Human Capital and Social Position in Britain: Creating a Measure of Wage-Earning Potential from BHPS Data

Man Yee Kan and Jonathan Gershuny

Institute for Social and Economic Research

University of Essex

ISER Working Paper
2006-03

Institute for Social and Economic Research

The Institute for Social and Economic Research (ISER) specialises in the production and analysis of longitudinal data. ISER incorporates the following centres:

- ESRC Research Centre on Micro-social Change. Established in 1989 to identify, explain, model and forecast social change in Britain at the individual and household level, the Centre specialises in research using longitudinal data.

- ESRC UK Longitudinal Studies Centre. A national resource centre for promoting longitudinal research and for the design, management and support of longitudinal surveys. It was established by the ESRC as independent centre in 1999. It has responsibility for the British Household Panel Survey (BHPS).

- European Centre for Analysis in the Social Sciences. ECASS is an interdisciplinary research centre which hosts major research programmes and helps researchers from the EU gain access to longitudinal data and cross-national datasets from all over Europe.

The British Household Panel Survey is one of the main instruments for measuring social change in Britain. The BHPS comprises a nationally representative sample of around 9,000 households and over 16,000 individuals who are reinterviewed each year. The questionnaire includes a constant core of items accompanied by a variable component in order to provide for the collection of initial conditions data and to allow for the subsequent inclusion of emerging research and policy concerns.

Among the main projects in ISER's research programme are: the labour market and the division of domestic responsibilities; changes in families and households; modelling households' labour force behaviour; wealth, well-being and socio-economic structure; resource distribution in the household; and modelling techniques and survey methodology.

BHPS data provide the academic community, policymakers and private sector with a unique national resource and allow for comparative research with similar studies in Europe, the United States and Canada.

BHPS data are available from the UK Data Archive at the University of Essex
http://www.data-archive.ac.uk

Further information about the BHPS and other longitudinal surveys can be obtained by telephoning +44 (0) 1206 873543.

The support of both the Economic and Social Research Council (ESRC) and the University of Essex is gratefully acknowledged. The work reported in this paper is part of the scientific programme of the Institute for Social and Economic Research.

Acknowledgement:

This paper forms part of the scientific programme of the ESRC Research Centre on Micro-social Change, and also part of the project "Gender, Time Allocation and the 'Wage Gap'" funded by the ESRC Gender Equality Network.

Readers wishing to cite this document are asked to use the following form of words:

> Kan, Man Yee and Gershuny, Jonathan (February 2006) 'Human Capital and Social Position in Britain: Creating a Measure of Wage Earning Potential from BHPS Data', ISER Working Paper 2006-03. Colchester: University of Essex.

For an on-line version of this working paper and others in the series, please visit the Institute's website at: http://www.iser.essex.ac.uk/pubs/workpaps/

Institute for Social and Economic Research
University of Essex
Wivenhoe Park
Colchester
Essex
CO4 3SQ UK
Telephone: +44 (0) 1206 872957
Fax: +44 (0) 1206 873151
E-mail: iser@essex.ac.uk
Website: http://www.iser.essex.ac.uk

ABSTRACT

This paper develops a continuously scaled indicator of social position (the Essex Score), which is estimated as individuals' potential wage in the labour market. The Essex Score is designed as a tool to investigate patterns of differentiation in life chances. It is constructed based on individuals' educational qualifications, recent experience in employment and non-employment, and occupational attainment using data from all the currently available 13 waves of the British Household Panel Survey. The Essex Score represents those embodied economic resources salient to individuals' participation in the labour market, equivalent to "human capital" in economic literature, and sometimes indicated by social class categories in sociological research. It has advantages over other social class measures. Being based on educational levels and on degrees of present and past attachment to the labour market as well as on present or previous occupational membership, it covers the entire adult population irrespective of their employment status and employment history. Its continuous level measurement also allows aggregation of scores from an individual to a household level, as well as the sensitive investigation of the determinants and consequences of changes in social position during the life course.

Human Capital and Social Position in Britain: Creating a Measure of Wage Earning Potential from BHPS data

INTRODUCTION

This paper reports an exercise continuing Gershuny's (2002) work on developing a continuously scaled indicator of social position, estimated as individuals' wage earning potential in the labour market (the "Essex Score"). The Essex Score is designed as a tool to investigate the distribution of economic power within society. It is also aimed to be a non-categorical index of personal embodied resources and skills salient to labour-market performance, i.e., what economists conventionally label as "human capital". Gershuny (2002) has used the first 9 waves of the British Household Panel Survey (BHPS) data to develop the scale. This exercise extends his work by using all the currently available 13 waves of BHPS data and modifying the estimating procedures.

MEASURING SOCIAL POSITION

By social position we mean a set of circumstances that gives an individual access to a particular range of future production and consumption activities. Neo-classical economists (e.g., Schultz, 1961; Becker, 1993) have differentiated between physical capital (which is embodied to productive tools and machinery) and human capital (which is embodied to individuals, such as skills, experience, knowledge and other personal attributes), and have advocated that both capitals can be invested or forgone and have direct impact on economic productivity and earnings. The theory of human capital has been applied to explain differentials in the economic advantages among different social groups (e.g., why is there a segregation of occupations based on gender? See, for example, Polachek, 1985) and individual or family decisions related to labour market work (e.g., why is there a gendered division of labour in the household? See, for example, Becker, 1991[1981]). Since attributes related to productivity and earnings are difficult to measure directly, economists often use educational attainment as a proxy measure of human capital. But such a proxy obviously misses out other important dimensions of human capital, e.g., specific knowledge and personal attributes that might not be reflected in educational attainment, and is based on the assumption that educational attainment is positively associated with performances in the labour market.

Sociologists, on the other hand, have conventionally measured social position by the concept of social class, which is commonly accepted to reside in an individual's possession of various sorts of resources (e.g., Scott, 1996), including fixed and financial capital (e.g., land, machinery and equipments) and capital embodied to an individual (e.g., personal skills and resources that determine an individual's position in the labour market). It is thought that individuals' capacities to achieve their own life aims, or their life chances, are crucially dependent on the mixture of fixed, financial and embodied capital resources that are at their disposal. As we can see, there are some similarities in how sociologists and economists define fixed and embodied capitals. One of the main distinctions between the economic and the sociological traditions perhaps lies in the fact that the former emphasizes much an individual's own choice about investment in human capital and hence social position, whereas the latter focuses on how societal constraints (e.g., social norms and family backgrounds) affect an individual's chances of acquiring certain forms of capital and hence attaining to a certain social position.

From modern sociological research, we can identify three distinct forms of embodied capital: "Human capital" refers to economically salient personal resources (e.g., skills, educational attainment, and specific knowledge) of the sort that might be considered by prospective employers as justifying offers of employment (Becker, 1993; Coleman, 1988); "social capital" refers to the range of and nature of personal connections between an individual and others in the society, which may be deployed for some further purposes beyond the immediate enjoyment of their company; "cultural capital" refers to specific knowledge related to the participation in, and enjoyment of, the various forms of consumption in the society. In sociological literature, the term "human capital" is sometimes used to cover all the above three sorts of capital. To avoid confusions and be consistent with the definition of the term used in economic literature, we define "human capital" in a more limited sense set out above as the capital that has direct economic salience in the labour market.

