) + \tfrac{1}{2}\left[\beta'\Sigma_{xx}\beta + \sigma^2_\epsilon\right]$$
$$+\ln\left[\frac{\Lambda^a_{\sigma_{\epsilon\nu}}}{\Lambda^a}\right] + \ln\left[\frac{\Phi^a_{\sigma_{\epsilon\nu}}}{\Phi^a}\right].$$

To summarize, there are three aggregation factors that need to be accounted for in examining the evolution of aggregate wages. The first term, $\frac{1}{2}\left[\beta'\Sigma_{xx}\beta + \sigma^2_\epsilon\right]$, describes the variance of returns (observable and unobservable). The second term, $\ln\left[\Lambda^a_{\sigma_{\epsilon\nu}}/\Lambda^a\right]$, measures the adjustment for composition changes in hours and depends on the size of the covariance between wages and hours. The final term, $\ln\left[\Phi^a_{\sigma_{\epsilon\nu}}/\Phi^a\right]$, highlights the importance of composition changes within the selected sample of workers from which measured wages are recorded. As in the standard selection bias literature, it too depends on the covariance between participation and wages.

2.6. Remarks on the Nature of the Aggregation Bias

To anticipate our application, we now discuss various features of how the aggregation biases can manifest themselves in data on labor participation and wages. Setting $\beta_0 + \beta' x_{it} = \ln r_t + \delta_{st}$ in (2.10) generates our baseline formulation (2.2). Participation follows the simple reservation wage rule (2.4), that is

$$\begin{aligned} \Pr[I_{it} > 0] &= \Phi\left(\frac{\ln r_t - \alpha \ln b_{it} + \delta_{js} - \eta_{js}}{\sigma_v}\right) \\ &\equiv \Phi\left(\frac{\alpha_0 + \alpha_1' z_{it}}{\sigma_v}\right). \end{aligned} \tag{2.16}$$

The time series evolution of the log aggregate hourly real wage, measured among workers, is characterized by

$$\begin{aligned} \ln \overline{w}_t = \ln r_t + E(\delta_{st}) + \tfrac{\sigma^2_{\delta,t} + \sigma^2_{\epsilon,t}}{2} \\ + \ln\left[\tfrac{\Lambda^a_{\sigma_{\epsilon v},t}}{\Lambda^a_t}\right] + \ln\left[\tfrac{\Phi^a_{\sigma_{\epsilon v},t}}{\Phi^a_t}\right] \end{aligned} \tag{2.17}$$

The latter term is the adjustment to the aggregate wage to allow for the selectivity on unobservable attributes ϵ_{it} in the log wage equation induced by participation.

Focusing on the aggregation factor $\ln\left[\Phi^a_{t\sigma_{\epsilon v},t}/\Phi^a_t\right]$, for the typical case in which $\sigma_{\epsilon v} > 0$, selection induces an upward bias in the average wage. Consider what happens as the return $\ln r_t$ increases over time with $E(\delta_{st})$ constant. For $\sigma_{\epsilon v} > 0$ this results in a decrease in $\ln\left[\Phi_{\sigma_{\epsilon v},t}/\Phi_t\right]$ and the corresponding downward bias in the average wage. Aggregation can therefore offset the procyclicality of wages through the entry in the upturn of individuals drawn from lower values of unobserved attributes ϵ_{it}. That is

$$\begin{aligned} d \ln E\left[w_t | I_t\right] &= d \ln r_t + d \ln \left[\tfrac{\Phi_{\sigma_{\epsilon v},t}}{\Phi_t}\right] \\ &= (1 + \lambda_{\sigma_{\epsilon v},t} - \lambda_t) d \ln r_t \end{aligned}$$

The composition bias term

$$\lambda_{\sigma_{\epsilon\nu},t} - \lambda_t = \left(\frac{\phi_{\sigma_{\epsilon\nu},t}}{\Phi_{\sigma_{\epsilon\nu},t}} - \frac{\phi_t}{\Phi_t} \right)$$

is negative for a increase in $\ln r_t$ over time since

$$\left(\frac{\phi_{\sigma_{\epsilon\nu},t}}{\Phi_{\sigma_{\epsilon\nu},t}} - \frac{\phi_t}{\Phi_t} \right) = \left(\frac{\phi_t \Phi_{\sigma_{\epsilon\nu},t} - \phi_{\sigma_{\epsilon\nu},t} \Phi_t}{\Phi_t \Phi_{\sigma_{\epsilon\nu},t}} \right) \leq 0 \text{ for } \sigma_{\epsilon v} \geq 0.$$

This analysis is easily extended to the case of two (or more) education or skill groups. Suppose there is a decrease in returns for the lower skilled workers. That is, suppose $\ln r_t^L$ in (2.6) falls. The decline in r^L reduces participation among lower skilled workers and the conditional wage may rise, since the remaining participants will be a more severely selected sample with higher ϵ_{it} on average. This implies that the average wage could show growth *even though* $\ln r_t^L$ *is declining.*

3. British Aggregate Wages and Participation

3.1. The Data

The microeconomic data used for this study are taken from the UK Family Expenditure Survey (FES) for the years 1978 to 1995. The FES is a repeated continuous cross-sectional survey of households which provides consistently defined micro data on wages, hours of work, employment status and education for each year since 1978.[14] Our sample consists of all men aged between 18 and 59 (inclusive).[15] For the purposes of modeling, the participating group consists of employees; the non-participating group includes individuals categorized as searching for work as well as the unoccupied. The hours measure for employees in FES

[14]Prior to 1978 the FES contains no information on educational attainment.

[15]We exclude individuals classified as self-employed. This could introduce some composition bias, given that a significant number of workers moved into self employment in the 1980's. However, given that we have no data on hours and relatively poor data on earnings for this group, there is little alternative but to exclude them.

is defined as usual weekly hours including usual overtime hours. The weekly earnings measure includes usual overtime pay. We divide nominal weekly earnings by weekly hours to construct an hourly wage measure, which is deflated by the quarterly UK retail price index to obtain real hourly wages. The measure of education used in our study is the age at which the individual left full-time education. Individuals are classified in three groups; those who left full-time education at age 16 or lower (the base group), those who left aged 17 or 18, and those who left aged 19 or over.[16] We model cohort effects on wage levels by a set of cohort dummies; five date-of-birth cohorts (b.1919-34, b.1935-44, b.1945-54, b.1955-64 and b.1965-77).

Our measure of out-of-work income (income at zero-hours) is constructed for each individual as follows. This measure is evaluated using the tax and benefit simulation model[17], which constructs a simulated budget constraint for each individual given information about his age, location, benefit eligibility and partner's income (if married/cohabiting). The measure of out-of-work income is largely comprised of income from state benefits; only small amounts of investment income are recorded. For married men we do not include the spouse's income from employment. We control for the spouse's characteristics, in particular her level of education and full set of interactions between, age, region and calendar time. State benefits include eligible unemployment benefits[18] and housing benefits.

Since our measure of out-of-work income will serve to identify the participation

[16] An alternative to our method for constructing the education dummy would use those who left education at the statutory minimum age as the base group. This method is equivalent to ours from 1973 onwards in the UK; before this date the minimum school leaving age was a year lower, at 15. Nonetheless, interactions between date-of-birth cohort effects and the education dummy will capture any effects of the change in minimum leaving age on the relative returns to education enjoyed by the 17+ group. See Gosling et. al (1996).

[17] The IFS tax and benefit simulation model TAXBEN (see Giles and McRae (1995)), designed for the British Family Expenditure Survey data used in this paper.

[18] Unemployment benefit included an earnings-related supplement in 1979, but this was abolished in 1980.

structure, it is important that variation in the components of out-of-work income over the sample is as exogenous to the decision to work or the level of wages as possible. In the UK, the level of benefits which individuals receive out of work varies with age, time, household size and (in the case of Housing Benefit) by region. As mentioned before, housing benefit varies systematically with time, location and cohort. One of the primary features of Housing Benefit is that older cohorts had much higher availability of public housing during their household formation period and would have been likely to stay in public housing. Since 1978 the rents in public housing have risen dramatically. For those out of work, Housing Benefit would have covered these increases, which may have had the effect of increasing the reservation wage for those in public housing.

After making the sample selections described above, our sample contains 50,825 observations. The number of employees in the data is 41,290, or 81.2% of the total sample. Tables 3.1 and 3.2 provide a description of the cell proportions by marital status and education level over the period of our analysis. As Table 3.1 shows, the proportions of single and married men in the data are relatively constant over the 1980s and 1990s, although there are rather less single men in 1979. Table 3.2 shows that single men are on average slightly better educated than the married men in the sample.

3.2. Results

We consider a number of possible specifications for our individual level participation and wage equations which relate to the various specifications discussed in Section 2.[19] Our model of participation includes out-of-work income interacted with marital status, as well as the variables included in the log wage equation.

[19]A full set of results is available from the authors. It also appears as Appendix B in the Institue for Fiscal Studies (www.ifs.org.uk) working paper version.

The results of estimating the participation (probit) equation show a strong significance of this benefit income variable. This is important as it is our primary source of identification.[20] The sheer number of interactions makes it hard to discern the impact of the various regressors, and we conduct joint significance tests for sets of regressors and interactions between them. These are presented in Table 3.3 for the participation probit and the wage equation with the selectivity correction via the inverse Mills ratio.

In estimation we are unable to use data on housing benefit for the years 1980-1983. This is because the Family Expenditure Survey does not appear to contain sufficient information to accurately calculate benefit entitlement for those years. We do, however, have a consistent series for 1978-1979 and the period 1984-1995. Below we present results for the complete period 1978-1995 but exclude results that rely on the benefit variable for the 1980-1983 period.

Our chosen specification, which the results below focus on, models participation and wages as a function of the three education groupings, cohort dummies, a cubic trend, and region, plus interactions between the cubic trend and education, cubic trend and cohort, education and cohort, linear trend by education and cohort, and a quadratic trend times region. This specification was chosen in comparison to a number of alternatives through a standard specification search.[21] Further details of the validation of this model are presented in the model validation section below.

The necessity of the inclusion of the interaction terms means that our preferred specification of the log wage equation departs from the full proportionality hypothesis as set out in Section 2. The additional interactions between cohort

[20] The full results are available on request.

[21] It is also in accordance with much of the literature on the evolution of British male wages (see Meghir and Whitehouse (1996), for example).

and education and trend which we introduce could reflect many differences in minimum educational standards across cohorts such as the systematic raising of the minimum school leaving age over the postwar period in the UK. Meanwhile the prices of different (education level) skills are allowed to evolve in different ways, by including an interaction between the education dummies and the trend terms. The selectivity correction using the inverse Mills ratio from the participation equation is interacted with marital status and by education group, because first, the way out-of-work income is defined implies that it is at quite different levels for single and married people, and second, it is quite possible that selection may have had different effects at different skill levels. As Table 3.3 shows the benefit income terms are strongly significant in the participation equation and the Mills ratio, education, cohort and trend terms are all significant in the wage equation.

3.2.1. Aggregate Wages and Corrections: Overall Sample Measures

We now consider aggregate wages and the corrections due to heterogeneity, the distribution of hours and labor participation.[22] We plot the values over time, to allow a quick assessment of the path of aggregate wages and the relative importance of the corrections, as well as how well the corrected aggregate wage matches up with the mean log wage implied by the micro-level wage equations. We have found this graphical approach much more straightforward than trying to directly analyze the numerous estimated coefficients underlying the graphs.

Overall aggregate wages and the various correction terms are plotted in Figure 3.1. The upper panel of Figure 3.1 displays the behavior of all the measures of wages we look at over the entire period. First there is the selectivity-adjusted

[22]The disturbance "variance" terms are computed by standard variance estimates from the estimated truncated regression structure.

prediction from the micro-level wage equation. Second, there is the aggregate measure of wages calculated as the log of average wages for those in work.[23] The remaining three lines shown on the figure give the (cumulative) application of the correction terms to aggregate wages. First is the correction for the distribution of hours. As we may have expected given the relatively stable pattern of hours worked, this has little impact on the time-series evolution of wages. Second is the selection correction for covariance between wages and participation. This has a more dramatic effect, with growing gaps over time associated with large decreases in participation. Finally, we apply the correction for the heterogeneity (dispersion) of individual wages. This gives the impact of the increasing heterogeneity in wages that is separated from participation effects.

In sum, this final series gives the aggregate wage after all corrections.[24] For comparison, we plot the mean log wage implied by the micro regressions (adjusted for participation, or omitting the selection term). Finally, in order to see the relative growth of the various series more clearly, the lower panel of Figure 3.1 shows exactly the same series for the micromodel prediction, the aggregate wage measure and the fully-corrected aggregate series, but rebased to 1979.[25] Plotting each series starting at the 1979 level makes it easier to see what the implementation of the adjustment formula does to the measured aggregate hourly earnings growth.

A key evaluation of our framework is whether the fully corrected aggregate series lines up with the selectivity-adjusted micromodel prediction. The upper panel of Figure 3.1 shows that there is a very close correspondence between the series.

[23]This is also calculated from the FES and corresponds closely to the measure of 'average earnings' which media commentators in the UK have focused on.

[24]Although we only have benefit information for the sample in 1979 and the years from 1984 through 1995 and hence we are not able to compute the selection correction for these years or include them in the figure. We do plot the micromodel prediction and the uncorrected aggregate series for 1978 and 1980-83.

