

Employee Heterogeneity and Within-Firm Experience-Earnings Profiles: A Nonparametric Analysis

Ross Campbell

The University of Aberdeen*

**First draft
5 October 2006**

Abstract

Motivated by a priori uncertainty with respect to the parametric specification of the earnings function, I model the earnings function as semiparametric partially linear model and follow the estimation approach described in Robinson (1988). Using data from the personnel records of a large major UK based financial sector employer, I let years of within-firm and pre-firm experience form the nonparametrically modelled component of the earnings function. It is shown that the estimated within-firm experience earnings profiles, which are conditional upon a given number years of pre-firm experience accumulated before entry, converge and even overtake as years of pre-firm experience increases. This result can be explained with the recognition of unobservable explanatory variables, such as the match and individual quality of the employees, both of which are a function of years of within- and pre-firm experience and wages.

* Address for correspondence: The University of Aberdeen, 24 High Street • Room EIF7 • Old Aberdeen • Scotland • AB24 3EB • r.campbell@abdn.ac.uk. This paper has benefited from useful discussion with Tim Barmby and Rainer Schulz.

1. Introduction

The estimation of and interpretability of experience-earnings profiles are two topics which have generated significant attention in modern labour economics.

With regard to the former, the conventional starting point for the specification of earnings functions¹ stems from Mincer's (1974) human capital earnings function in which the natural logarithm of earnings is expressed as the sum of a linear function of the number of years of school completed and as a quadratic function of years of total experience. Obviously more complex models of earnings determination necessitate the expansion of this benchmark earnings function with other covariates. Nonetheless the conventional Mincer equation - dictating the functional form relationship between earnings against schooling and experience - still remains the "workhorse" of empirical research on earnings determination and is probably the most widely used specification in empirical economics (Lemieux, 2006).

The popularity of the quadratic term in experience can be attributed to the fact that it was derived by Mincer (1974) as closed form solution to a formal theoretical model of rational human capital accumulation in post-schooling training decisions (Ben-Porath, 1967). Though whether this specification is actually the truly most parsimonious benchmark model of earnings determination was systematically questioned in a seminal paper by Murphy and Welch (1990). Using CPS data from 1964 to 1987, they illustrated that the low dimensionality of the quadratic specification could not capture important features of the profile. Finding this bias to be stable over time and across educational groups at a point in time, they proposed a quartic polynomial in total experience to be a specification with estimates of sufficiently small enough bias to be considered for use as the standard parametric specification. The important point is that higher order polynomials in the specification of total experience can still yield estimates of declining earnings growth with time and this finding still retains consistency with the predictions of human capital theory. It was Mincer's (1974) rather ad hoc assumption of a linearly declining investment ratio over the lifecycle that led to the quadratic in experience and such a specification naturally facilitates only a constant rate of decrease in earnings growth. The specification issue has attracted only somewhat moderate attention since Murphy and Welch (1990), though Robinson (2003), Lemieux (2006) and Zheng (2000) have all presented evidence which have lent support to the conjecture that higher order polynomials are needed in total experience.²

One must note, however, that the entry of total years of experience as an independent variable in an earnings function is itself somewhat restrictive and is in fact a restriction which conventional human capital theory rejects. Total years of experience can be decomposed into a number of elements. To name two: years of experience before entering a firm and years of experience at a firm. The entry of total years of experience implicitly implies that the accumulation of each element has identical affect upon wages, so that their sum is a sufficient explanatory variable. Theory suggests otherwise and whilst empirical recognition of this insight is hardly new to labour economics, the appropriate specification of an earnings function which

¹ An earnings function refers to a regression of some measure of earnings on a given set of personal, market, or environmental variables thought to influence earnings (Willis, 1986).

² Also see Basu and Ullah (1992) who respond to the specification uncertainty by estimating an earnings-age profile nonparametrically, but however make no comparison to any parametric estimates.

explicitly differentiates between the two elements of total experience has attracted little, if any, significant attention.

It is recognition for the need to differentiate between years of pre-firm and within-firm experience, as well as the a priori uncertainty regarding the entry of these variables in a parametric earnings function, which motivates the econometric approach in this paper. Using personnel data from a single large UK based financial sector employer and by modelling the earnings function as a semiparametric partially linear model where years of within-firm and years of pre-firm experience enter nonparametrically, and by estimating the earnings function across schooling groups, all a priori specification uncertainty is essentially lost.