How can we measure individuals' embodied resources and hence their social positions? Sociologists have attempted to construct continuous or non-categorical measures of social position. In the U.K., most notably, these include the Hope-Goldthorpe Scale, which measures social prestige or standing on the basis of an individual's occupation (see

Goldthorpe & Hope, 1974), and the Cambridge Scale, which is constructed on the basis of an occupational ordering from accounts of the occupational affiliations of friends, on the assumption of affinities in the prestige levels in friendships (see Stewart, et al., 1980 and Prandy, 1990). These measures, however, rely heavily on the assertion of an association between prestige and economic power. Although the above studies did find a positive association between measures of occupational prestige and indicators of economic success, it is doubtful whether the process is a causal one, i.e., it is not clear whether economic power is really derived from prestige, or the association between them is to be explained mainly by some quite the reverse causal narratives so that those who are economically successful are as a result often also prestigious. Another disadvantage of these measures is that they are based (primarily at least) on an individual's present occupation so we cannot use them to study individuals who are currently not in employment. In such circumstances, these scales sometimes use *past* occupations for individuals who do not have a current job. But past and present occupations do not necessarily have the same implications for social positions. For example, it will obviously be problematic to assume that a nurse practitioner, who has been working continuously for 15 years, to have the same level of economic and social resources as a former colleague of her, who has left the occupation for 10 years since marrying a doctor. By contrast, the Essex Score approach includes an explicit process that adjusts for the absence of recent labour market experience.

Our approach to measure social position is to focus on the direct economic advantages that individuals' could potentially obtain through their participation in labour market work. We intend the Essex Score to address the real underlying issues of the core notion of social class—long term advantages and disadvantages, maintained through life-course and transmitted across generations, which result from differentials in possession of economically salient resources. We use the coefficients from a wage equation as the basis for our index of human capital. Insofar as human capital takes *direct* account of various economically salient personal resources, it is a more appropriate way of representing long term patterns of advantage and disadvantage that constitute social class than non-categorical scales that are constructed on the basis of present or past occupational attachment.

Certainly employers might seek a range of characteristics from potential employees, such as professional qualifications and educational attainment, commitment to work, diligence, reliability, and specific skills and work experience. Our task is to select a set of relevant personal characteristics that can be appropriately weighted and then combined to make up a single score. We are interested particularly in how the duration of employment or absence from employment will affect an individual's earnings potential in the labour market. We take the stance adopted by some economists (e.g., Becker, 1965) that investment in human capital in a particular activity is positively related to the amount of time at such activity, and the accumulation of human capital is a recursive process: Individuals gain experience and acquire labour market work skills in jobs, which in turn enhance their employability and life chances; on the other hand, being unemployed or outside the labour market due to family care responsibilities bring an opposite effect. We will therefore need to include work and life history information in the calibrated scale. The BHPS is an ideal source of data for this purpose, since it contains comprehensive work and life history data in its panel, as well as detailed relevant retrospective information collected from respondents in Waves 2 and 3.

THE ESTIMATING PROCEDURES

We first estimate a "wage equation", which calculates the value of various personal characteristics from their effect on the wage earning capacity. The coefficients of the wage equation will then be used as a basis for constructing an index of human capital. For people who are not in employment, their lack of a wage does not necessarily imply that they have no wage-earning capacity. People might well be constrained to be outside employment, or prefer not to be involved in labour market work. In fact their absence from the labour market may be systematically related to their embodied resources. Hence it is important that we include those who are not in employment in our procedures for calculating the potential economic value of various characteristics.

To achieve this aim, we follow the conventional economic procedure (Heckman, 1976) of combining an estimation of the probability of an individual's selection into employment, with an appropriately adjusted regression estimate of the economic value of the various characteristics for those actually in employment. There is, however, a constraint on using the Heckman's approach. Since the Essex Score is designed as a tool to measure patterns of differentiation in economic power, we have to be careful in the selection of variables that are of theoretical interest to us when calibrating the index. That is, we have to avoid the circularity of arguments by using the score to predict some socio-economic status variables (e.g., gender and family status) that have already been included for constructing the index. Therefore, although we expect some socio-economic status variables to be strongly associated with individuals' potential wages, we will exclude them from the composite index of economically salient resources.

Data

The data for this exercise come from the British Household Panel Survey (BHPS). This is a longitudinal survey that interviewed all members of a random selection of British households in 1991, and re-interviews all the original household members, their natural descendents, and all their current household co-residents on an annual basis. The BHPS provides a detailed and careful collection of wage data, which enables us to connect various historical and other accumulated personal characteristics with their current consequences in terms of respondents' wage rates. We have currently thirteen years of data from the panel itself, together with a considerable collection of retrospective information on employment and other circumstances prior to the start of the panel[1].

Measuring Job Quality

To estimate the wage equation, it is essential to tap the skills and characteristics that are required by a particular sort of jobs. To recapitulate, we are interested in how people's work and life history will affect their economic advantages in a particular sort of jobs, and in particular how their work experience might help them accumulate human capital for such jobs. Certainly different jobs have different "qualities" and will require different levels of skills, knowledge and commitment. We therefore need an index that can reveal the quality of jobs effectively, and preferably, in a neat manner.

In an interim attempt to develop the Essex Score, Gershuny (2000) used the Hope-Goldthorpe Scale as a rough approximation to a measure of job quality. Nevertheless, as Gershuny (2002) pointed out, the Hope-Goldthorpe Scale was calculated explicitly as an indicator of occupational standing or prestige. As such we might expect it to be correlated

[1] The work-life history data used in this exercise were complied and updated to Wave 12 by Brendan Halpin, and were further updated to Wave 13 by the first author of this paper using a different set of program codes. This might result in some incongruities in the estimation of the scale for the Wave 13 data when compared to the rest of waves.