[25]That is, the 1979 values are subtracted from all values in the series.

Later on we use bootstrap methods to check whether any difference which does arise between the micromodel and the corrected aggregate series is statistically significant.

Several features of this figure are noteworthy. For instance, the direction of movement of the uncorrected log aggregate wage does not always mirror that of the mean micro log wage. During the recession of the early 1980s, aggregate wages grow rather more than the corrected micromodel wage. Whilst there is a reasonably close correspondence between the trend of the two lines in the latter half of the 1980s, in the 1990s we find that there is a reasonably substantial increase in log aggregate wages but essentially no growth in the corrected measure. The lower panel of Figure 3.1, which rebases to 1979, shows these patterns even more vividly. Correcting for selection over the period reduces our estimate of real aggregate wage growth from about 30% to less than 20%.

3.2.2. Wage Measures by Education Group

Next we break our sample up by the three education groups used in the analysis. We plot the wage series defined just as before but this time we are taking the micromodel prediction, the 'aggregate' wage series and the corrections to the aggregate series *within education group* for each year. Hence we have three plots in Figure 3.2, which present the path of the series for each education group.

For the low education group who are those that left full time education at age 16 or younger, the picture is particularly clear. This is presented in the first panel of Figure 3.2. Controlling for the biases induced by shifts in participation rates over the 1980s and 1990s reduces our estimate of average wage growth for this group from over 20% to around 10%. The corrected aggregate series and the selectivity-adjusted micromodel prediction appear to line up very well here.

For those individuals with more schooling, presented in the subsequent two panels of figure 3.2, the fit between the two series is less good largely because these are smaller subsamples, and so the data on wages for them is more noisy. Nevertheless, there appears to be evidence that selection effects do bias measured wage growth estimates upwards for both of the better-educated groups.

3.2.3. A Regional Breakdown

There are several further breakdowns of the FES wage data which are interesting to look at in our framework in addition to the split by educational group. Regional differences in real wages and labor market participation are characteristic of Britain as they are of many European economies. We examine differences in the path of measured average wages and the wages predicted by our micromodel and corrections to the average measure for two broad regions, the 'North' and the 'South' of Britain[26].

As Figure 3.3 shows, the two regions experienced marked differences in male wages and participation over this period. In 1978 participation for the South was only around 3-4 percentage points higher than it was in the North. By 1983 this North-South gap had widened to more than 10% as the North was affected a lot more severely by the decline of traditional manufacturing sectors than was the South (mainly because the old industries were mainly located in the North). Growth in participation in the late 1980s in the North then closed some of the increase in the gap, and in the 1990s recession both regions appear to have been affected a lot more equally. The lower panel of Figure 3.3 shows that wages grew faster on average in the South than they did in the North over the 1980s; in the

[26]More precisely, our definition of the 'North' comprises the FES standard regions Northeast, Northwest, Yorkshire & Humberside, West Midlands, Wales and Scotland. The 'South' comprises London and the Southeast. The Southwest, East Midlands and East Anglia are omitted.

1990s the experience of both regions has been relatively similar.

Figure 3.4 presents the corrected figures. For the North, there is much slower growth in the early eighties than the aggregate figures portray and a reasonably continuous divergence between the uncorrected aggregate wage measure and the micromodel prediction from 1979 until 1995. The corrected aggregate measure tracks the micromodel prediction closely for the most part. In the South, the aggregate measure and the micromodel prediction grow at a similar rate between 1979 and 1990, although there are some fluctuations around the trend for the aggregate measure. After 1990, the gap between the two measures opens out as falling participation increases the importance of selection. Indeed in the 1990s, the corrected figure indicates that average wages actually *fell* back in the South. Again there is a close correspondence of the corrected aggregate measure and the micromodel prediction although there is some divergence between the two in the mid-80s. Selection biases induced by differential employment behavior in the North and the South of Britain appear to indicate that the behavior of individual wages was very different from that which would be surmised from the aggregate figures.

3.2.4. A Lifecycle Perspective: The Cohort Breakdown

Disaggregating wages and the pattern of wage growth reveals another important aspect of the impact of participation on aggregate wages. In Figure 3.5 the participation rate and real wages by date-of-birth cohort are presented. The employment rate fell sharply over the period for the oldest cohort in our study, born between 1919 and 1934, a fall which coincides with the onset of early retirement for many members of this group. This decline in employment is mirrored to a lesser extent in the next oldest cohort (born 1935-44). The other cohorts show more of a cyclical movement in participation, with a slight downward trend (except for the youngest cohort, born between 1965 and 1977). Meanwhile, the lower panel of Figure 3.5 shows that the younger a cohort was, the higher the rate of wage growth it achieved over the sample period. With these facts in mind, we performed the adjustments to the aggregate wage measure within each cohort group and compared them with the selectivity-adjusted micromodel wage prediction for the same cohort group. This exercise produced very different results according to the cohort studied.

Figure 3.6 graphs the results for two different cohorts: those born between 1935 and 1944 (who were the oldest cohort with representatives in every sample year) and those born between 1955 and 1964 (who were the youngest). For the older cohort in the upper panel, selection effects are clearly very important. Wage growth over the sample period is predicted to be about 25% on the basis of the unadjusted aggregate measure, whereas the micromodel predicts less than 10%. In words, as the cohort ages, the composition of those remaining in employment changes in such a way as to bias up the estimated real hourly wage – this might also be interpretable as an upward bias in the returns to experience.

3.3. Model Validation

Our model and the econometric assumptions underlying have been tested as far as is possible in order to ascertain their plausibility. The validation procedures undertaken include a check to see whether the corrections to the aggregate wage measure line it up sufficiently well with the predictions from the selectivity - adjusted micromodel, relaxing the normality assumption on the unobservables by estimating an analogous model using semiparametric methods, and plots of the predicted indices from the probit and the wage equation to assess whether the distributions of observable attributes conform to normality. We now assess each of these in turn.

3.3.1. Bootstrapping the Accuracy of the Model Fit

To assess the accuracy with which the corrections which we make to the aggregate average male log wage series 'line up' against the prediction from our micro-model of wages (with the selectivity correction included), we used bootstrap methods to simulate the difference between the two measures[27]. These pictures are shown in Figure 3.7 for the overall sample (upper panel), and broken down by educational group (lower panel). They show that the difference between the two measures is not significantly different from zero in most of the years covered by the sample. Occasionally the difference is significantly positive (indicating that the corrected aggregate measure is higher than the micromodel prediction), particularly for the best educated group in 1989 and 1992, but in general the corrections to the aggregate measure and the selectivity-adjusted micromodel line up very well. This provides a very positive validation of the model framework.

3.3.2. Semiparametric Estimation

Our model, as set out in Section 2, makes the assumption that the unobservable factors affecting participation and wages are normally distributed. This can of course be called into question. The properties of the estimator rely on the parametric distributional assumptions on the joint distribution of the errors. However, given our exclusion assumption on the continuous out-of-work income variable, semiparametric estimation can proceed in a fairly straightforward manner. To estimate the slope parameters we follow the suggestion of Robinson (1988)

[27]The number of repetitions in the bootstrap simulation was 500.

which is developed in Ahn and Powell (1993). These techniques are explored in a useful application to labor supply by Newey, Powell and Walker (1990). In Figure 3.8 we graph a comparison between the predicted wages estimated using semiparametric techniques and the wage predictions from the selectivity-adjusted micromodel which we use. Bootstrap confidence bands (95%) refer to the parametric selectivity model. The corrected aggregate wage measure is also plotted. The upper panel shows the results for the overall sample, and the lower panel those for the lower education group. In both cases there is a very close correspondence between the predictions from the parametric micromodel and the semiparametric version. We conclude that the assumption of normality of the unobservables in the model is not unduly restrictive.

3.3.3. Normality of the Wage and Participation Indexes

In addition to checking the validity of the normality assumption on the unobservables, we are also interested in the normality of the probit index and of the fitted wage distribution from the selectivity-adjusted wage equation. Taking the participation probit first of all, Figure 3.9 plots the distribution standardized probit index $\widehat{\alpha}'z$ over all years of the sample (plots for individual years are all quite similar). The index is distributed roughly normally although with a slight negative skew.[28]

We also checked the validity of the normality assumption on log wages by plotting the standardized wage predictions from the model overlaid with a standard normal curve. This is shown in Figure 3.10. The distribution is not obviously skewed left or right, and there appears to be a higher density of observations around the mean than is the case with a standard normal. In any case, while

[28] For further validation, kernel regressions of participation on $\widehat{\alpha}'z$ show a normal shape, details of which are available from the authors on request.

these plots do not show exact concordance with the normal distribution assumptions, we feel that the proximity of the empirical distributions to normal helps explain the close correspondence between corrected aggregate wages and the mean wages implied by the micro regressions.[29]

4. Conclusion

This aim of this paper has been to provide a systematic assessment of the way changes in labor market participation affect our interpretation of aggregate real wages. We have developed and implemented an empirical framework for understanding this relationship which reduces to the calculation of three aggregation factors. These can be interpreted as correction terms reflecting changes in selection due to participation, changes in the distribution of returns and changes in hours of work, respectively. We have shown that they do a remarkably good job of explaining the differences between individual and aggregate wages in the British context.

British data was used for three reasons. First, there have been significant changes in labor market participation over the last two decades. Participation rates for men have seen a secular decline and have displayed strong cyclical variation. The secular decline is largely reflected in increasing decline in participation among older men across cohorts while the cyclical variation shows strong regional variation. This phenomena is common to many other developed economies. The second argument for studying British wages is that there are strong changes in real wages and the distribution of real wages over this period. Finally, there is important exogenous variation in certain components of out of work incomes across

[29] While there are some visible departures from normality, the entire impact of those departures on the analysis is summerized in the difference between the plots from the corrected aggregate measure and the micro model. As we have noted above these plots are extremely close.

time and across individuals that allows the identification of the correction terms.

The empirical analysis of aggregate wages is shown to provide a coherent picture of the relationship between individual male wages and aggregated wages over this period. Moreover, the statistical model adopted appears to accord well with the empirical facts. The correction terms explaining the differences between log aggregate wages and the average of log wages implied by our analysis. And the differences are interesting, and have valuable implications. Most noteworthy is how mean individual log-wages are largely flat throughout the early 1990's, whereas measured aggregate wages are rising. As such, we see our estimates as giving fairly clear evidence that the biases in log aggregate real wages are substantial and can lead to misleading depictions of the progress of wages of individual male workers.

References

[1] Ahn, Hyungtaik and Powell, James L. (1993),"Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism", *Journal of Econometrics,* 58, 3-29.

[2] Bils, Mark J. (1985), "Real Wages Over the Business Cycle: Evidence from Panel Data," *Journal of Political Economy*, 93, 666-689.

[3] Cogan, John F. (1981), "Fixed Costs and Labor Supply", *Econometrica*, 49, 945-964.

[4] Gosling, Amanda, Machin, Stephen and Meghir, Costas, (1996), "The Changing Distribution of Male Wages in the UK, 1966-92" IFS Working Paper No. W94/13.

[5] Heckman, James J. (1979), 'Sample Selection Bias as a Specification Error', *Econometrica*, 47, 153-161.

[6] Heckman , James J. and Sedlacek, Guilherme, (1985), "Heterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market", *Journal of Political Economy*, 93, 1077-1125.

[7] Heckman , James J. and Sedlacek, Guilherme, (1990), "Self-Selection and the Distribution of Hourly Wages", *Journal of Labor Economics*, 8, S329-S363.

[8] MaCurdy, Thomas E. (1987), "A Framework for Relating Microeconomic and Macroeconomic Evidence on Intertemporal Substitution," in T.F. Bewley, ed., *Advances in Econometrics, Fifth World Congress*, vol. II, 149-176.

[9] McFadden, Dan. and F. Reid (1975), "Aggregate Travel Demand Forecasting From Disaggregated Behavioral Models," Transportation Research Record, No. 534.

[10] Meghir, Costas, and Whitehouse, Edward, (1996), 'The Evolution of Male Wages in the UK: Evidence from Micro Data', *Journal of Labor Economics*, 14(1), 1-25.

[11] Newey, Whitney K., Powell, James L. and Walker, James R. (1990), 'Semiparametric Estimation of Selection Models: Some Empirical Results', *American Economic Review*, 80(2), 324-328.

[12] Robinson, Peter M. (1988), "Root-N Consistent Semiparametric Regression", *Econometrica*, 56, 931-954.

[13] Roy, Andrew D. (1951),"Some Thoughts on the Distribution of Earnings", *Oxford Economic Papers*, 3, 135-146.

[14] Schmitt, James, (1995),"The Changing Structure of Male earnings in Britain, 1974-1988", in *Differences and Changes in Wage Structures*, University of Chicago Press: Chicago.

[15] Solon, Gary, Barsky, Robert and Parker, Jonathan A. (1994), "Measuring the Cyclicality of Real Wages: How Important is Composition Bias," *Quarterly Journal of Economics*, 109, 1-25.

A. Appendix A: Proof of Lemma 2.1 and Related Results

Under the conditions of Lemma 2.1, we begin with some familiar derivations, and then proceed to the Proof. The event $\mathcal{V} < 0$ is equivalent to the event $(\mathcal{V} - \mu_{\mathcal{V}})/\sigma_{\mathcal{V}} < -\mu_{\mathcal{V}}/\sigma_{\mathcal{V}}$, so that

$$E[I] = \Phi\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]$$

follows by definition where $\Phi[\cdot]$ is the standard normal c.d.f.