With respect to the interpretability issue, modern human capital theory (Becker, 1964; Ben-Porath, 1967; Hashimoto, 1981; Mincer, 1974), of course, forms the conventional basis for explaining experience-earnings profiles. The theory argues that the cost of the accumulation of general human capital, through on-the-job training, is borne entirely by the worker as are the returns. The accumulation of such capital increases employees' productivity and hence wage in all firms. The investment in specific human capital however – capital which is of value only in the firm in which it was acquired - is a joint investment between the worker and the firm, with the costs and returns to the investment being shared. A direct implication of specific human capital accumulation therefore is that there exists a seniority-earnings profile and that, once specific capital is accumulated, a given worker at a point in time will earn more at the firm in which the specific capital was acquired relative to the opportunity wage offered by other firms who offer wages based upon only the workers stock of general human capital.

The interpretation of wage growth with years in the labour market solely as a reflection of returns to these human capital investments may not necessarily follow however due to the existence of other plausible explanations of wage growth. Two well known alternatives are the delayed payment model of Lazear (1979, 1981) - a model which belongs to a much wider class of long-term incentive-compatible contracting models which seek to explain life-cycle wage growth independently to contemporaneous life-cycle productivity growth and instead view wage contracts as optimal responses to the asymmetries in information between employees and firms³- and search models which are based on the existence of a match specific value which is attached to any given employee-firm relationship.

Lazear's (1979, 1981) delayed payment or shirking model is an alternative explanation for the seniority-earnings profile.⁴ Essentially this model- as well as other models in the same class -stems from the fact that employers often have considerable difficulty in monitoring employees and in assessing their marginal product. Such

³ These agency themed models with lifecycle implications fall roughly into three categories: Those which emphasise the seniority- earnings profile as stemming from either 1) The provision of life-cycle incentives by the firm; 2) selection and sorting mechanisms; or 3) Insurance motives. Becker and Stigler's (1974) model was perhaps the first to fall into the first category, arguing that firms should separate the payment of workers from their output to keep workers from shirking. Salop and Salop (1976) proposed that firms implement steep seniority-earnings profiles with earnings less than productivity upon entry in order to discourage those with high propensity to quit from joining the firm (i.e. a self-selection model to alleviate asymmetric information problems) and consequently falls into the second category. Harris and Holmstrom's (1982) model assumes that risk averse workers are uncertain about their future productivity and seek insurance against the possibility that they will turn out to be relatively unproductive, and falls into the latter category.

⁴ That is, it refers to the relation between earnings and tenure holding general experience effects constant (Hashimoto and Raisian, 1985, pp.728).

cases being intuitively most relevant in circumstances in which the job entails team production, or where the job consists of numerous and varied tasks, or in particularly large firms in which there is an arms length relationship between employees and managers. Given such environments, employees have an incentive to put forward less effort, or shirk. Lazear's model is one solution to this underlying agency problem.⁵ By supplying a contract which pays less than the value of marginal product at the start of tenure and more near the end, Lazear argues that the employer will illicit work effort from the employees as exerting low effort runs the risk of dismissal and the loss of the collateral which they put up early in their career. This therefore increases the expected present value of each employee's net contribution to the firm and has a corresponding effect upon the expected present value of each employee's lifetime wages. The model in fact predicts an upward sloping seniority- earnings profile even in the absence of any specific human capital accumulation.⁶ Empirical support for the hypothesis of an upward sloping earnings profile in the absence of increases in productivity was presented in Medoff and Abraham (1980, 1981) and Flabbi and Ichino (2001), though criticism of their approach centres on the inherently difficult task of measuring workers productivity.⁷ However, other studies have found support for the human capital explanation by directly investigating its core hypothesis: namely investments in *training* lead to increases in worker productivity and to increases in wages (see, Bartel, 1995; Brown, 1989; and Barron et al, 1989)

Job search models typically assume that wages can be expressed as a function of a time-invariant match quality parameter which is specific to a particular job-employee match. As such, they emphasise that the source of wage growth is through mobility across jobs, and that this wage growth can occur independently to any productivity enhancing investments within jobs. These models generally fall into two categories: the 'search good' model of job matching which assumes that the quality of job matches is known upon inspection (Burdett, 1978) and the 'experience good' model of Jovanovic (1979) in which information as to the quality of the match is not known for certain ex ante, but updated as the job is 'experienced.' In the former model, the movement of a worker is therefore dependent on whether there exists a higher match quality in alternative employment and whether the gain from moving - due to the higher wage offered in the better match - exceeds the loss from moving from the current employer - a loss which, in the human capital framework, stems from the share of the returns from specific human capital investment which was undertaken at this existing employer. In the latter model, turnover is generated by the revelation of match quality and its relationship to the workers reservation match quality on the current job. Early in the match the reservation match quality will start low due to option value that the true match quality may turn out to be very high. A worker who receives a sequence of poor match quality signals, however, will continually revise down their belief about the true match quality, and as such the

⁵ The most obvious other solution being that the firm devotes additional resources to the supervision of its employees.