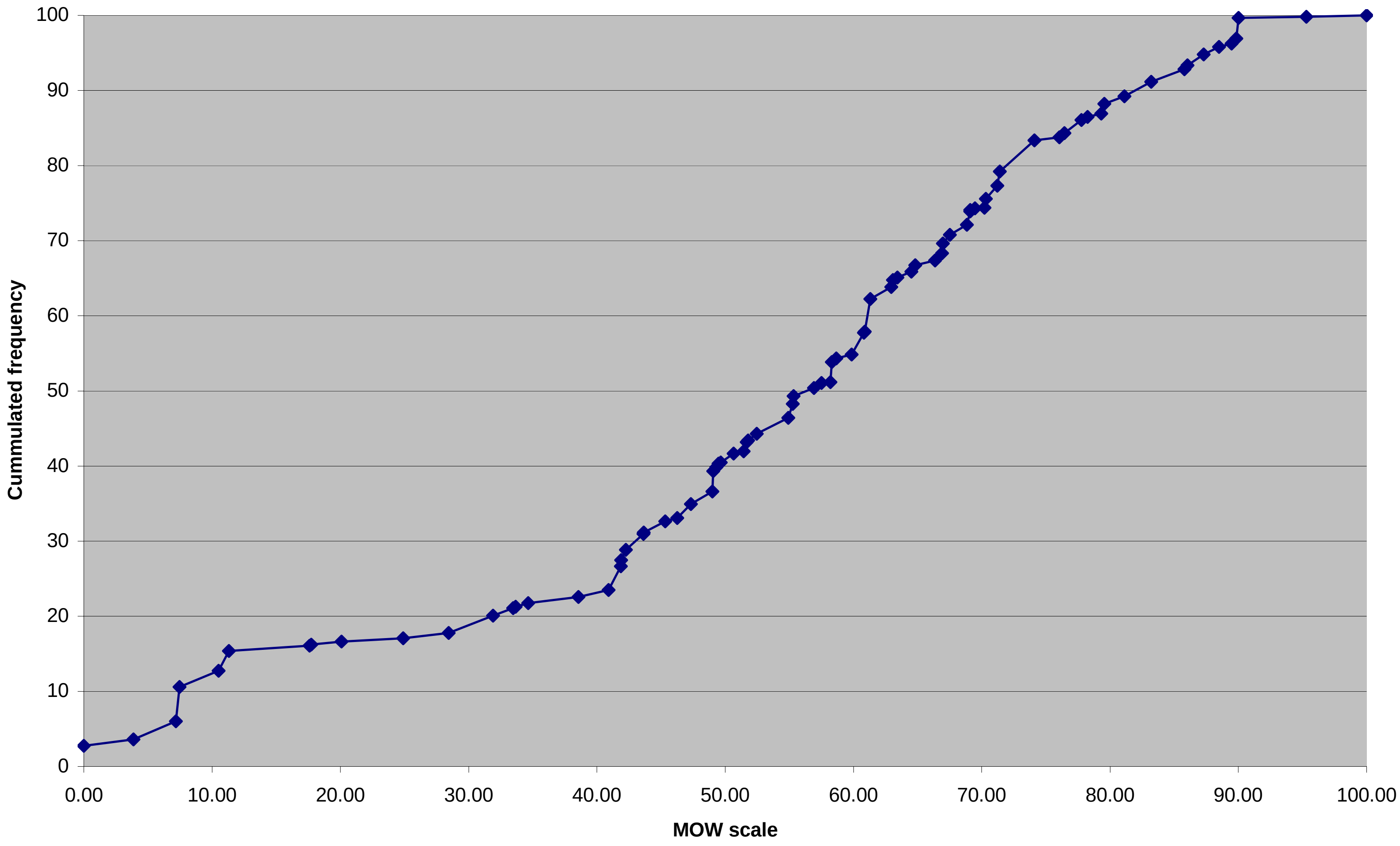
quite strongly with some aspects of the sort of job quality that we are interested in, but also confound them with other historical reflections on the status and authority of the job of the sort. (In fact, the Hope-Goldthorpe Scale was established on the basis of empirical research conducted nearly thirty years ago.) Therefore, following Gershuny (2002), in this exercise we estimate the quality of jobs by their market valuation (i.e., the expected income levels of people who are engaged in these jobs). A central component of the Essex Score is the Mean Occupational Wage (MOW) scale of job quality. This is constructed by pooling all the 13 waves of BHPS responses (yielding 29,100 observations), adjusting monthly incomes by the Consumer Price Index, and then calculating the mean for each 2-digit (or 3-digit) group in the standard occupational classification. Some selected large 2-digit SOC groups are broken down by their third digit (i.e., SOC41, SOC41.3 and SOC41.6; SOC72, SOC72.3, and SOC72.6; SOC95 and SOC95.5) so as to capture more variatiance in the hourly wage rates within each of them.

Incomes, as we would expect, are widely spread, with a cluster of relatively poorly paid occupations, and progressively sparser numbers of people in increasingly well-paid jobs. A straightforward representation of mean income would therefore provide a relatively inefficient indicator of job quality. Hence we take the natural log of income for each occupational category, and then normalized the result so that the lowers-income job is scored 0, and the highest is scored 100. A description of the MOW Scale by standard occupational classification groups is given in Appendix A.

Figure 1 shows the cumulated distribution of BHPS respondents by the MOW Scale.

As we can see, there is a slight cluster of some large occupations with low MOW scores. For example *Childcare and related occupations* and *Sales assistant* have a difference of 0.30 in their scores but account for 4.57% of all employees. Apart from this, however, the scale provides a reasonably even distribution of respondents across the full range of the scale, with around 50% of all respondents located below a MOW score of 55.

Figure 1. Distribution of Employed BHPS Respondents by MOW Scale



As we have expected, the MOW score correlates strongly with the Hope-Goldthorpe Scale (a measure of occupational grading or prestige) and the Cambridge Score[2] (a measure of similarity of lifestyle and hence social status), with the correlation coefficients being 0.668 and 0.488 respectively.

A Heckman Wage Estimation of Log Hourly Wage

The regression stage of the Heckman procedure provides estimates for the equations:

$\text{lwage}_f = f($age agesq mow mowsq higra agegr agrsq medgra agemd agmsq educ1 to educ6, jobtot1 to jobtot4, famtot1 to famtot4, unmtot1 to unmtot4), for all respondents age > 15 and age < 65 -- (1)

$\text{lwage}_g = g($age agemow higra agegr medgra agemd educ1 to educ6, jobtot1 to jobtot4, famtot1 to famtot4, unmtot1 to unmtot4), for all respondents age > 54 -- (2)

where:

- lwage_f and lwage_g are the logarithms of hourly wages;
- mow and mowsq are the MOW score and its square
- higra is a dummy variable indicating membership of the top 10% of the MOW scale (83 – 100) and medgra indicates membership of the next 30% (61 – 83);
- agegr agrsq agemd agmsq are the products and squared products of age and the high and medium grade dummies, introduced to allow for differing age/earning curves;
- agemow is the product of age and the MOW score;
- educ1 to educ6 provide dummy variables of educational attainment for, respectively, *Higher Degree, First Degree, Other tertiary qualification, A Level, O Level/Higher grade GCSE and other GCSE/CSE* (the reference category being *All other qualifications/No qualifications*);
- jobtot_ famtot_ unmtot_ represent respectively the number of months in employment, family care and unemployment in each of the four years immediately preceding the date of interview.

These variables are chosen on the basis of the theoretical considerations mentioned before: Age, educational attainment, employment and life history are associated with one's employability and earnings capacity; job quality, as estimated by MOW score, taps the

[2] Throughout this paper, we use the revised version of the Cambridge Score, which was calculated from a larger data set than the earlier version and where female occupations were specifically incorporated. For more details,

specific skills and personal characteristics required by a particular occupation. These variables are selected also because they are available throughout all waves of the BHPS, and also in the retrospective occupational histories collected in waves 2 and 3 (which allows us to make Essex Score estimates for the earliest waves of the panel study).

The selection stage of the Heckman procedure includes the same set of variables mentioned above plus gender to help identify the selection equation. The inclusion of the control for gender is aimed to adjust the size of the coefficients in the regression stage of the equations. The coefficient for the gender variable, however, is not used in the imputation of the Essex Score; our intention is to avoid any circularity of argument when we use the score to estimate the difference in economic advantages between men and women.

We again use a pooled file of the full set of 13 waves of BHPS data. The four previous years' employment history data, where appropriate, are taken from retrospective materials. First, the equation is estimated for a pooled sample of respondents aged between 16 and 64 (n = 132,747). The output from the Stata programme is provided in Appendix B. The coefficients for age and gender are both significant in the selection equation, indicating that they both affect an individual's likelihood of being in employment. In fact, we have also expected to find that age would affect our estimation of earnings potential. For people approaching their retiring age, we expect that both age and MOW would have a more or less linear effect on their log wage earning rates. We therefore run a second regression, restricted the sample to respondents aged over 54 (n = 49,244), where the square terms for age and MOW are removed and replaced by an interaction between age and MOW. The output from the Stata programme for this part of analysis is provided in Appendix C.

The Essex Score

Finally, we use the coefficients from the Heckman regression stage to estimate a predicted value for the log (shadow) wage rate for each respondent of each wave of BHPS. Since we have run two Heckman regressions to take account of the older age effect on potential wage rate, we will therefore need to use weights to adjust the estimated values from the two equations. For respondents aged between 16 and 54, we estimate their log shadow wage rate by equation 1 only; for respondents aged over 64, we estimate their log shadow

---

see Prandy (1990).

wage rate by equation 2; for those aged between 55 and 64, we estimate their log shadow wage rate by the following equation: $lwage_{age55-64}$ = (age - 54)*$lwage_g$+(1 - (age - 54))*$lwage_f$. We refer the exponential of the predicted log shadow wage for each respondent as the Essex Score.

Table 1 presents a description of the Essex scores by year[3].