Note first that

$$\frac{\partial}{\partial\eta}\left[e^{-\frac{\eta^2}{2\sigma^2}}\right] = -\frac{\eta}{\sigma^2}e^{-\frac{\eta^2}{2\sigma^2}}.$$

With

$$\begin{aligned} u &\equiv \mathcal{U} - \mu_{\mathcal{U}} \\ v &\equiv \mathcal{V} - \mu_{\mathcal{V}} \end{aligned} \tag{A.1}$$

normality implies that

$$u = \frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v + s \tag{A.2}$$

where s is independent of v. Therefore,

$$\begin{aligned} E(uI) &= \tfrac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}\int_{-\infty}^{-\mu_{\mathcal{V}}}\tfrac{v}{\sqrt{2\pi}\sigma_{\mathcal{V}}}e^{-\frac{v^2}{2\sigma_{\mathcal{V}}^2}}dv \\ &= -\tfrac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\tfrac{1}{\sqrt{2\pi}}e^{-\frac{(-\mu_{\mathcal{V}})^2}{2\sigma_{\mathcal{V}}^2}} \\ &= -\tfrac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\phi\left(\tfrac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right). \end{aligned} \tag{A.3}$$

Noting that $E(\mathcal{U}I) = \mu_{\mathcal{U}}E(I) + E(uI)$ and $E(\mathcal{U}|I) = E(\mathcal{U}I)/E(I)$, we have that

$$E[\mathcal{U}I] = \mu_{\mathcal{U}}\Phi\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right] - \frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\phi\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right] \tag{A.4}$$

where $\phi[\cdot]$ is the standard normal density function. Consequently, we have

$$E[\mathcal{U}|I=1] = \frac{E[\mathcal{U}I]}{E[I]} = \mu_{\mathcal{U}} - \frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\lambda\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right] \tag{A.5}$$

where $\lambda[\cdot] = \phi[\cdot]/\Phi[\cdot]$ is the inverse Mill's ratio.[30]

Applying (A.4) to the case with $\mathcal{U} = a + b\mathcal{V}$ gives

$$E[(a + b\mathcal{V})I] = (a + b\mu_{\mathcal{V}})\Phi\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right] - b\sigma_{\mathcal{V}}\phi\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right]$$

and

$$E[(a + b\mathcal{V})|I = 1] = (a + b\mu_{\mathcal{V}}) - b\sigma_{\mathcal{V}}\lambda\left[\frac{-\mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}}\right].$$

A.1. Proof of Lemma 2.1

For A, first note that

$$E[\mathcal{W}I] = e^{\mu_{\mathcal{U}}}E(e^{s})E\left(e^{\frac{\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}^2}v}I\right) \tag{A.6}$$

since s (of (A.2)) is independent of v. Now

$$\begin{aligned} E(e^{s}) &= e^{E(s)+\frac{1}{2}\sigma_s^2} \\ &= e^{\frac{1}{2}\sigma_{\mathcal{U}}^2(1-\rho_{\mathcal{U}\mathcal{V}}^2)} \end{aligned} \tag{A.7}$$

where $\rho_{\mathcal{U}\mathcal{V}} = \sigma_{\mathcal{U}\mathcal{V}}/\sigma_{\mathcal{U}}\sigma_{\mathcal{V}}$.

The final term of (A.6) is developed as

$$\begin{aligned} E\left(e^{\frac{\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}^2}v}I\right) &= \int_{v<-\mu_{\mathcal{V}}} e^{\frac{\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}^2}v}\left[\frac{1}{\sqrt{2\pi}\sigma_{\mathcal{V}}}e^{-\frac{v^2}{2\sigma_{\mathcal{V}}^2}}\right]dv \\ &= \int_{v<-\mu_{\mathcal{V}}} \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{V}}}e^{\left[-\frac{v^2}{2\sigma_{\mathcal{V}}^2}+\frac{\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}^2}v\right]}dv \end{aligned}$$

This term is simplified by completing the square in the exponent of the latter integral. The exponent is

$$\begin{aligned} -\frac{v^2}{2\sigma_{\mathcal{V}}^2} + \frac{\sigma_{\mathcal{U}\mathcal{V}}}{\sigma_{\mathcal{V}}^2}v &= -\frac{1}{2\sigma_{\mathcal{V}}^2}\left[v^2 - 2\sigma_{\mathcal{U}\mathcal{V}}v\right] \\ &= -\frac{1}{2\sigma_{\mathcal{V}}^2}\left[v - \sigma_{\mathcal{U}\mathcal{V}}\right]^2 + \frac{\sigma_{\mathcal{U}\mathcal{V}}^2}{2\sigma_{\mathcal{V}}^2} \end{aligned}$$

[30]Recall that our notational convention is that $E(\cdot|I)$ denotes expectation conditional on $I = 1$.

This implies that

$$\begin{aligned} E\left(e^{\frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v}I\right) &= e^{\frac{\sigma_{\mathcal{UV}}^2}{2\sigma_{\mathcal{V}}^2}}\int_{v<-\mu_{\mathcal{V}}}\left[\frac{1}{\sqrt{2\pi}\sigma_{\mathcal{V}}}e^{-\frac{1}{2\sigma_{\mathcal{V}}^2}[v-\sigma_{\mathcal{UV}}]^2}\right]dv \\ &= e^{\frac{\sigma_{\mathcal{UV}}^2}{2\sigma_{\mathcal{V}}^2}}\int_{v^*<-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}\left[\frac{1}{\sqrt{2\pi}\sigma_{\mathcal{V}}}e^{-\frac{1}{2\sigma_{\mathcal{V}}^2}(v^*)^2}\right]dv^* \\ &= e^{\frac{\sigma_{\mathcal{UV}}^2}{2\sigma_{\mathcal{V}}^2}}\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right] \end{aligned}$$

Collecting all of the terms gives

$$\begin{aligned} E[\mathcal{W}I] &= e^{\mu_{\mathcal{U}}}E(e^s)E\left(e^{\frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v}I\right) \\ &= e^{\mu_{\mathcal{U}}}e^{\frac{1}{2}\sigma_{\mathcal{U}}^2\left(1-\rho_{\mathcal{UV}}^2\right)}e^{\frac{\sigma_{\mathcal{UV}}^2}{2\sigma_{\mathcal{V}}^2}}\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right] \\ &= e^{\mu_{\mathcal{U}}+\frac{1}{2}\sigma_{\mathcal{U}}^2}\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right] \end{aligned}$$

Dividing by the formula for $E[\mathcal{W}I]$ by $E[I]$ gives the result for $E[\mathcal{W}|I]$, or (2.8). Noting that the mean of $\mathcal{W}$ in the population (without selection) is $E[\mathcal{W}] = e^{\mu_{\mathcal{U}}+\frac{1}{2}\sigma_{\mathcal{U}}^2}$ yields the remark following the statement of Lemma 2.1.

For part B, using (A.1) and (A.2), we have that

$$\mathcal{VW} = \mu_{\mathcal{V}}\mathcal{W} + e^{\mu_{\mathcal{U}}}e^s\left(v\cdot e^{\frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v}\right)$$

so that

$$E[\mathcal{VW}I] = \mu_{\mathcal{V}}E[\mathcal{W}I] + e^{\mu_{\mathcal{U}}}E(e^s)E\left(v\cdot e^{\frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v}I\right)$$

The first term can be solved for from part A, so we focus on the second term. We have

$$E\left(v\cdot e^{\frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v}I\right) = \int_{v<-\mu_{\mathcal{V}}} v\cdot e^{\frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v}\left[\frac{1}{\sqrt{2\pi}\sigma_{\mathcal{V}}}e^{-\frac{v^2}{2\sigma_{\mathcal{V}}^2}}\right]dv$$

$$= \int_{v<-\mu_{\mathcal{V}}} \frac{v}{\sqrt{2\pi}\sigma_{\mathcal{V}}} e^{\left[-\frac{v^2}{2\sigma_{\mathcal{V}}^2}+\frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v\right]} dv$$

$$= e^{\frac{\sigma_{\mathcal{UV}}^2}{2\sigma_{\mathcal{V}}^2}} \int_{v^*<-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}} (v^* + \sigma_{\mathcal{UV}}) \left[\frac{1}{\sqrt{2\pi}\sigma_{\mathcal{V}}} e^{-\frac{1}{2\sigma_{\mathcal{V}}^2}(v^*)^2}\right] dv^*$$

$$= e^{\frac{\sigma_{\mathcal{UV}}^2}{2\sigma_{\mathcal{V}}^2}} \left\{\int_{v^*<-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}} v^* \left[\frac{1}{\sqrt{2\pi}\sigma_{\mathcal{V}}} e^{-\frac{1}{2\sigma_{\mathcal{V}}^2}(v^*)^2}\right] dv^* + \sigma_{\mathcal{UV}} \Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right]\right\}$$

$$= e^{\frac{\sigma_{\mathcal{UV}}^2}{2\sigma_{\mathcal{V}}^2}} \left\{-\sigma_{\mathcal{V}}\phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right] + \sigma_{\mathcal{UV}}\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right]\right\}$$

where the third equality follows from completing the square as in part A, and the last equality follows from direct integration as in (A.3) above. Now, collecting terms gives

$$\begin{aligned}
E[\mathcal{VW}I] &= \mu_{\mathcal{V}} E[\mathcal{W}I] + e^{\mu_{\mathcal{U}}} E(e^s) E\left(v \cdot e^{\frac{\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}^2}v} I\right) \\
&= \mu_{\mathcal{V}} e^{\mu_{\mathcal{U}}+\frac{1}{2}\sigma_{\mathcal{U}}^2} \Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right] \\
&\quad + e^{\mu_{\mathcal{U}}+\frac{1}{2}\sigma_{\mathcal{U}}^2}\left\{-\sigma_{\mathcal{V}}\phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right] + \sigma_{\mathcal{UV}}\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right]\right\} \\
&= e^{\mu_{\mathcal{U}}+\frac{1}{2}\sigma_{\mathcal{U}}^2}\left\{(\mu_{\mathcal{V}}+\sigma_{\mathcal{UV}})\Phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right] - \sigma_{\mathcal{V}}\phi\left[\frac{-\mu_{\mathcal{V}}-\sigma_{\mathcal{UV}}}{\sigma_{\mathcal{V}}}\right]\right\}
\end{aligned}$$

Equation (2.9) follows from dividing by $E[I]$. This completes the proof of the Lemma 2.1.

Table 3.1: Proportions of single and married in FES data by year, whole sample

Year	Single Number	Single %	married number	married %	Total
1979	978	23.6	3166	76.4	4144
1984	1110	27.2	2974	72.8	4084
1985	1138	27.8	2954	72.2	4092
1986	1279	31.0	2852	69.0	4131
1987	1210	29.3	2922	70.7	4132
1988	1232	30.8	2765	69.2	3997
1989	1247	30.8	2801	69.2	4048
1990	994	27.4	2640	72.6	3634
1991	1080	28.7	2679	71.3	3759
1992	1181	29.8	2785	70.2	3966
1993	1136	30.4	2599	69.6	3735
1994	1040	29.1	2532	70.9	3572
1995	1012	28.7	2519	71.3	3531
Total	**14637**	**28.8**	**36188**	**71.2**	**50825**

Table 3.2: Proportions of single and married in FES data by education group, whole sample

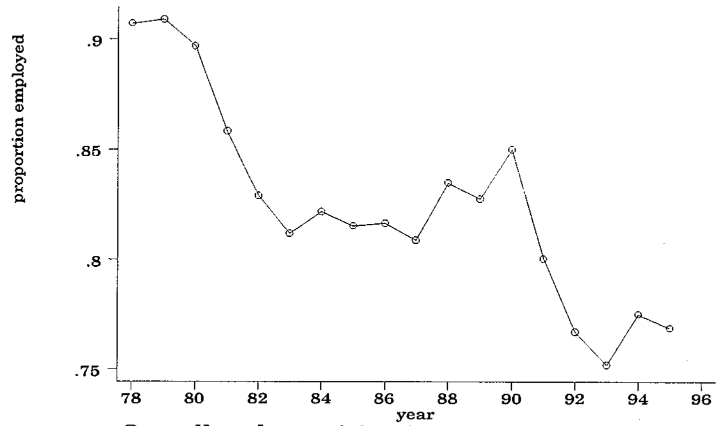
	Education group (i) left school at <=16	(ii) left 17-18	(iii) left 19+	TOTAL
Single	9889	2459	2289	14637
(%)	67.6	16.8	15.6	100
Married	26279	4963	4946	30270
(%)	72.6	13.7	13.7	100
TOTAL	36168	7422	7235	50825
(%)	71.2	14.6	14.2	100

Table 3.3. Significance tests for regression specification

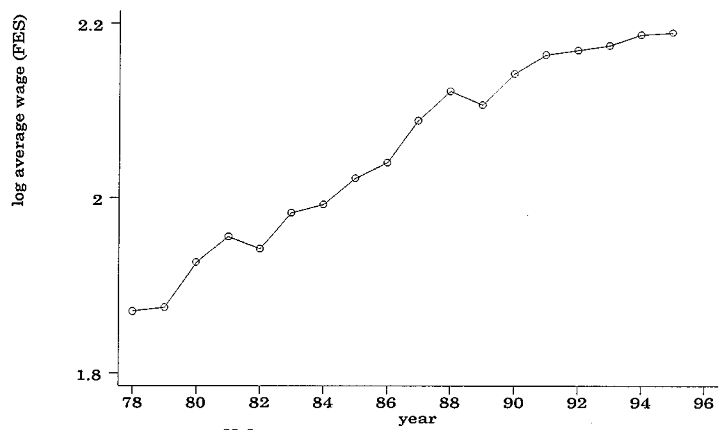
(significant results are shaded).