⁶ An implicit assumption of the model is that the gain the firm would make from terminating a workers contract after they begin earning more than the value of their marginal product is less than the loss which would be incurred by such systematic unscrupulous behaviour by the firm. This conjecture can be supported if one assumes that a firm's long-run productivity is a function of their reputation to society in general and to future and current employees. However this means that the model is best applied to large, visible firms.

⁷ For papers in which the central objective is to test the human capital explanation against the agency themed explanation, see Hellerstein and Neumark (1995), Lazear and Moore (1984), Hutchens (1987) and Brown and Sessions (2006).

option value declines and the reservation match quality increases and they will choose to leave the firm. Ultimately therefore, the model predicts that the worst matches leave first, followed in succession with marginally better matches. Eventually only the higher quality matches remain.

It is recognition of this match heterogeneity as well as individual heterogeneity, both of which constitute the driving force for voluntary (or involuntary) movement across firms in the labour market, which has stemmed a rather substantial literature which attempts to estimate the returns to tenure. Such heterogeneity implies that a standard earnings function will suffer from omitted variable bias since both match quality and individual quality are unobservable to the econometrician but are correlated with years of tenure. There have been a number of methodological approaches derived to solve this problem. Altonji and Shakotko (1987) used data from the Panel Study of Income Dynamics and instrumented years of tenure with the deviation of tenure from its mean for the sample observation on a given job match— an instrument they argue to be uncorrelated with the unobservable explanatory variables missing from the earnings function. They concluded that the true returns to tenure are minor (10 years of tenure leading to a 6.8 percent increase in earnings in their preferred specification) and that accumulation of general labour market experience (through either general human capital accumulation or labour market search) accounts for the lions share of wage growth over the career. This finding carries the implication that tenure at a firm has a much smaller role in shaping the structure of earnings as was previously thought and in turn categorically undermines the importance of an entire compensation literature in which wages are set as a device for affecting worker productivity. Nevertheless Topel (1991) addressed the issue using essentially the same data, but a different methodological approach and found that with his method there are estimated to be substantial tenure effects upon wages, with a tenure effect of 28 percent in the same period of time (also see, Bratsberg and Terrell, 1998; Williams, 1991; Abraham and Farber, 1987; Mincer and Jovanovic, 1981; and Altonji and Williams, 1997).

To shed light on how to interpret the estimated experience- earnings profiles consider a simple wage determination model of the form:⁸

$$\ln y_{it} = \beta_1 X_{it} + \beta_2 S_{it} + \varepsilon_{it} \quad (1)$$

Equation (1) states that the wage, y , of individual i in firm f in period t is a function of years of pre-firm labour market experience, x , and years of service s with the current employer and a disturbance ε .⁹ Disregarding heterogeneity in match quality or individual quality across employees, then β_1 and β_2 will represent the average returns to pre-firm experience and within-firm experience. Returns to pre-firm experience reflect general human capital accumulated out-with the firm, while returns to within-firm experience reflect general human capital accumulated within the firm as well as the returns stemming from the seniority earnings profile - a profile which may be associated to employees return to specific human capital accumulated

⁸ For notational simplicity, I abstract here from other explanatory variables and from higher order terms in pre firm or within firm experience or controls for secular wage growth with time.

⁹ This model is analogous to the one which appears in Altonji and Shakotko (1987) and Topel (1991). However in their model, years of pre-firm experience is replaced with years of total experience. Their papers were concerned with estimating the returns to tenure and as such, in that scenario, the earnings function must be specified holding total experience constant, in order to capture tenures net effect.