Table 1. *Essex Score by Year*

	Mean	*SD*	*N*
1991	5.07	3.06	9898
1992	5.21	3.19	9430
1993	5.33	3.29	8995
1994	5.43	3.38	9027
1995	5.49	3.43	8792
1996	5.55	3.45	9090
1997	5.63	3.49	9066
1995	5.73	3.54	8872
1999	5.75	3.51	8728
2000	5.87	3.59	8607
2001	5.90	3.58	8503
2002	5.92	3.60	8279
2003	5.94	3.58	8123

Data source: BHPS, 1991 – 2003. The sample contains all respondents aged over 15. All values are weighted.

[3] Readers may notice that the Essex scores estimated in this exercise are smaller than that reported in Gershuny (2002). The differences are due to: (a) In our earlier attempt to construct the index, Gershuny restricted the sample to respondents aged 20 – 59 and ran one Heckman regression for the whole sample, whereas in our current analysis we have included all respondents aged over 15 and have run two regressions respectively for respondents aged between 16 and 64 and for those aged over 54. (b) There was an arithmetic error in Gershuny's (2002) report: The scores reported should have been scaled down by 75%.

Figure 2. Correlations between the 1991 Essex Score and Essex Scores in subsequent years

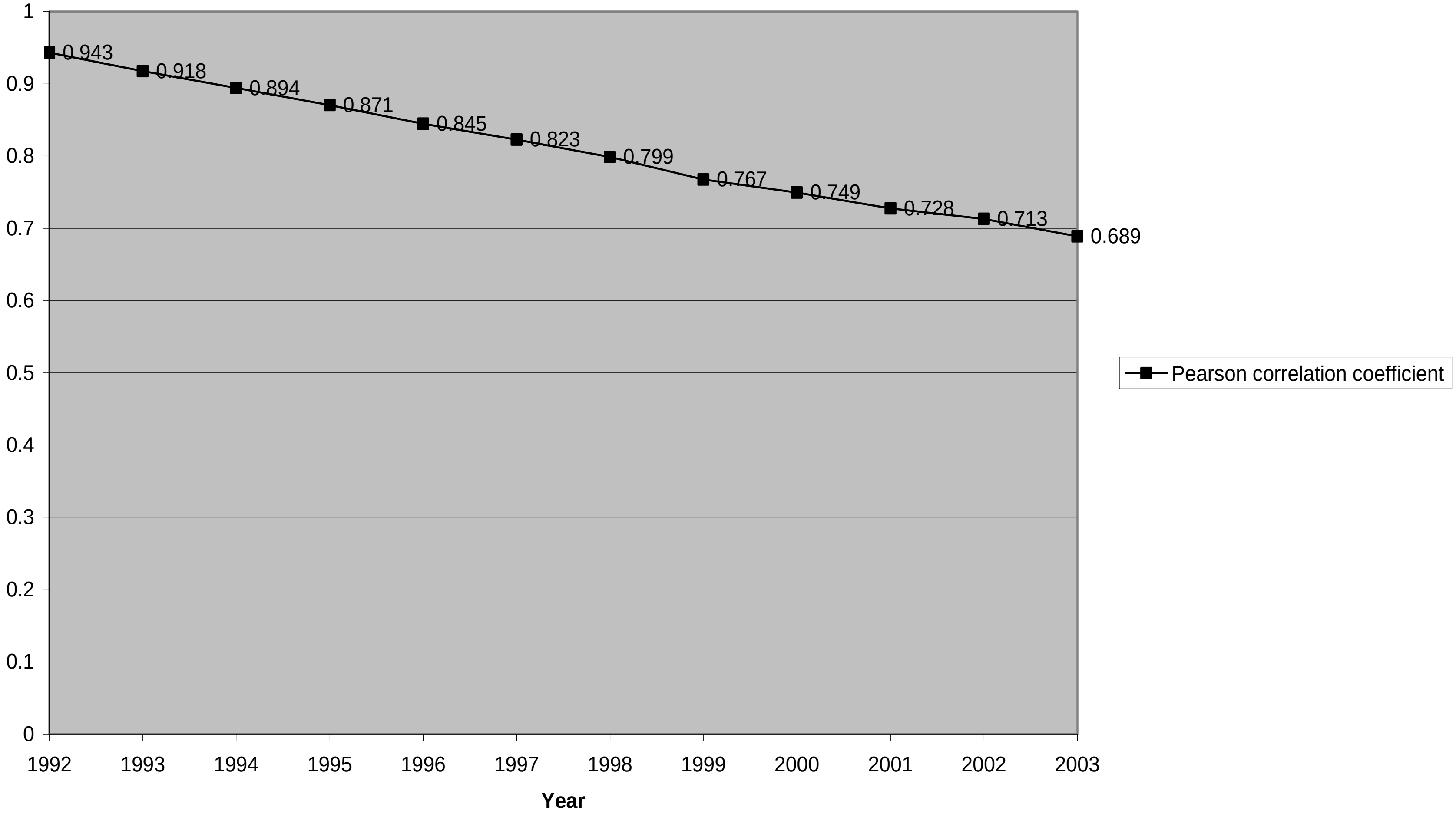


Figure 2 gives the correlations between the 1991 Essex Score and Essex Scores in subsequent years.

As have expected, the correlations between the 1991 score and the scores in subsequent years are very high. The correlation with the 1992 score equals 0.943 and the figure decreases quite proportionally over the years. The correlation between the 1991 and 2003 scores is still strong and equal to 0.689.

Since the Essex Score is designed as an indicator of potential wage rate, it is interesting to compare it with the current wage rate of employed respondents. Table 2 describes the mean hourly wage derived from BHPS data[4].

Table 2. *Mean Hourly Wage of Employees by Year*

	Mean	*SD*	*N*
1991	5.88	4.11	4946
1992	6.41	5.48	4563
1993	6.40	4.19	4348
1994	6.73	4.94	4415
1995	7.01	5.00	4316
1996	7.20	5.61	4460
1997	7.57	5.95	4580
1998	7.89	6.52	4580
1999	8.37	7.81	4496
2000	8.65	5.82	4410
2001	9.13	6.13	4377
2002	9.48	6.58	4271
2003	9.84	7.56	4192

Note: Data from BHPS, 1991 – 2003. The sample contains all employed respondents. All values are weighted.

As we can see from Tables 1 and 2, the Essex Scores generally agree with the trend of gradual increments in hourly wage between 1991 and 2003, although the degree of increase in Essex Scores is smaller than that in the real wage rate. The predicted potential hourly wages on average are also lower than that of the actual hourly wages estimated by wage information reported by respondents. These results are within our expectation. To estimate

[4] This is derived from the monthly gross income from paid work and the number of hours normally worked per week reported by the respondent. Hourly wage = (Normal monthly wage*12)/(Normal weekly work hours*52).

potential hourly wage, we have included those who were outside employment in our analysis. These respondents presumably had lower average value of the economically salient characteristics than those who were employed. That is to say, the premium on Essex Score earned by certain qualifications will be smaller than that actually observed in the real earnings data estimated from a sample with non-zero earnings.

Table 3. *Essex Score by Gender and Employment Status in 1991*

	Men			Women		
Employment status	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*
Full-time employed	7.50	3.03	2872	6.24	2.62	1709
Part-time employed	5.91	3.97	220	5.15	2.02	1079
Unemployed	3.91	1.94	371	3.56	1.54	160
Non-employed	3.02	1.43	1358	2.83	1.19	2479
Total	5.89	3.35	4821	4.39	2.45	5427
	Mean	*SD*	*N*			
All (men+women)	5.09	3.01	10248			

Note. The sample contains respondents aged over 15 in Wave 1 (1991) of BHPS. All values are unweighted.