Coefficients	Participation equation		Wage equation	
	χ^2 (d.o.f.)	P - value	F-test (k) n=41209	P - value
Instruments (out-of-work income * marial status)	2514.5 (3)	.000	N/A	
Education (left 17-18, left 19+)	54.76 (2)	.000	167.76 (2)	.000
Trend (3rd order polynomial)	26.49 (3)	.000	27.61 (3)	.000
Cohort (b. 1919-34, b. 1935-44, b. 1955-64, b. 1965-77)	99.91 (4)	.000	70.96 (4)	.000
Education * trend	4.35 (6)	.630	10.71 (6)	.000
Education * cohort	14.02 (8)	.081	25.72 (8)	.000
Trend * cohort	189.91 (12)	.000	16.70 (12)	.000
Education * trend (1st order only) * cohort	11.46 (8)	.17	5.93 (8)	.000
Region (11 standard regions)	219.17 (10)	.000	45.50 (10)	.000
Region*trend, region*trend^2	38.18 (20)	.004	3.00 (20)	.000
Mills ratio * education *marital status	N/A		20.18 (6)	.000
Married (single coefficient)	.517	.000	.287	.000
Spouse's education (single coefficient)	.031	.000	N/A	

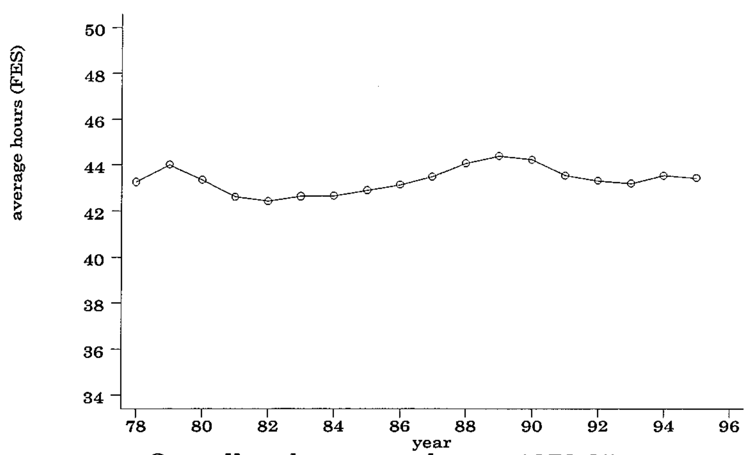
Figure 1.1: British males – participation, wages and hours



Overall male participation rate, 1978-95



overall log average wages, 1978-95



Overall male average hours, 1978-95

Figure 1.2: Participation, wages and hours by education group

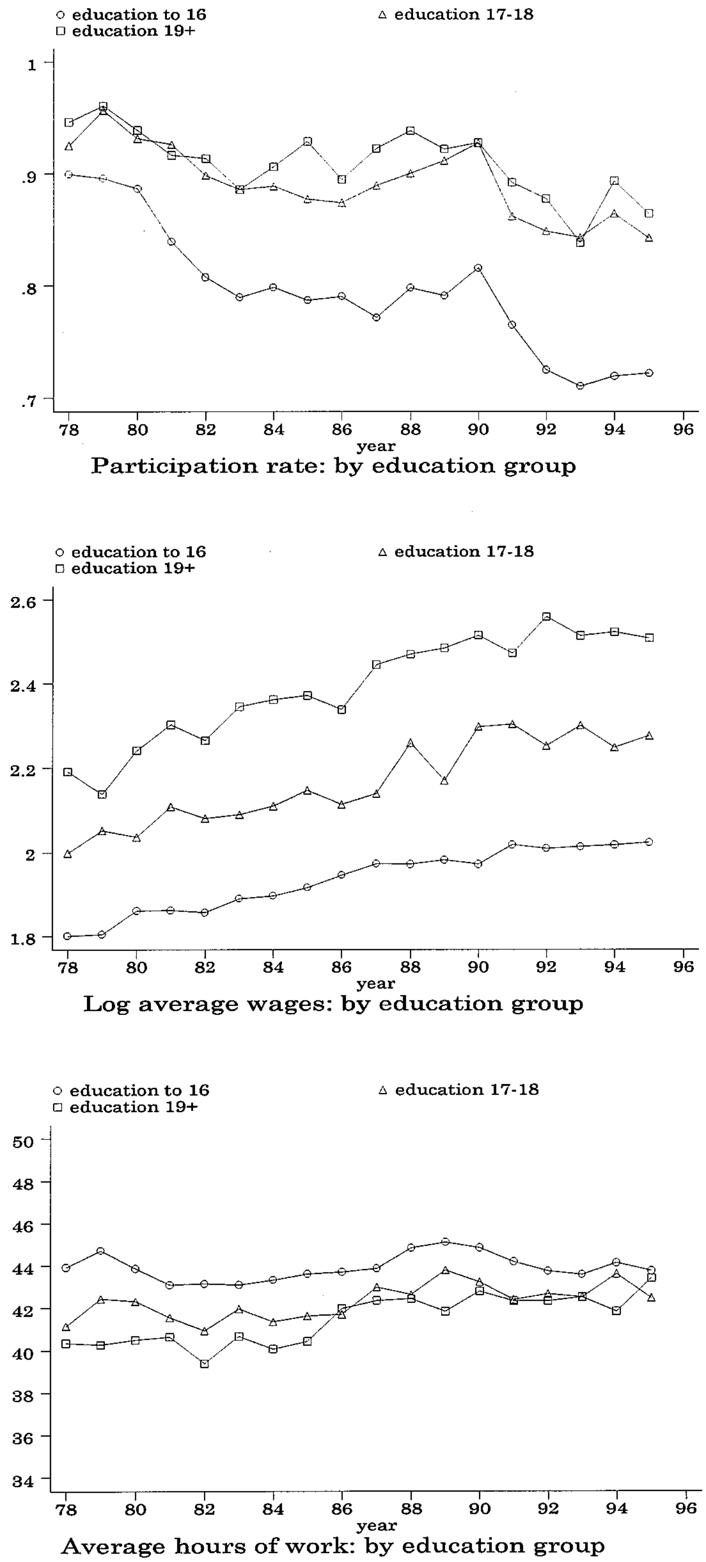
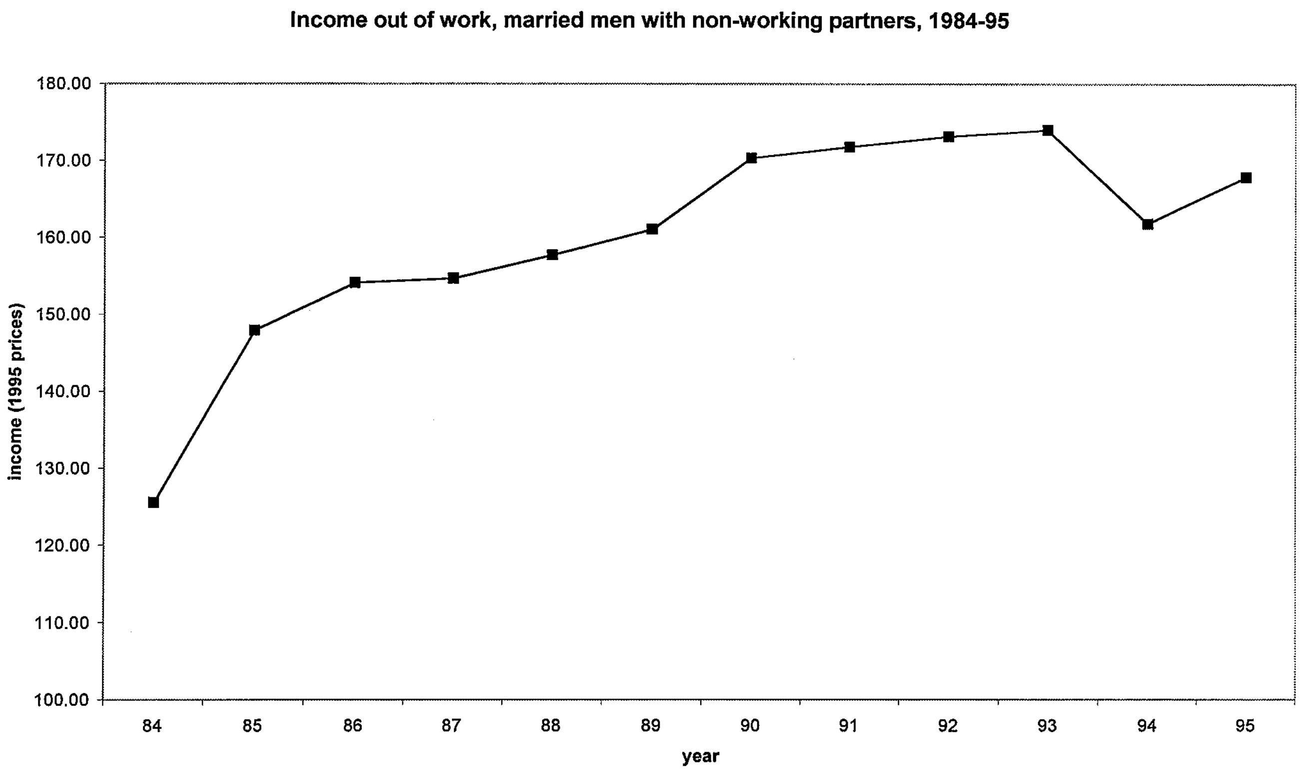
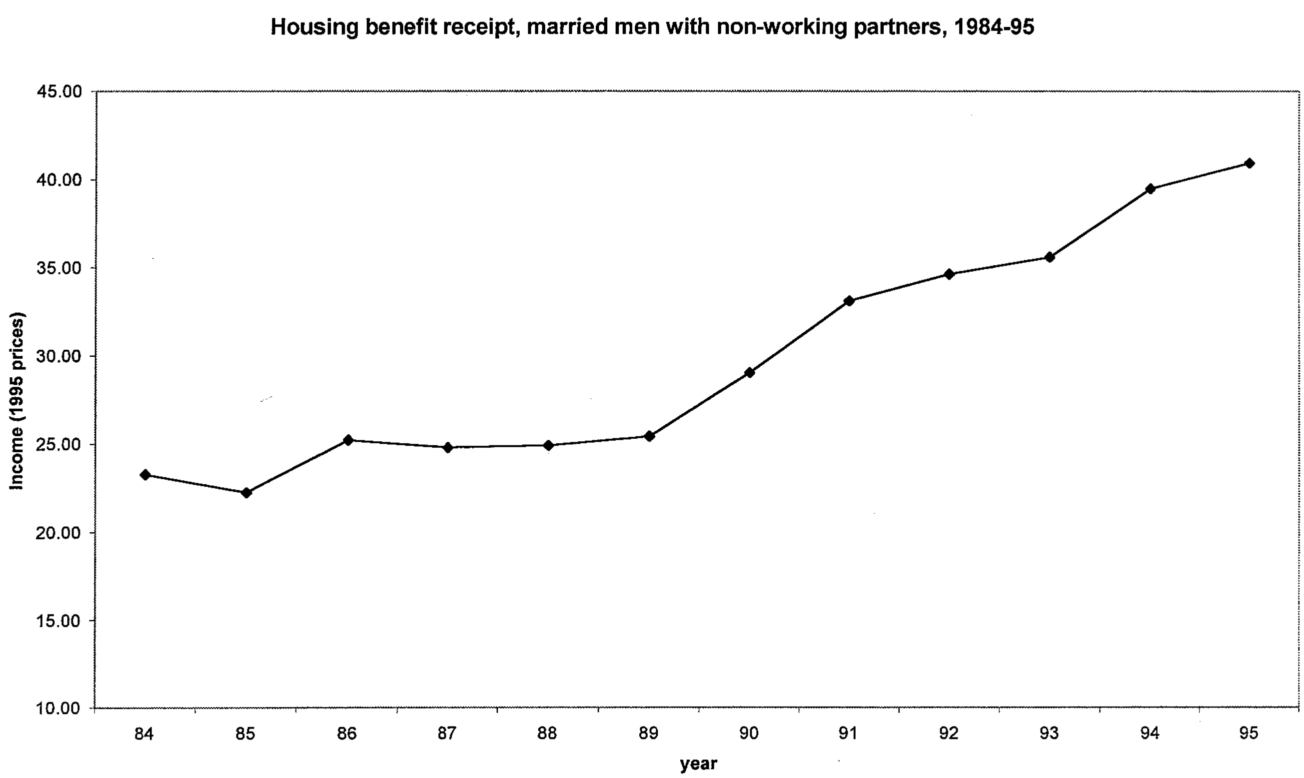