Table 3 gives a basic breakdown of the Essex Score estimation by gender and employment status. The patterns of results generally agree with our expectations. For example, in all categories, the predicted shadow hourly wage of women is lower than that of men; the part-time employed have a lower potential hourly wage than the full-time employed. Moreover, the potential wage rate of those outside employment is lower than that of the employed.

Given that the Essex Score is designed as an indicator of social position, it is not uninteresting to compare it with the subset the respondents that can be allocated a Goldthorpe occupational class. The results are shown in Table 4.

Table 4. *Essex Score by Gender and Goldthorpe Class in 1991*

Goldthorpe Occupational Class: present job	Men			Women		
	Mean	*SD*	*N*	*Mean*	*SD*	*N*
Service class, higher grade	11.29	3.16	614	10.24	2.96	228
Service class, lower grade	8.33	2.72	554	7.39	2.54	648
Routine non-manual employees	5.64	1.76	198	5.10	1.11	699
Personal service workers	3.48	0.85	61	4.25	0.92	386
Small proprietors with employees	6.65	1.54	108	6.65	1.69	47
Small proprietors without employees	6.23	1.72	284	5.27	1.60	94
Farmers, smallholders	6.46	1.57	48	6.25	2.09	6
Foreman, technicians	6.33	1.90	346	4.82	0.94	121
Skilled manual workers	6.05	1.76	450	4.72	1.23	86
Semi/unskilled manual workers	5.05	1.37	594	4.31	0.93	521
Agricultural workers	4.53	1.01	34	4.13	0.77	17
Total	7.22	3.14	3291	5.77	2.46	2853

	Mean	*SD*	*N*
All (men+women)	6.55	2.93	6144

Note. The sample contains respondents aged over 15 in Wave 1 (1991) of BHPS. All values are unweighted.

In general we find the expected sorts of ordering, with, for example, the higher service class with much the highest levels of human capital, and personal service workers with the lowest scores. We see less expected results in the gender differentials, with women proprietors with employees having as much human capital as their male counterparts, and women personal service workers having a higher level of human capital than men. These results presumably reflect gender differences of process of selection into and out of these occupational groups. The statistical association between Goldthorpe Class and the Essex Score is very high; the Eta equals 0.721, implying that 52% of the variation in the Essex Score could be explained by membership of Goldthorpe classes, while sex and Goldthorpe class together account for around 54% of this variation. We also find that the Essex Score are strongly correlated with Hope-Goldthorpe Scale and Cambridge Score[5], with the correlations being 0.659 and 0.605 respectively.

[5] See footnote number 2.

One particular advantage of the Essex Score approach over that of the Goldthorpe occupational class, Hope-Goldthorpe Scale and Cambridge Scores, however, is that it is *in principle* comprehensive: Each member of the population is provided with an index, irrespective of their present or previous employment status. So for example, the 10,248 respondents in Table 3 have been reduced to 6,144 in Table 4. The missing 4,104 are people aged over 15 who do not have a current classifiable employment. But these people are not in fact entirely without potentially economically salient personal resources.

Predicting Future Earnings in the Labour Market

In the following, we will compare the leverage of the Essex Score and that of the Goldthorpe Class in predicting future wage earnings. To restrict our analysis to respondents at normal working age throughout the panel, here we select cases where respondents were aged between 20 and 48 and had both an Essex Score and a Goldthorpe Class in 1991 (i.e., all respondents were aged between 20 and 60 throughout the period between 1991 and 2003). Table 5 describes the fit of the variously specified models involving the 1991 Essex Score, the 1991 Goldthorpe Class, and the combinations of these with each other and with the control variables (age, sex, age square, the interaction between age and sex, and the interaction between age square and sex). The first six rows of Table 5 set out the levels of variance in individual employment income at various points in time. We have exactly the same cases involved in each model; those cases with an Essex Score but no Goldthorpe Class in 1991 have been dropped (For example, of the 2,295 cases where there was an income from the labour market in 2003, 349 had an Essex Score but no Goldthorpe Class in 1991). Therefore, by carefully comparing the explanatory power of these models, we can derive straightforward estimates of the causal impact of the variables, using an extension of the simple but effective causal modelling technique suggested by Simon (1954).

Table 5. *OLS models of future earnings in the labour market: 1991 Essex Score Vs Goldthorpe Class Measure*

	Adjusted R-Square						
	1991	1993	1995	1997	1999	2001	2003
1. Essex Score + controls	0.495	0.469	0.440	0.418	0.422	0.406	0.388
2. Essex Score alone	0.372	0.340	0.316	0.289	0.284	0.260	0.246
3. Goldthorpe Class + controls	0.483	0.461	0.425	0.412	0.416	0.405	0.378
4. Goldthorpe Class alone	0.326	0.311	0.280	0.268	0.258	0.248	0.213
5. Goldthorpe Class + Essex Score + controls	0.531	0.503	0.465	0.447	0.448	0.433	0.411
6. Controls alone	0.270	0.254	0.239	0.229	0.237	0.235	0.225
7. Essex Score contribution (1-6)	0.225	0.215	0.201	0.190	0.185	0.170	0.163
8. Goldthorpe Class contribution (3-6)	0.214	0.207	0.186	0.183	0.179	0.170	0.152
9. Essex Score unique (5-3)	0.048	0.042	0.040	0.035	0.032	0.028	0.033
10. Goldthorpe Class unique (5-1)	0.036	0.033	0.026	0.028	0.027	0.028	0.023
11. % of Essex Score by Goldthorpe Class	0.789	0.806	0.799	0.817	0.825	0.836	0.796
12. % of Goldthorpe Class by Essex Score	0.833	0.838	0.862	0.846	0.852	0.838	0.848
N	3666	2763	2570	2480	2301	2143	1946

Note: The dependent variables are the natural logarithms of the usual gross monthly wage earnings in various years. The control variables are age, sex, age square, the interaction between age and sex, and the interaction between age square and sex in 1991. All values are unweighted. The samples contain respondents aged between 20 and 48 in 1991 who had both an Essex Score and a Goldthorpe Class.

Rows 1 and 3 give the proportions of variance in wage earnings at various points from 1991 to 2003. The effects of the control variables in these two models are summarized in row 6. In predicting wage earnings, there is a small margin of predictive power of the row 1 Essex Score models over the row 3 Goldthorpe Class models. However, we cannot see the scale of the advantage from these two rows, since in each case clearly part of the variance in income explained by the class variable is also associated with the row 6 control variables. We can obtain an irreducible minimum estimate of the respective explanatory contributions of the Essex and Goldthorpe indicators by subtracting the whole of the variance explained by the row 6 variables from rows 1 and 3. These figures are set out in rows 7 and 8. They indicate that the Essex Score is slightly stronger than Goldthorpe Class. For example, the Essex Score explains 16% of wage earnings in 2003 while Goldthorpe Class explains 15%.

The most important question, though, concerns whether the two indicators explain broadly the same variation. Row 5 shows the fit of the models where both indicators and the control variables are included. As we can see, there is not a great margin of difference either between row 5 and the Goldthorpe Class models in row 3, or the Essex Score models in row 1. Furthermore, by comparing rows 9 and 10 with rows 7 and 8, we see that the Essex Score explains some 83 – 86% of all the variance explained also by Goldthorpe Class, while Goldthorpe Class explains some 79 – 84% of the same variance explained by the Essex Score. Again, the results suggest that the predictive powers of the two indicators are similar but the Essex Score is slightly stronger.

In fact we may be surprised at how well the Goldthorpe categories do perform at predicting future earnings in the labour market. Paradoxically, it is with those who are, at a particular point in time, *outside* the labour market, that the wage-equation-based Essex Score's main advantage lies: in its capacity to incorporate the entire population's employment and non-employment experience within an analysis of advantaged and disadvantaged social positions, as well as maintaining a robustly strong leverage in explaining future wage levels.

CONCLUDING REMARKS

We have constructed the Essex Score, a non-categorical indicator of personal resources salient to labour market performance. The Essex Score has advantages over conventional occupational class index since it covers both men and women, and all adult age groups irrespective of their current or past employment status. It is calibrated from an index that takes account of durations of both employment and non-employment activities. Furthermore, its continuous measurement also allows aggregation from individual to household levels, as well as sensitive investigation of the determinants and consequences of changes in social position during the life course. We have deposited a data set of the MOW score and the Essex Score in the UK Data Archive (a description of the variables in the data set is provided in Appendix D). It will be a useful too for a wide variety of research that involves comparing possession of economic power among different social groups, as well as for social mobility studies that focuses on intra- or intergenerational changes in life chances.

REFERENCES

Becker, G. S. (1965). A Theory of the Allocation of Time. *Economic Journal, 75*, 493-517.

Becker, G. S. (1991 [1981]). *A Treatise on the Family* (Enl. ed.). Cambridge, Mass; London: Harvard University Press.

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). New York: Columbia University Press.

Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology, (S)95*, 95-120.

Gershuny, J. (2000). Social position from narrative data. In R. Crompton & F. Devine & M. Savage & J. Scott (Eds.). *Renewing class analysis (Sociological review monographs).* Oxford: Blackwell.

Gershuny, J. (2002). A new measure of social position: social mobility and human capital in Britain. In *Working Papers of the Institute for Social and Economic Research* (pp. paper 2002-02). Colchester: University of Essex.

Goldthorpe, J. H., & Hope, K. (1974). *The social grading of occupations*. Oxford: Oxford University Press.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection, and limited dependent variables, and a simple estimator for such models. *The Annals of Economic and Social Measurement, 5,* 475-492.

Polachek, S. (1985). Occupational Segregation: A Defence of Human Capital Predictions. *Journal of Human Resources, 20*(3), 437-440.

Prandy, K. (1990). The revised Cambridge Scale of occupations. *Sociology, 24*(4), 629-655.

Schultz, T. W. (1961). Investment in human capital. *the American Economic Review, LI*(1), 1-17.

Scott, J. (1996). *Stratification and power: Structures of class status and command*. Cambridge: Polity.

Simon, H. A. (1954). Spurious correlation: A causal interpretation. *Journal of the American Statistical Association*, *49*, 467-479.

Stewart, A., Prandy, K., & Blackburn, R. M. (1980). *Social stratification and occupations*. London: Macmillan.

APPENDIX A

The MOW Scale

(Calibrated from Waves 1 – 13 of the BHPS)

	SOC	*N*	CUM FREQ	LOG INCOME	MOW
Other sales and servs: cleaners, domestics	95.5	2319	2.75	5.43	0.00
Retail cash desk & check-out operators	72.3	719	3.60	5.52	3.87
Childcare & related occupations	65	2039	6.02	5.59	7.17
Sales assistants	72	3855	10.59	5.60	7.47
Other occs in sales & service (part)	95	1815	12.74	5.67	10.51
Catering occupations	62	2222	15.37	5.69	11.31
Hairdressers, beauticians & related occupations	66	608	16.09	5.83	17.59
Petrol pump forecourt attendants	72.6	115	16.23	5.83	17.72
Mobile market & door-to-door salespersons & agents	73	353	16.64	5.89	20.08
Personal & protective service occupations nec	69	363	17.08	5.99	24.89
Domestic staff &related occupations	67	606	17.79	6.08	28.44
Health & related occupations	64	1925	20.07	6.15	31.90
Receptionists, telephonists & related occupations	46	862	21.10	6.19	33.46
Travel attendants & related occupations	63	155	21.28	6.19	33.67
Sales occupations nec	79	395	21.75	6.22	34.63
Other occupations in agriculture, forestry & fishing	90	687	22.56	6.31	38.55
Other occupations in communication	94	782	23.49	6.36	40.90
Clerks (not otherwise specified)	43	2669	26.65	6.38	41.86
Textiles, garments & related trades	55	707	27.49	6.38	41.89
Counter clerks & cashiers	41.3	1162	28.87	6.39	42.26
Secretaries, personal ass, typists, word processor operators	45	1783	30.98	6.42	43.62
Debt, rent & other cash collectors	41.6	145	31.15	6.42	43.65
Other routine process operatives	86	1258	32.64	6.46	45.33
Clerical & secretarial occupations nec	49	365	33.08	6.48	46.26
Filing & records clerks	42	1567	34.93	6.50	47.32
Stores & despatch clerks, storekeepers	44	1411	36.60	6.54	48.99
Accounts & wages clerks, book-keepers etc	41	2287	39.32	6.54	49.07
Food, drink & tobacco process operatives	80	450	39.85	6.55	49.29
Food preparation trades	58	418	40.34	6.55	49.48
Textiles & tannery process operatives	81	96	40.46	6.56	49.66
Other craft &related occupations nec	59	1030	41.68	6.58	50.65
Other occupations in transport	93	228	41.95	6.60	51.43
Social welfare associate professionals	37	1055	43.20	6.60	51.67
Other occupations in mining & manufacture	91	167	43.40	6.61	51.79
Assemblers/lineworkers	85	756	44.29	6.62	52.46
Construction trades	50	1783	46.41	6.68	54.92
Admini/clerical officers etc in civil service & local govt	40	1562	48.26	6.69	55.27
Vehicle trades	54	894	49.32	6.69	55.33
Woodworking trades	57	921	50.41	6.72	56.91
Other occupations nec	99	556	51.07	6.74	57.50
Librarians & related professionals	27	74	51.15	6.75	58.20
Road transport operatives	87	2262	53.84	6.75	58.30
Other occupations in construction	92	423	54.34	6.76	58.67
Printing & related trades	56	428	54.84	6.79	59.85
Health associate professionals	34	2441	57.74	6.81	60.80
NCOs & other ranks, armed forces	60	115	57.87	6.81	60.88
Managers & proprietors in service industries	17	3667	62.22	6.82	61.32

APPENDIX A – *Continued*

	SOC	*N*	CUM FREQ	LOG INCOME	MOW
Artists, musicians, athletes	38	1356	63.83	6.86	62.94
Professional occupations nec	29	793	64.77	6.86	63.08
Metal working process operatives	84	277	65.09	6.87	63.44
Managers in farming, horticulture, forestry & fishing	16	656	65.87	6.90	64.51
Chemicals, paper, plastics & related process operatives	82	721	66.73	6.90	64.81
Other transport & machinery operatives	88	525	67.35	6.94	66.36
Scientific technicians	30	815	68.31	6.95	66.90
Plant & machine operatives nec	89	1100	69.62	6.95	66.96
Metal forming, welding & related trades	53	987	70.79	6.96	67.52
Sales representatives	71	1117	72.11	6.99	68.85
Security & protective service occupations	61	1503	73.89	7.00	69.10
Metal making &treating process operatives	83	152	74.07	7.00	69.11
Buyers, brokers & related agents	70	198	74.31	7.01	69.47
Legal associate professionals	35	61	74.38	7.03	70.21
Associate professional & technical occupations nec	39	1005	75.57	7.03	70.33
Electrical/electronic trades	52	1478	77.32	7.05	71.21
Metal machining, fitting &instrument making trades	51	1576	79.19	7.05	71.41
Teaching professionals	23	3511	83.35	7.11	74.11
Draftspersons, quantity & other surveyors	31	370	83.79	7.16	76.04
Managers in transport & storing	14	446	84.32	7.17	76.45
Financial inst office managers, civil service exec officers	13	1478	86.07	7.20	77.76
Natural scientists	20	365	86.50	7.21	78.23
Architects, town planners &surveyors	26	354	86.92	7.23	79.31
Business & financial associate professionals	36	1108	88.23	7.24	79.54
Computer analyst/programmers	32	842	89.23	7.27	81.11
Managers & administrators nec	19	1616	91.15	7.32	83.21
Engineers & technologists	21	1429	92.84	7.38	85.79
General man & admin in Government, big companies	10	419	93.34	7.39	86.04
Production managers in manufacturing, construction mining etc	11	1235	94.80	7.41	87.29
Business & financial professionals	25	843	95.80	7.44	88.50
Legal professionals	24	392	96.26	7.46	89.46
Health professionals	22	553	96.92	7.47	89.85
Specialist managers	12	2321	99.67	7.48	90.01
Ship & aircraft officers, air traffic planners & controllers	33	133	99.83	7.60	95.29
Protective service officers	15	145	100.00	7.70	100.00
	TOTAL	84380			