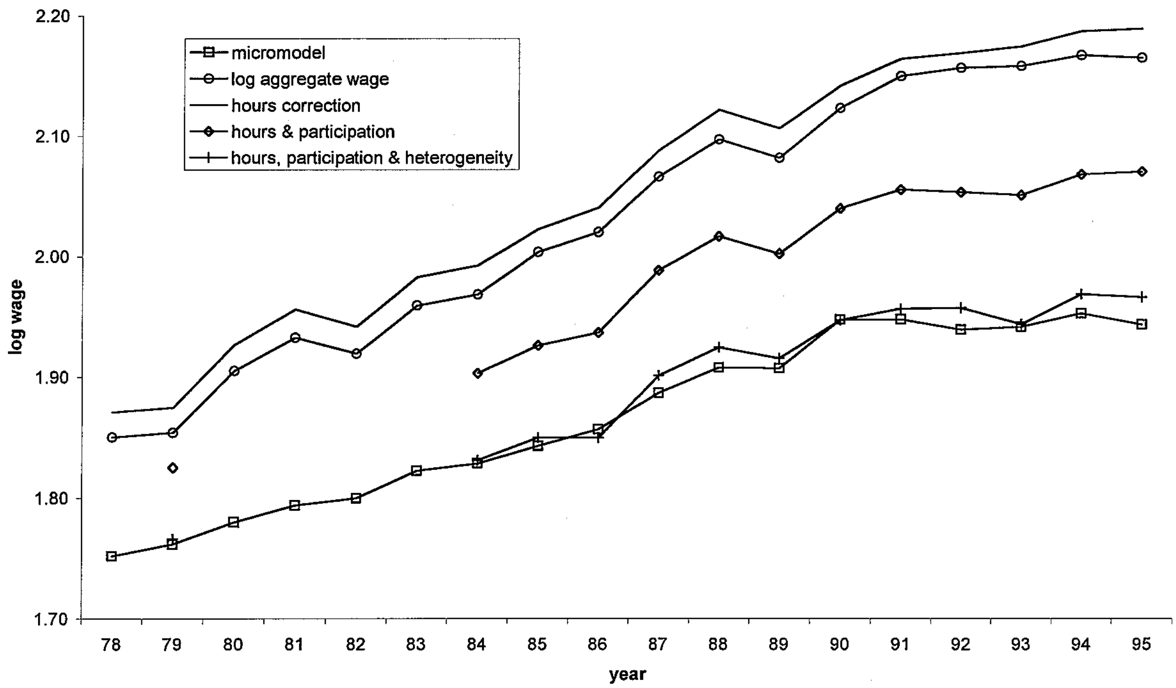
Figure 1.3: Simulated out-of-work income and benefit receipt





Source: IFS TAXBEN microsimulation model

Figure 3.1: wage predictions from micromodel, aggregate wage and corrections – whole sample



Rebased to 1979

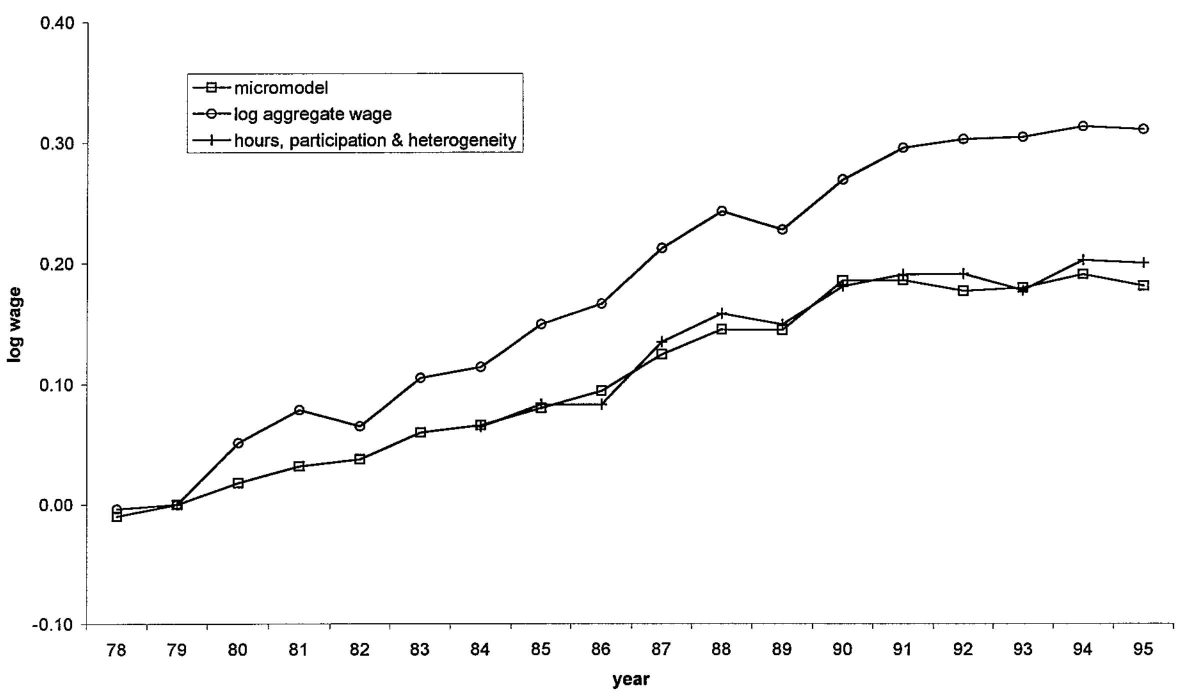
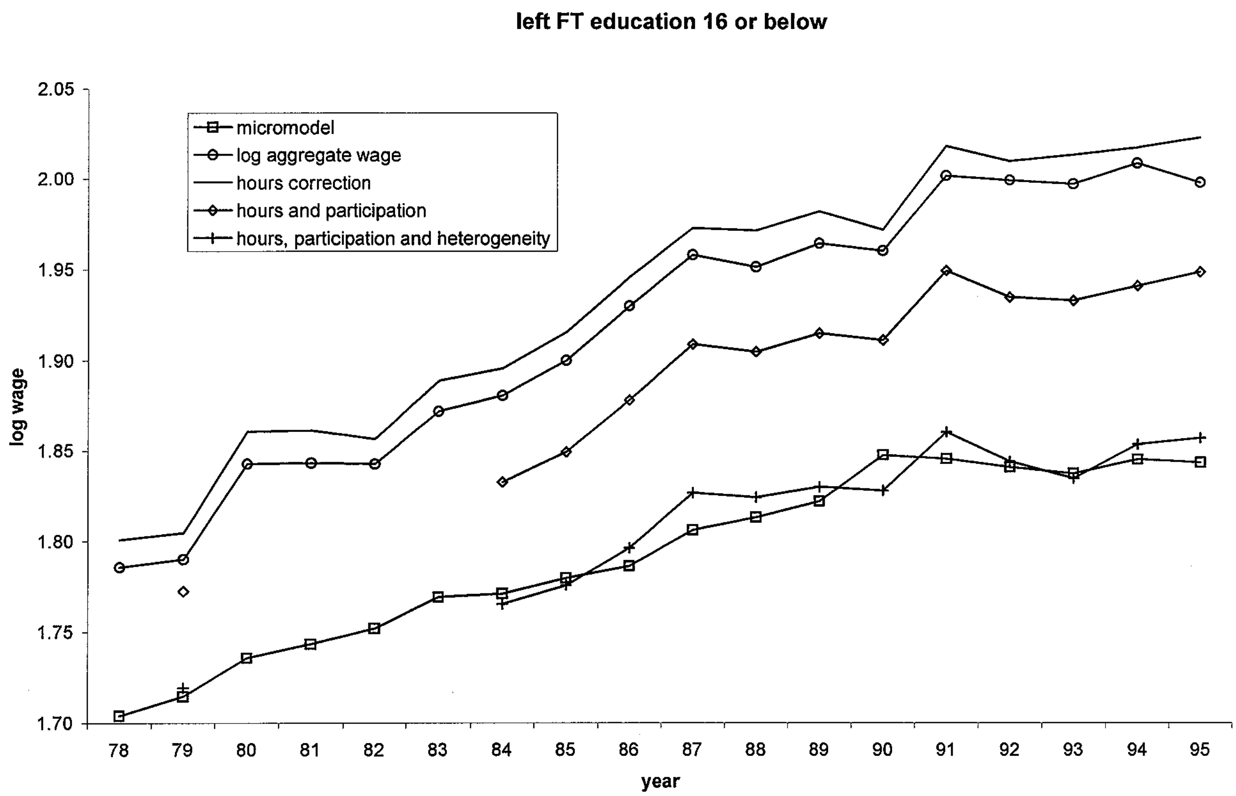
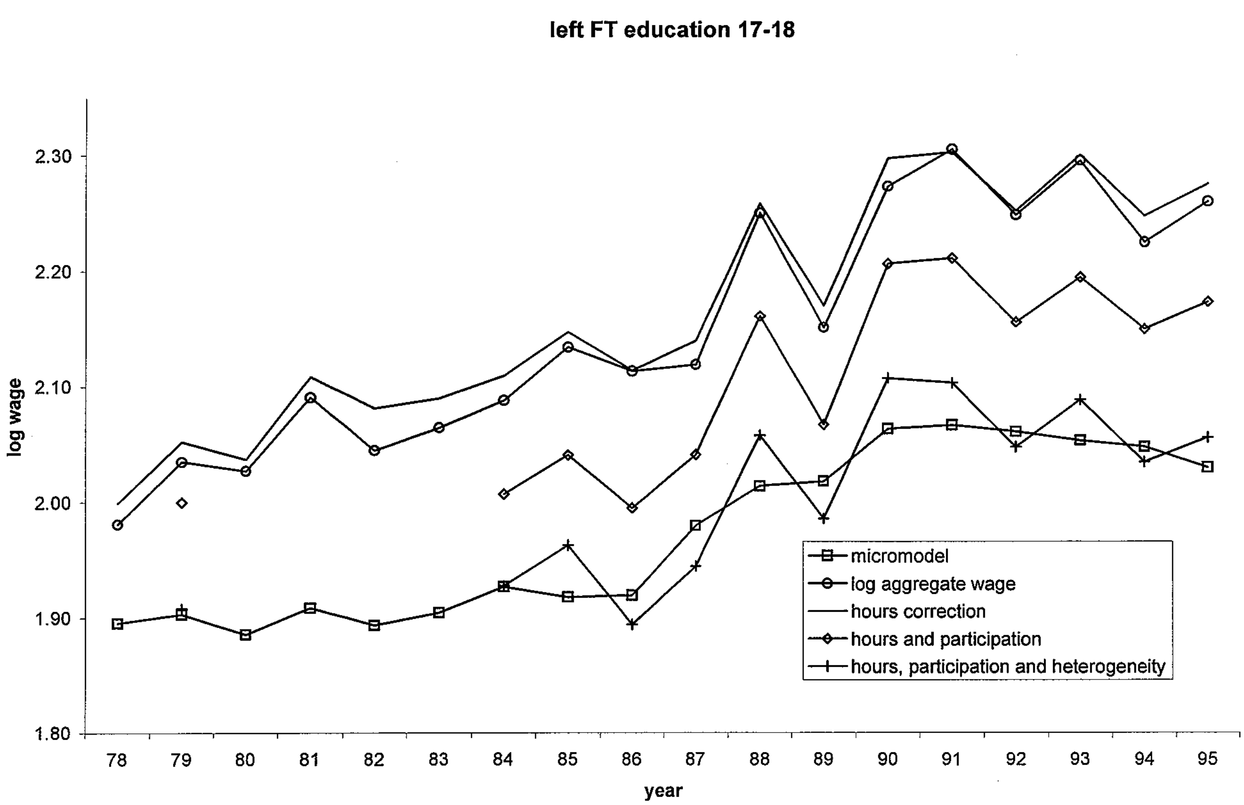