APPENDIX B

The Heckman Estimation of the Essex Score I

```
/******************************
Estimating wage equation ****
*****************************/

version 8
clear
set more off

capture log close
set matsize 150
set memory 50m
log using "M:\Time and Gender\Man Yee\humscore3.log",replace
use "M:\Time and Gender\poolfil2.dta"

. keep if (age > 15 & age < 65)

generate woman=sex-1

#delimit ;
delimiter now ;
heckman lwage age agesq mow mowsq higra agegr agrsq medgra agemd agmsq educ*
jobtot* famtot* unmtot*,
select (woman age agesq mow mowsq higra agegr agrsq medgra agemd agmsq educ*
jobtot* famtot* unmtot*)
//

Heckman selection model                         Number of obs      =    132747
(regression model with sample selection)        Censored obs       =     53303
                                                Uncensored obs     =     79444

                                                Wald chi2(28)      =  66831.22
Log likelihood = -102477.6                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lwage        |
         age |   .0456562   .0010851    42.08   0.000     .0435295    .0477829
       agesq |  -.0005182   .0000139   -37.33   0.000    -.0005454    -.000491
         mow |  -.0060967   .0002658   -22.93   0.000    -.0066177   -.0055757
       mowsq |   .0001516   3.91e-06    38.76   0.000     .0001439    .0001593
       higra |  -.9272212   .0758682   -12.22   0.000     -1.07592   -.7785223
       agegr |   .0378637   .0038067     9.95   0.000     .0304027    .0453246
       agrsq |  -.0003761   .0000466    -8.07   0.000    -.0004675   -.0002847
      medgra |  -.3886601   .0365314   -10.64   0.000    -.4602603     -.31706
       agemd |   .0195205   .0019567     9.98   0.000     .0156854    .0233556
       agmsq |  -.0002143    .000025    -8.58   0.000    -.0002633   -.0001654
       educ1 |   .6645064   .0103622    64.13   0.000     .6441968    .6848159
       educ2 |   .5590279   .0065033    85.96   0.000     .5462817     .571774
       educ3 |   .4078464   .0070378    57.95   0.000     .3940525    .4216404
       educ4 |   .2988307     .00532    56.17   0.000     .2884036    .3092577
       educ5 |   .1997042   .0049349    40.47   0.000     .1900319    .2093765
       educ6 |   .1130838   .0073891    15.30   0.000     .0986014    .1275661
     jobtots |   .0102185   .0008177    12.50   0.000     .0086159    .0118211
     jobtotr |   .0050216   .0007671     6.55   0.000     .0035181     .006525
     jobtotq |   .0031031   .0007718     4.02   0.000     .0015903    .0046158
     jobtotp |   .0084368   .0005628    14.99   0.000     .0073338    .0095398
     famtots |   -.003441   .0015155    -2.27   0.023    -.0064112   -.0004707
     famtotr |  -.0011458   .0014547    -0.79   0.431    -.0039969    .0017053
     famtotq |  -.0009802   .0015366    -0.64   0.524    -.0039918    .0020314
     famtotp |  -.0068435   .0012287    -5.57   0.000    -.0092517   -.0044353
     unmtots |   -.014525   .0018051    -8.05   0.000    -.0180628   -.0109871
     unmtotr |  -.0065449    .001639    -3.99   0.000    -.0097574   -.0033324
     unmtotq |  -.0042949   .0016564    -2.59   0.010    -.0075414   -.0010484
     unmtotp |  -.0057391   .0014622    -3.92   0.000     -.008605   -.0028732
       _cons |   .2600163   .0200754    12.95   0.000     .2206692    .2993635
-------------+----------------------------------------------------------------
```

APPENDIX B – *Continued*

```
select       |
       woman |   .3644978   .0094725    38.48   0.000     .3459322    .3830635
         age |  -.0158217     .00256    -6.18   0.000    -.0208392   -.0108043
       agesq |  -.0000259   .0000323    -0.80   0.421    -.0000892    .0000373
         mow |   .0241615   .0006379    37.88   0.000     .0229113    .0254116
       mowsq |  -.0003537   9.20e-06   -38.45   0.000    -.0003718   -.0003357
       higra |   3.156352   .2580882    12.23   0.000     2.650508    3.662195
       agegr |  -.0959091   .0124821    -7.68   0.000    -.1203735   -.0714447
       agrsq |   .0009966   .0001469     6.79   0.000     .0007088    .0012844
      medgra |   1.705382    .107804    15.82   0.000      1.49409    1.916674
       agemd |   -.060599   .0055879   -10.84   0.000     -.071551   -.0496469
       agmsq |   .0006042   .0000689     8.77   0.000     .0004692    .0007393
       educ1 |   .5979402    .030173    19.82   0.000     .5388021    .6570783
       educ2 |   .5534396   .0169597    32.63   0.000     .5201991      .58668
       educ3 |   .5306473   .0190975    27.79   0.000     .4932169    .5680776
       educ4 |    .410022    .012857    31.89   0.000     .3848226    .4352213
       educ5 |   .4447405   .0116929    38.04   0.000     .4218229    .4676582
       educ6 |   .3868779   .0195747    19.76   0.000     .3485121    .4252437
     jobtots |   .1282222   .0015589    82.25   0.000     .1251668    .1312776
     jobtotr |   .0137321   .0020576     6.67   0.000     .0096993    .0177649
     jobtotq |   .0052149   .0022122     2.36   0.018     .0008791    .0095507
     jobtotp |  -.0026495   .0016522    -1.60   0.109    -.0058878    .0005887
     famtots |  -.0675797   .0026866   -25.15   0.000    -.0728454    -.062314
     famtotr |   .0149684   .0035349     4.23   0.000     .0080402    .0218966
     famtotq |  -.0104241   .0039334    -2.65   0.008    -.0181335   -.0027148
     famtotp |  -.0117584    .003105    -3.79   0.000     -.017844   -.0056727
     unmtots |  -.0414752   .0029738   -13.95   0.000    -.0473038   -.0356466
     unmtotr |   .0153574   .0034799     4.41   0.000     .0085369     .022178
     unmtotq |  -.0113137   .0037993    -2.98   0.003    -.0187602   -.0038671
     unmtotp |  -.0112287   .0034102    -3.29   0.001    -.0179126   -.0045447
       _cons |  -.7997635   .0433015   -18.47   0.000    -.8846329   -.7148941
-------------+----------------------------------------------------------------
     /athrho |   .2205483   .0170036    12.97   0.000     .1872218    .2538748
    /lnsigma |  -.8415662   .0029279  -287.43   0.000    -.8473047   -.8358277
-------------+----------------------------------------------------------------
         rho |   .2170406   .0162026                      .1850645    .2485575
       sigma |   .4310349    .001262                      .4285685    .4335155
      lambda |   .0935521   .0071291                      .0795792    .1075249
------------------------------------------------------------------------------
LR test of indep. eqns. (rho = 0):   chi2(1) =   155.23   Prob > chi2 = 0.0000
------------------------------------------------------------------------------

//
corr lwage heckman
//
             |    lwage  heckman
-------------+------------------
       lwage |   1.0000
                            heckman |   0.6859   1.0000
```