Figure 3.2: wage predictions and corrections by education group





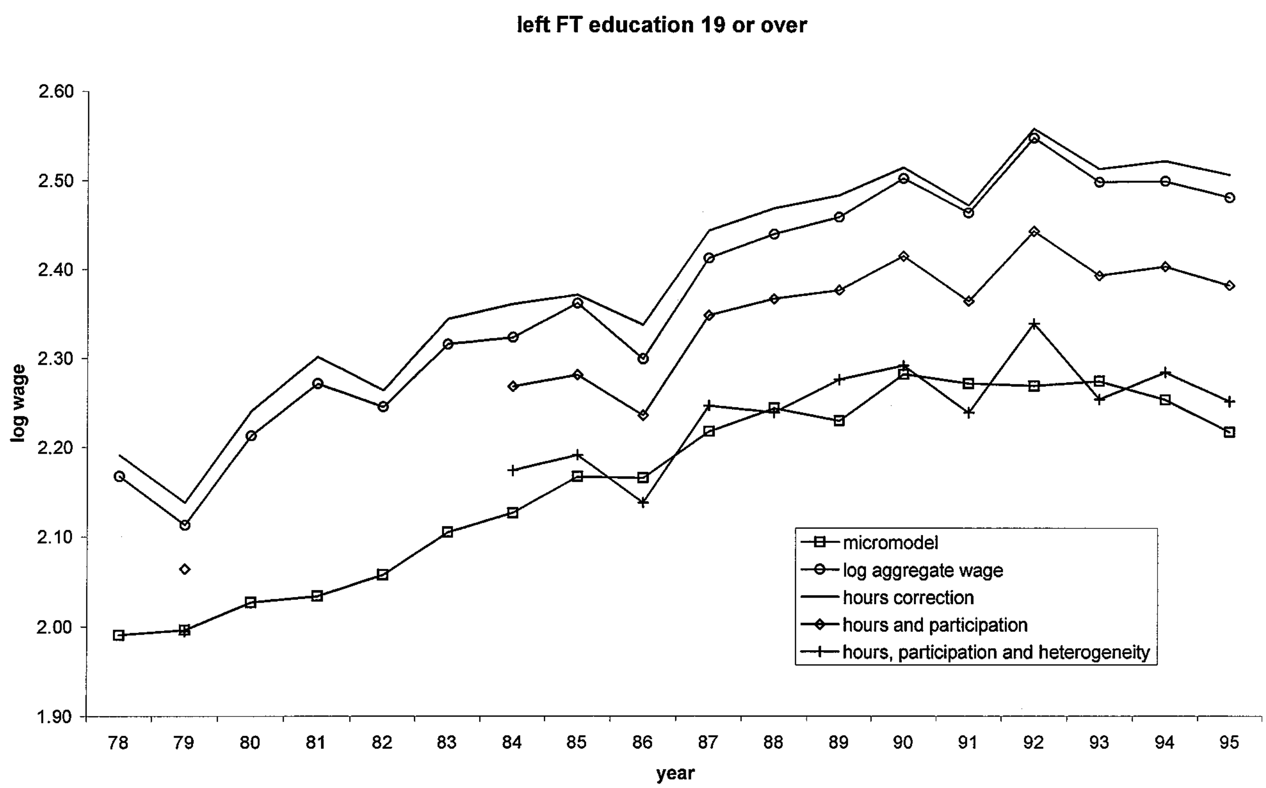
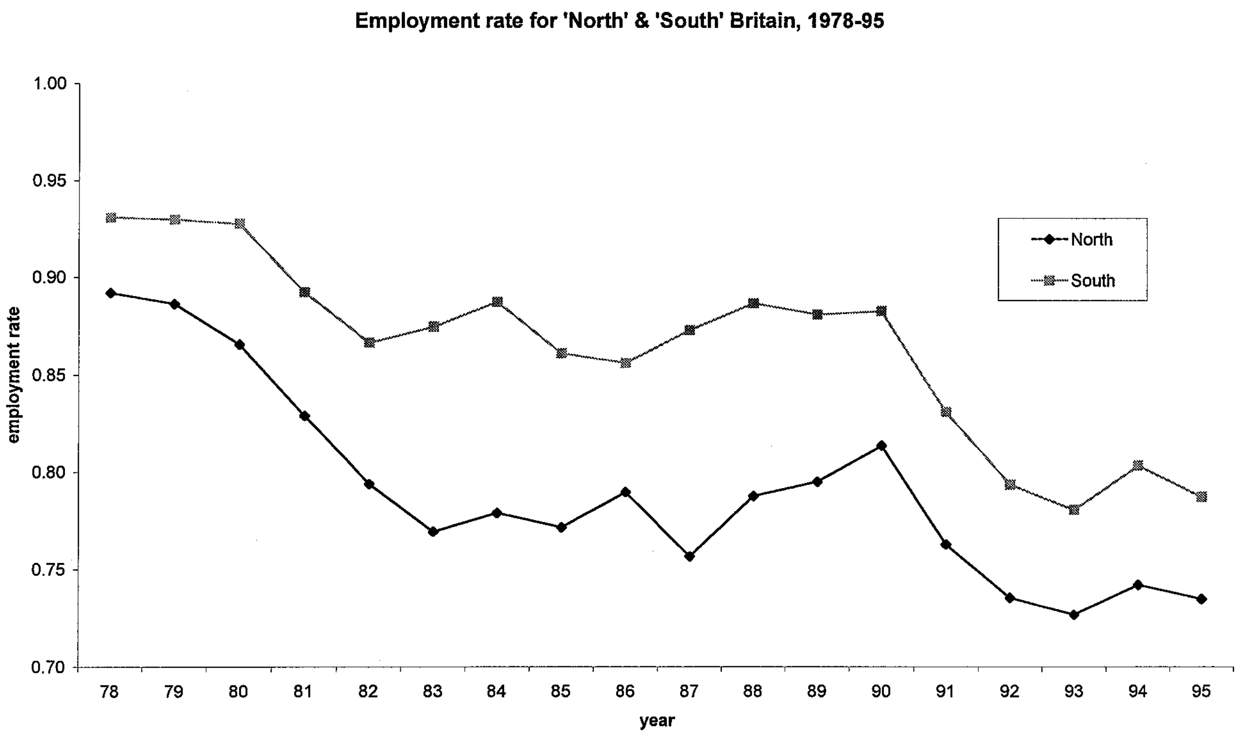

left FT education 19 or over
log wage
year
2.60
2.50
2.40
2.30
2.20
2.10
2.00
1.90
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
micromodel
log aggregate wage
hours correction
hours and participation
hours, participation and heterogeneity

Figure 3.3: participation and wages by broad region



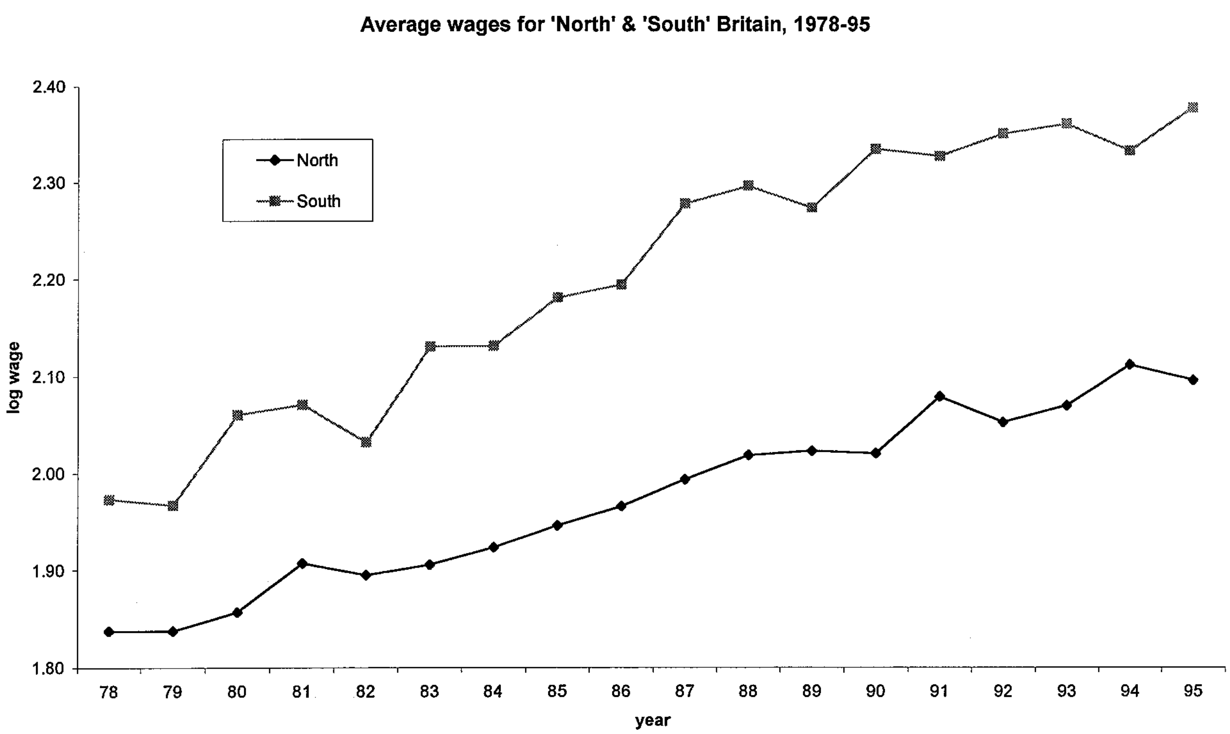
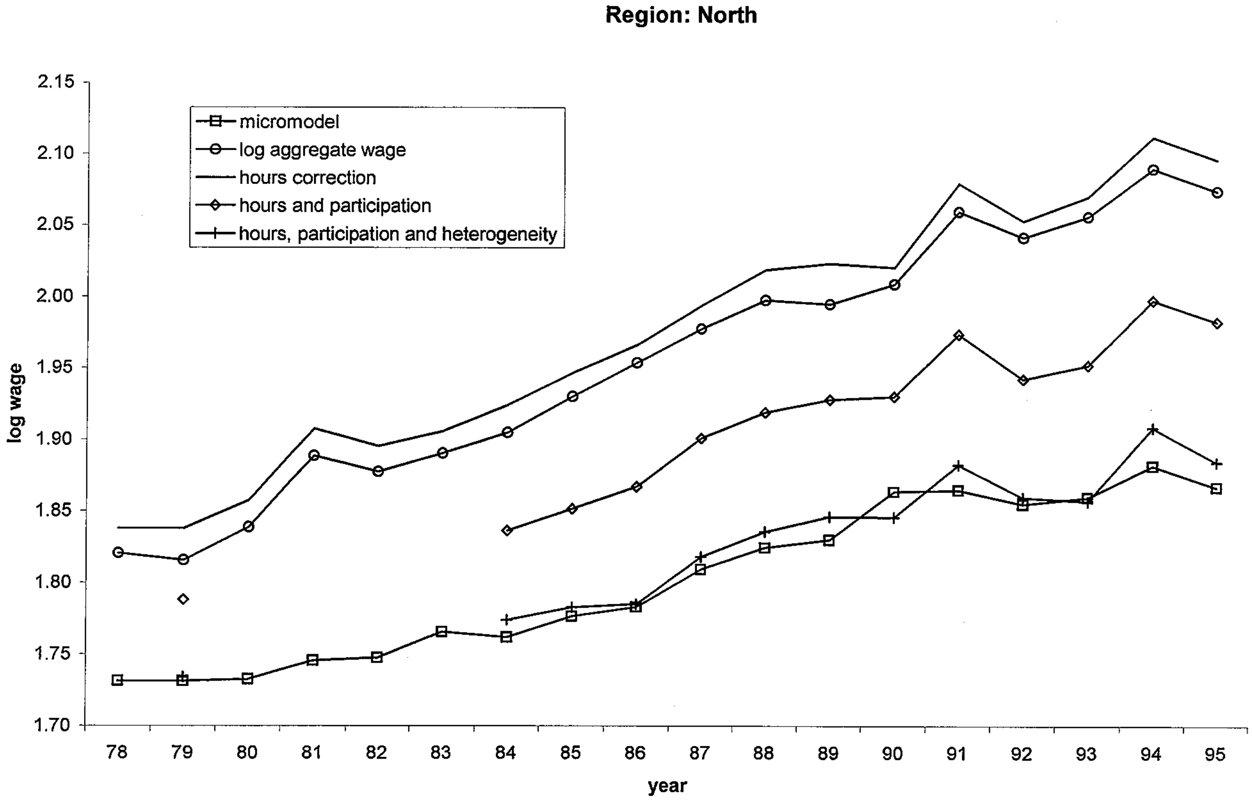


Figure 3.4: wage predictions and corrections by broad region



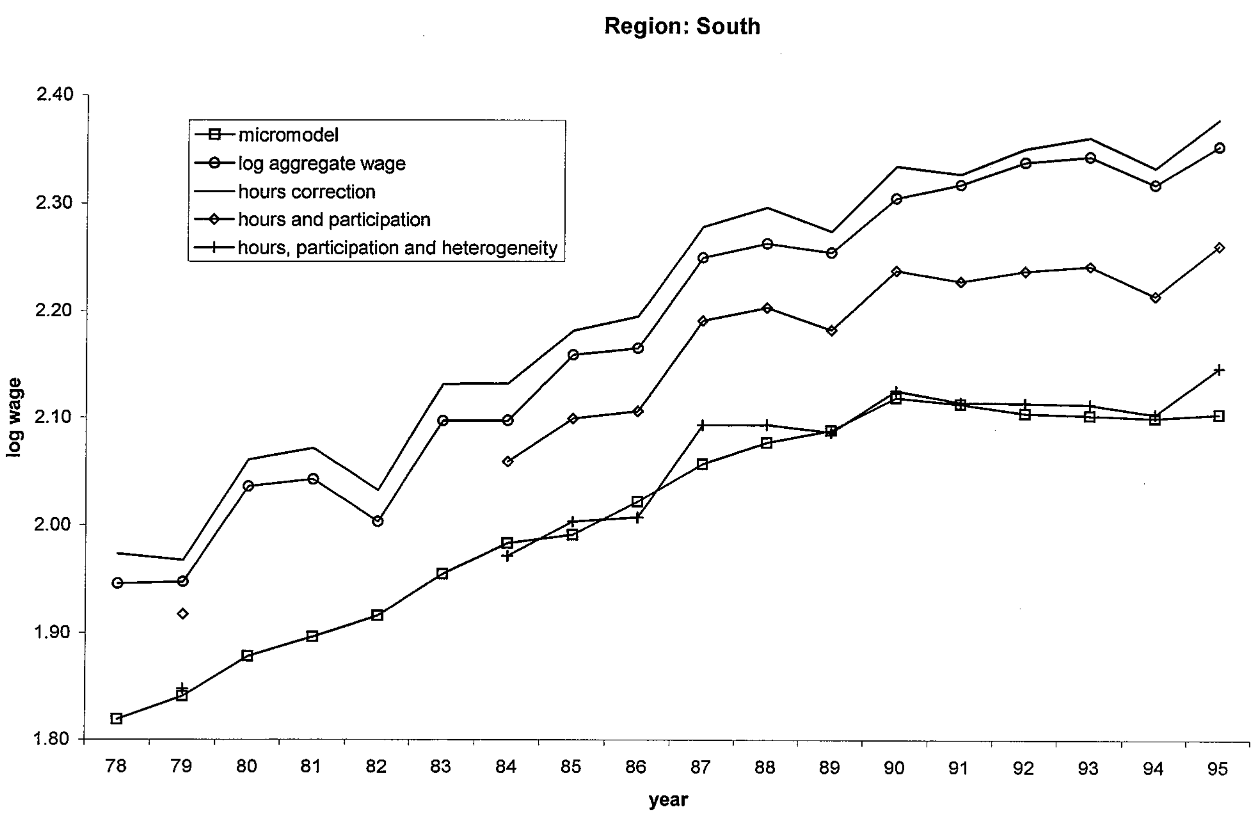
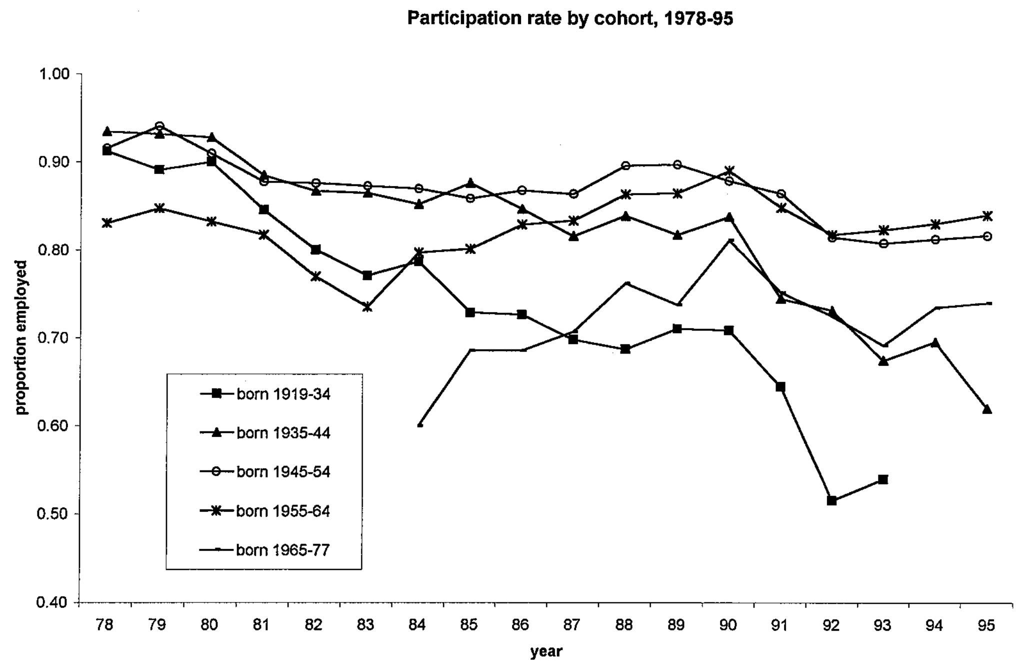


Figure 3.5: participation and wages by cohort group



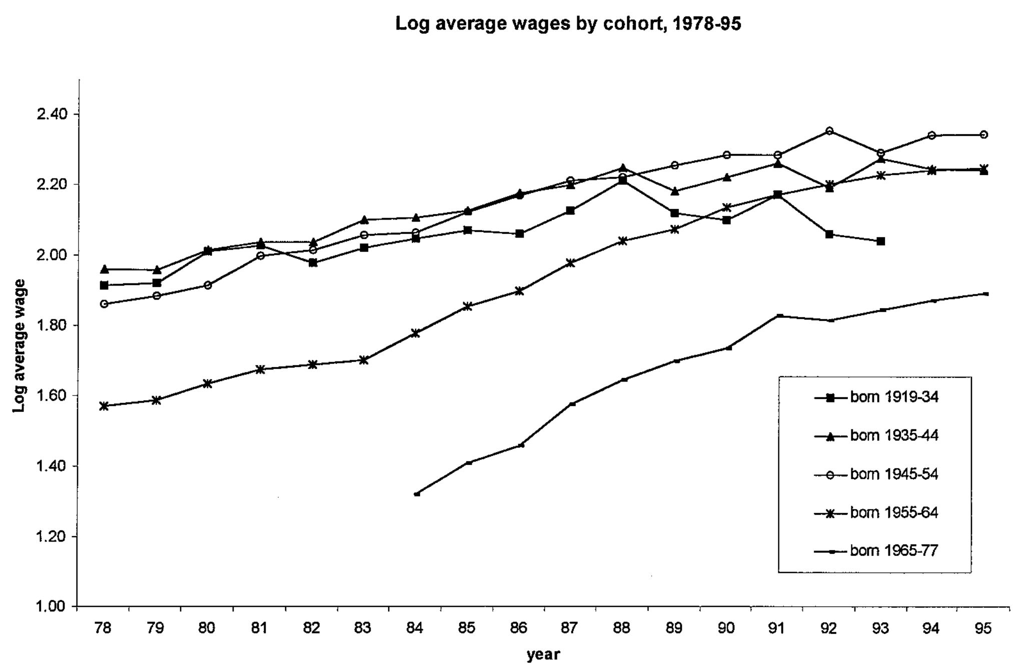
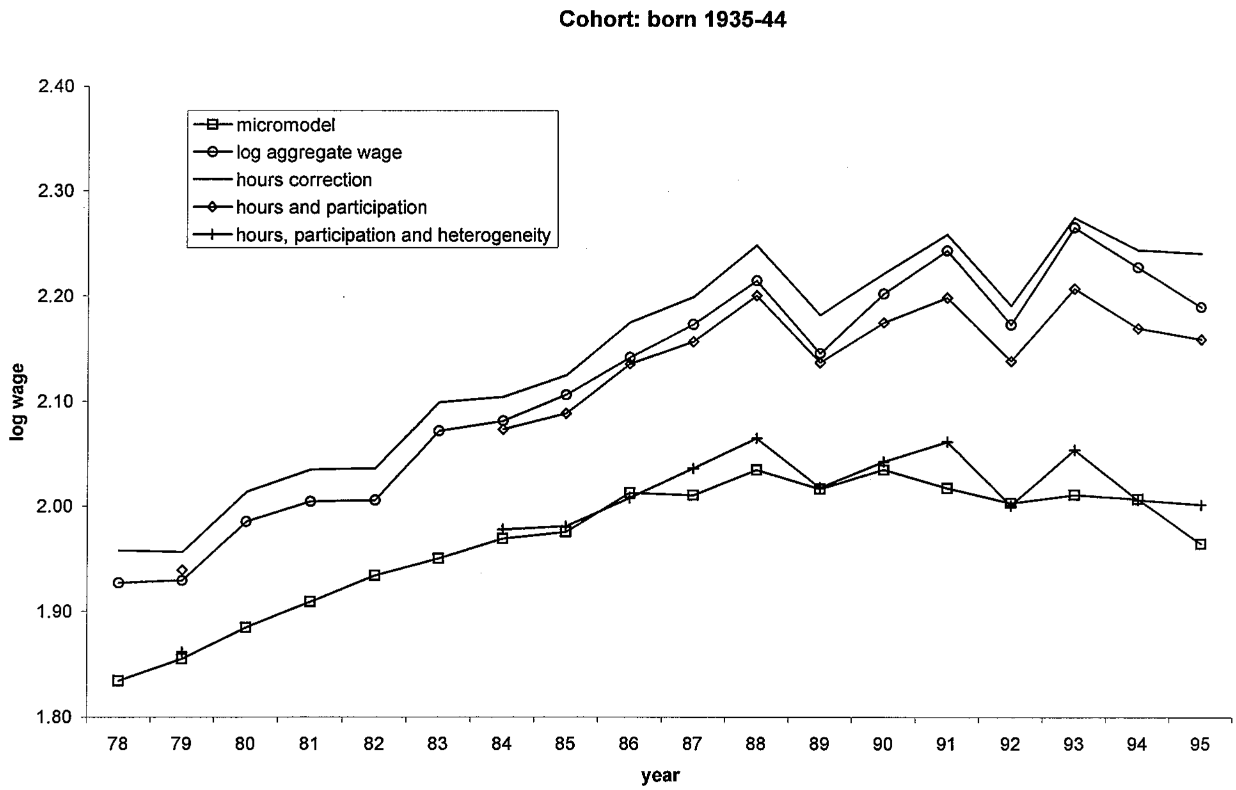


Figure 3.6: wage predictions and corrections by cohort group



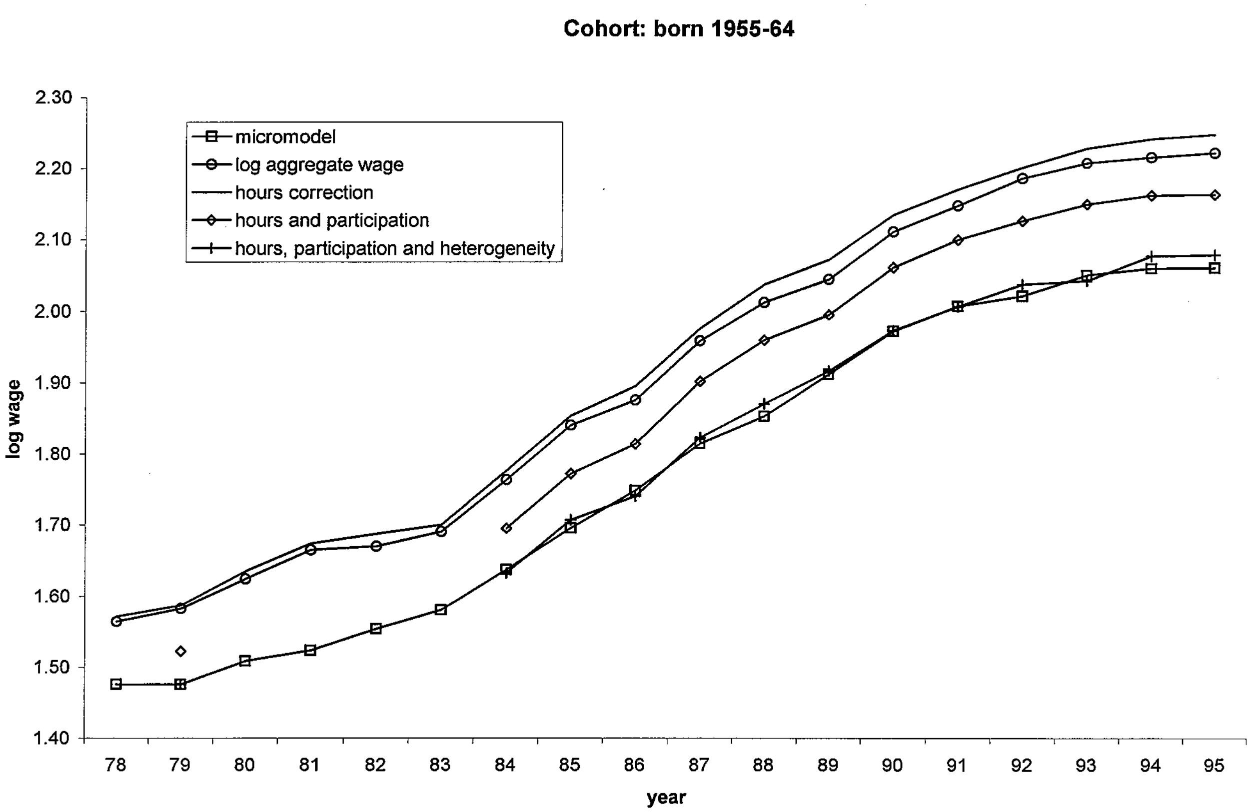
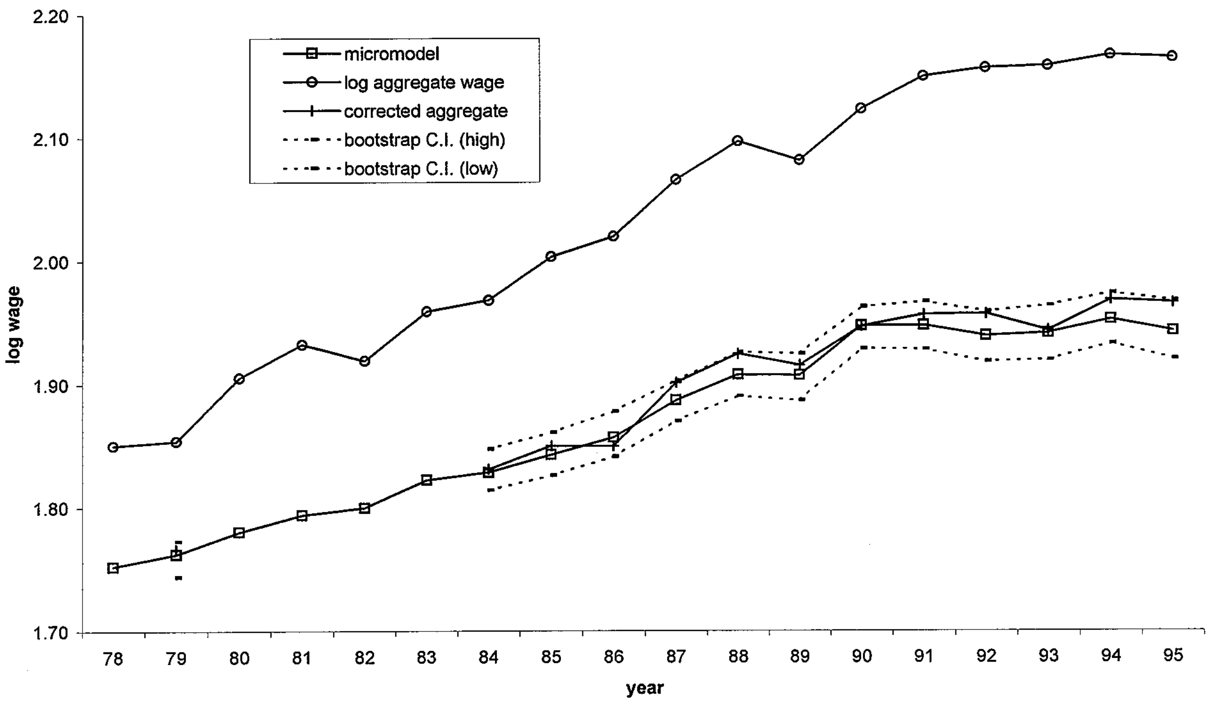
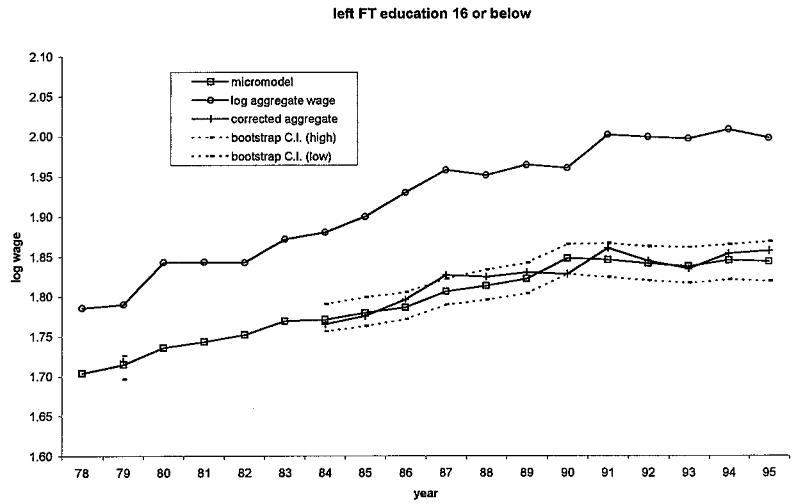