APPENDIX C

The Heckman Estimation of the Essex Score II

```
/*****************************
Estimating wage equation ****
****************************/

version 8
clear
set more off

capture log close
set matsize 150
set memory 50m
log using "M:\Time and Gender\Man Yee\humscore2.log",replace
use "M:\Time and Gender\poolfil2.dta"

replace age= . if age < 0
keep if (age > 54)
generate agemow=age*mow

generate woman=sex-1
#delimit ;
heckman lwage age mow agemow higra agegr medgra agemd educ* jobtot* famtot*
unmtot*,
select (woman age mow agemow higra agegr medgra agemd educ* jobtot* famtot*
unmtot*);

Heckman selection model                         Number of obs      =     49244
(regression model with sample selection)        Censored obs       =     41422
                                                Uncensored obs     =      7822

                                                Wald chi2(25)      =   4313.95
Log likelihood = -13396.25                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lwage        |
         age |  -.0201173   .0026701    -7.53   0.000    -.0253507    -.014884
         mow |   .0110809   .0040009     2.77   0.006     .0032392    .0189226
      agemow |  -.0001285   .0000666    -1.93   0.054    -.0002592    2.10e-06
       higra |  -.2896851   .3649277    -0.79   0.427     -1.00493      .42556
       agegr |   .0117154   .0061498     1.91   0.057     -.000338    .0237689
      medgra |    .258276   .2334242     1.11   0.269    -.1992269    .7157789
       agemd |  -.0018173   .0039123    -0.46   0.642    -.0094853    .0058508
       educ1 |   .6406815   .0451855    14.18   0.000     .5521196    .7292435
       educ2 |   .5214922   .0250182    20.84   0.000     .4724575    .5705268
       educ3 |   .4258312   .0236404    18.01   0.000     .3794968    .4721656
       educ4 |   .2472726    .018847    13.12   0.000     .2103332     .284212
       educ5 |   .1864716   .0139394    13.38   0.000     .1591508    .2137924
       educ6 |   .0782636   .0689721     1.13   0.256    -.0569192    .2134464
     jobtots |   .0236792   .0046331     5.11   0.000     .0145984      .03276
     jobtotr |   .0024946    .003249     0.77   0.443    -.0038732    .0088624
     jobtotq |   .0052401   .0034227     1.53   0.126    -.0014682    .0119485
     jobtotp |   .0085882   .0025025     3.43   0.001     .0036835     .013493
     famtots |  -.0051897   .0088185    -0.59   0.556    -.0224737    .0120943
     famtotr |  -.0004502    .007992    -0.06   0.955    -.0161143     .015214
     famtotq |   .0048431   .0092154     0.53   0.599    -.0132188    .0229049
     famtotp |  -.0075554   .0071206    -1.06   0.289    -.0215116    .0064008
     unmtots |  -.0176526   .0084612    -2.09   0.037    -.0342364   -.0010689
     unmtotr |  -.0148343   .0068534    -2.16   0.030    -.0282667   -.0014019
     unmtotq |  -.0068003   .0069075    -0.98   0.325    -.0203387    .0067381
     unmtotp |  -.0132043   .0055616    -2.37   0.018    -.0241049   -.0023037
       _cons |   2.119857   .1559491    13.59   0.000     1.814203    2.425512
-------------+----------------------------------------------------------------
select       |
       woman |   .4134829   .0247286    16.72   0.000     .3650158      .46195
         age |  -.0627178   .0034118   -18.38   0.000    -.0694048   -.0560309
         mow |  -.0348977   .0055505    -6.29   0.000    -.0457764    -.024019
```

APPENDIX C – *Continued*

```
      agemow |   .0005778   .0000897     6.44   0.000      .000402    .0007536
       higra |   2.963664    .633771     4.68   0.000     1.721496    4.205832
       agegr |   -.046819   .0104894    -4.46   0.000    -.0673779   -.0262601
      medgra |    1.01334   .3947885     2.57   0.010     .2395683    1.787111
       agemd |  -.0186382   .0065037    -2.87   0.004    -.0313852   -.0058912
       educ1 |  -.0768694   .0895878    -0.86   0.391    -.2524583    .0987195
       educ2 |  -.1706857   .0504576    -3.38   0.001    -.2695808   -.0717906
       educ3 |  -.1039701   .0470389    -2.21   0.027    -.1961647   -.0117755
       educ4 |   .0165699   .0392717     0.42   0.673    -.0604013     .093541
       educ5 |  -.0070006   .0292575    -0.24   0.811    -.0643443     .050343
       educ6 |   .3711487   .1636068     2.27   0.023     .0504852    .6918122
     jobtots |   .2051716   .0040335    50.87   0.000     .1972661     .213077
     jobtotr |   .0209092   .0056569     3.70   0.000     .0098218    .0319966
     jobtotq |   .0002871    .006521     0.04   0.965    -.0124939    .0130681
     jobtotp |  -.0014473   .0048585    -0.30   0.766    -.0109697    .0080751
     famtots |  -.0291415   .0095886    -3.04   0.002    -.0479347   -.0103483
     famtotr |   .0143029   .0128498     1.11   0.266    -.0108822     .039488
     famtotq |  -.0101951    .015116    -0.67   0.500    -.0398219    .0194317
     famtotp |  -.0280842   .0117508    -2.39   0.017    -.0511154    -.005053
     unmtots |   .0252529   .0100852     2.50   0.012     .0054862    .0450196
     unmtotr |    .029939   .0113856     2.63   0.009     .0076236    .0522545
     unmtotq |  -.0034517   .0123719    -0.28   0.780    -.0277002    .0207967
     unmtotp |   -.007013   .0100955    -0.69   0.487    -.0267999    .0127739
       _cons |   1.514328   .2216681     6.83   0.000     1.079867     1.94879
-------------+----------------------------------------------------------------
     /athrho |   .3665401   .0535138     6.85   0.000     .2616549    .4714253
    /lnsigma |  -.7238599   .0121279   -59.69   0.000    -.7476301   -.7000897
-------------+----------------------------------------------------------------
         rho |   .3509617   .0469223                      .2558428    .4393502
       sigma |   .4848771   .0058805                      .4734873    .4965408
      lambda |   .1701733   .0243421                      .1224637    .2178828
------------------------------------------------------------------------------
LR test of indep. eqns. (rho = 0):   chi2(1) =    38.69   Prob > chi2 = 0.0000
------------------------------------------------------------------------------

//
corr lwage heckman
//
             |    lwage  heckman
-------------+------------------
       lwage |   1.0000
     heckman |   0.5891   1.0000

//
```

APPENDIX D

The Essex Score estimates for the BHPS are kept in the *shadow_wage.sav* file. Data users can match this data set with other files in the BHPS readily by PID, the cross wave person identifier. The data set contains three end-variables MOW (the Mean Occupational Score), HRWAGE (hourly wage estimated from real wage data), and ESCORE (Essex Score, shadow hourly wage), which are prefixed with the wave number (w).

Description of Variables in shadow_wage.sav

Variable	Description
PID	Cross Wave Person Identifier
wMOW	Mean Occupational Wage (MOW) Score
wHRWAGE	Hourly wage estimated from real wage data
wESCORE	Shadow hourly wage (Essex Score)

Note: w = A – M, indicating Wave Number