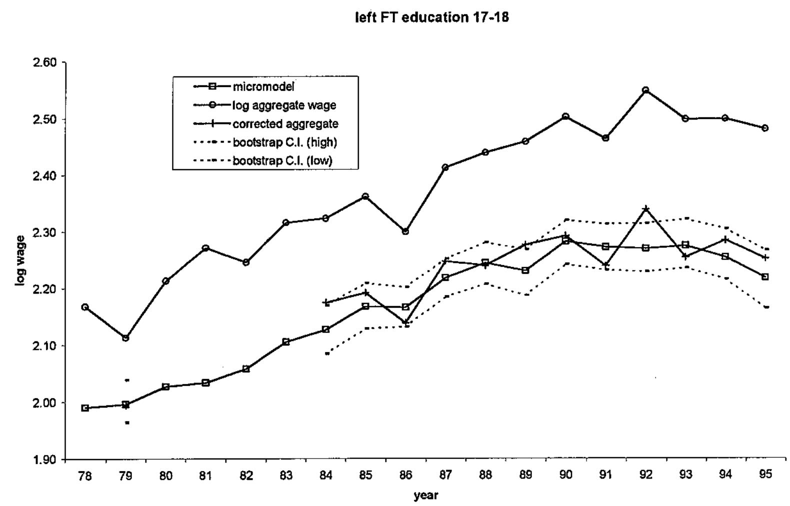
Figure 3.7: Plots of bootstrapped standard errors on micromodel predictions (95% confidence intervals)

a) overall sample



b) by education group





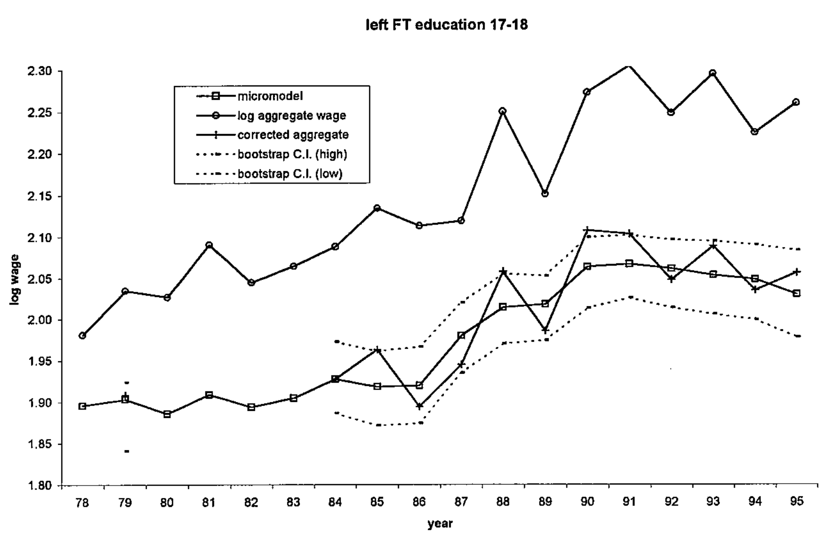
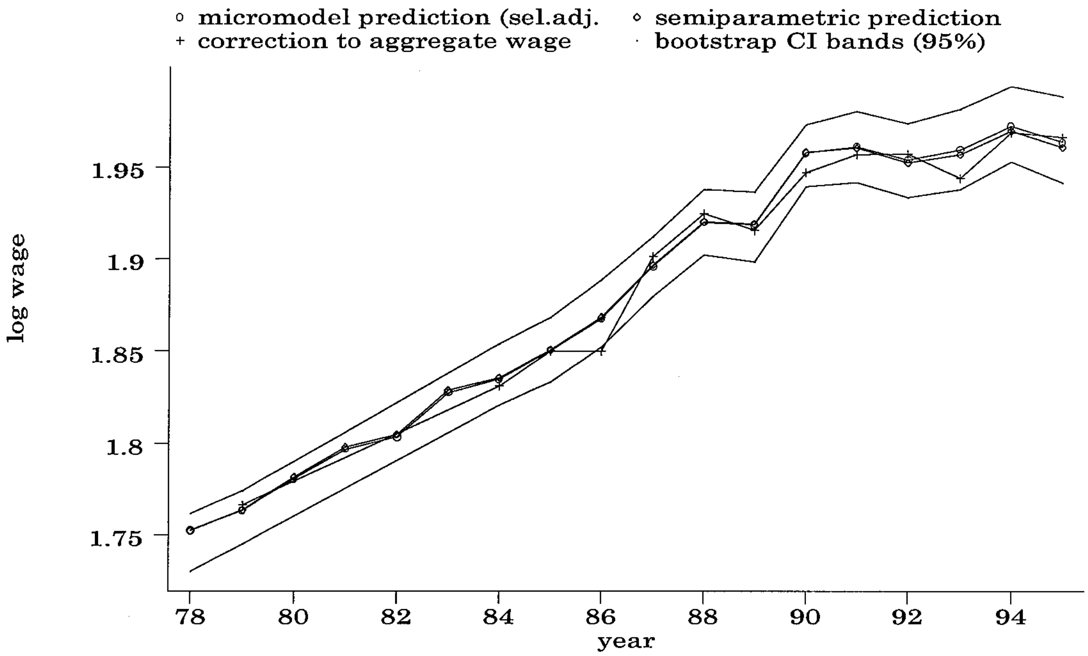


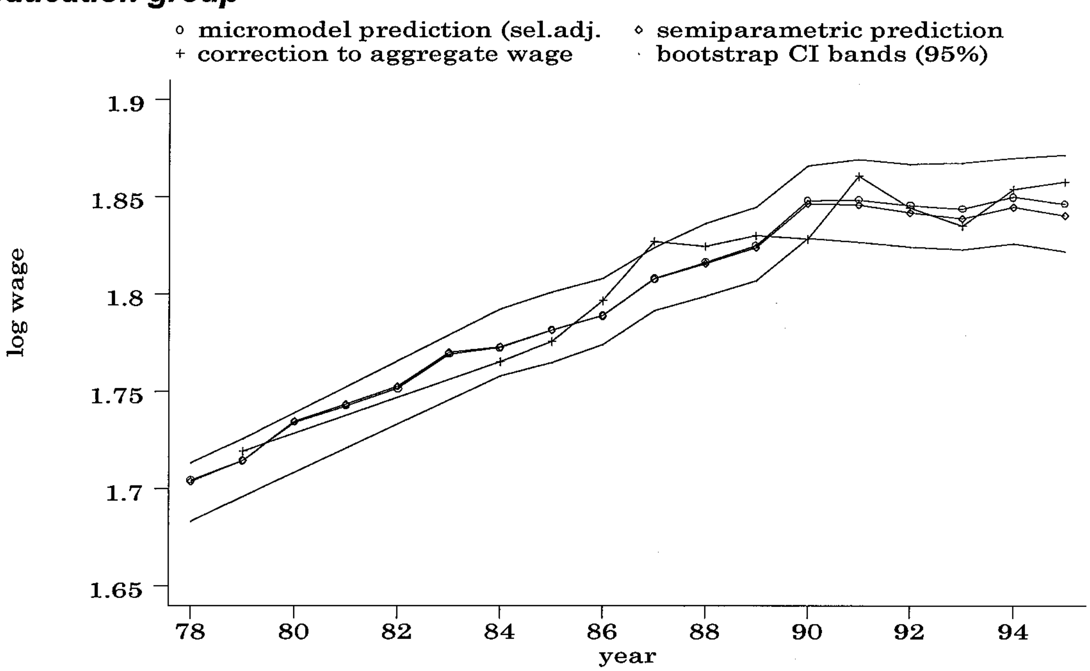
Figure 3.8: Semiparametric estimation results

a) overall



Semiparametric estimation, overall: plot comparison

b) low education group



Semiparametric estimation, low education group

Figure 3.9: Plot of z'α index from Probit, overall sample

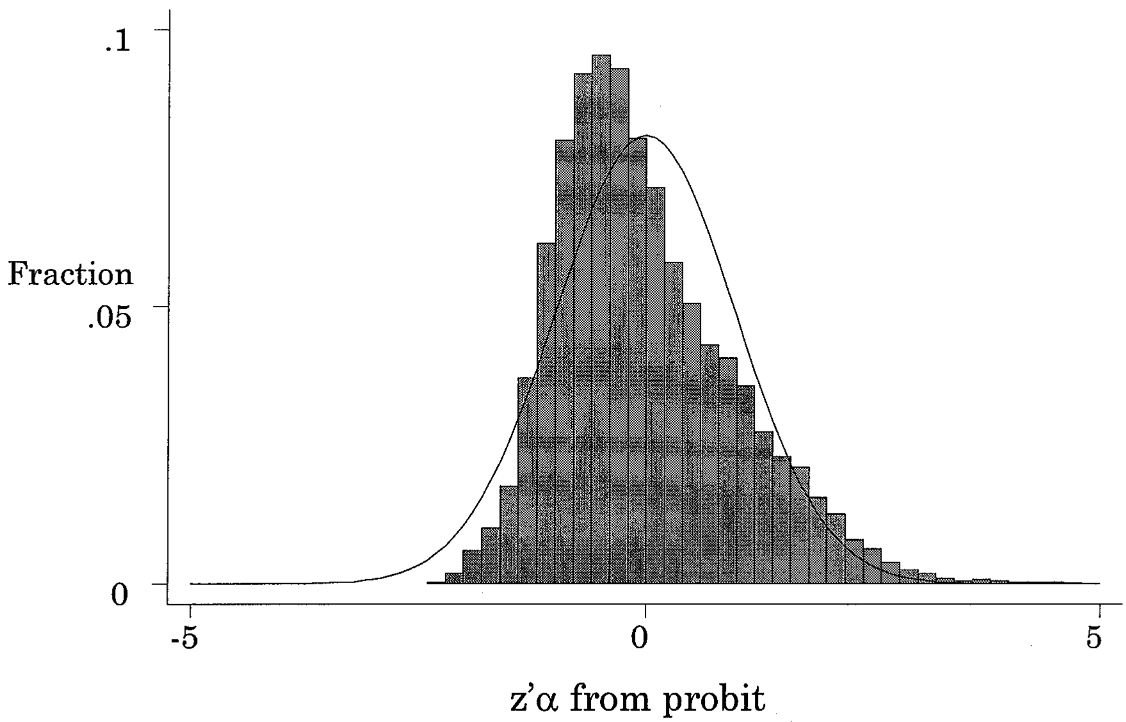


Figure 3.10: Plot of standardised predictions from wage equation, overall sample