and/or a contractual solution to alleviate the principal-agent problem inherent in the worker-firm relationship.¹⁰ However let us decompose the disturbance term to introduce unobservable individual and match heterogeneity into the model. Let the decomposition take the form of an individual effect (i.e. innate ability) a_i , which affects earnings equally across all firms, a firm-match effect m_{if} , representing the time-invariant productivity associated to a given employee-firm match,¹¹ and a white noise error u_{ijt} :

$$\varepsilon_{ijt} = a_i + m_{if} + u_{ijt} \quad (2)$$

Bias will be prevalent in the estimation of β_1 and β_2 if covariance exists between the regressors in (1) and the unobservables in (2). The preceding discussion already suggested the existence of such covariance with respect to unobservable match quality m_{if} . Therefore let us firstly focus upon the bias generated by the existence of this unobservable by letting its relationship to the observed explanatory variables be written as

$$m_{if} = b_1 X_{it} + b_2 S_{it} + e_{it} \quad (3)$$

Therefore $E \hat{\beta}_1 = \beta_1 + b_1$ and $E \hat{\beta}_2 = \beta_2 + b_2$. The ‘search’ model of Burdett (1978), at least, hypothesises that individuals’ match quality will rise with years of pre-firm experience as individuals have sampled a larger number of job-offers and have sequentially moved to better matches within the labour market before joining this particular firm. This implies that $b_1 > 0$ and the estimation of ‘returns’ to pre-firm experience are upward biased. In fact, in the extreme scenario, the earnings of the given representative worker need not increase *due* to accumulated years of pre-firm experience -that is $\beta_1 = 0$, yet in the cross-section of workers observed, those entering the firm with greater years of pre-firm experience are expected to be of a better match quality and naturally are rewarded with a higher wage upon entry –so $b_1 > 0$ - and the positive estimated return to pre-firm experience is completely spurious. The direction of the bias due to unobservable match quality upon the estimation of returns to within-firm experience, on the other hand, is ambiguous. An employee with long tenure at the firm has chosen not to move firm, so may conceivably be expected to be in a better match compared to an employee with short tenure (Jovanovic, 1979). Aggregating the same scenario across all employees means that we would expect $b_2 > 0$ and returns to within-firm experience are biased upward. On the other hand, an employee with a short observed length of tenure may have left not because they were a bad match, but because they have found a better match elsewhere (Burdett, 1978). Aggregation of this scenario suggests that we would expect $b_2 < 0$ and returns to within-firm experience to be biased downward.¹²

¹⁰ For instance, the solution of incentive problems could foster the accumulation of specific human capital. As such there may be a degree of non-mutual exclusivity between the implications of the competing - agency versus human capital - explanations for the seniority-earnings profile (Brunello and Ariga, 1997))

¹¹ Note that the m_{if} term may also reflect the variation in wages across firms as discussed in the efficiency wage literature (Parsons, 1986).

¹² Topel (1991) notes that significant mobility costs reinforce the $b_2 < 0$ hypothesis, while costly search suggests that $b_2 > 0$ as only those with relatively poor matches will actively search. Stevens (2003) also argues that $b_2 < 0$, as workers with high levels of specific human capital – and hence years of tenure- will stay with the firm even when the match quality is actually low. Therefore in a sample of

In addition to the unobservable matching bias, there is also bias expected to be generated due to the covariance between the unobservable individual effect a_i in equation (2) and years of within-firm experience. Let the relationship be written as:

$$a_i = c t_{if} + \epsilon_{if}. \quad (4)$$

We would expect individuals of higher ability¹³ naturally to earn more and would also expect that individuals of higher ability are less likely to be dismissed by the firm. Therefore it is expected that $c > 0$. This implies that we will observe individuals with high years of within firm experience being better paid relative to those with low years of within firm experience, though this differential will not be exclusively due to these high tenured individuals being more productive through human capital investment, or through reaping the benefits of some delayed payment incentive in operation in the firm.

This conceptual earnings determination model has highlighted the fact that the estimated experience-earnings profiles themselves cannot be given a strict a priori interpretation. They will reveal how wages vary with years of pre-firm and within-firm experience, though this estimated variation cannot be attributed exclusively as 'returns' to pre-firm or within firm experience. Nevertheless, the merit of the estimates is that they will naturally serve as a useful descriptive summary of the structure of compensation and its relationship with years of previous and within-firm experience of the individuals who are employed in the firm. The nature of the estimated wage evolution can then be reconciled with the theory which was imbedded in the model presented above.

The rest of the paper is structured as follows. Section 2 presents the semiparametric model and describes the nonparametric estimation technique. Section 3 discusses the data used and provides some preliminary statistical analysis of the dataset. Under the assumption that the grade in the hierarchy occupied by an employee is indicative to some extent of their individual ability (a_i) and match quality (m_{if}) some incite can be gauged in support of the hypothesis presented above and this is shown to be helpful in interpreting the within-firm experience earnings profiles, the estimates of which are presented and discussed in Section 4. Section 5 concludes with a summary of what has been achieved and a discussion.

2. Nonparametric Modelling and Estimation of the Earnings Function

2.1. Semiparametric Model

Let the earnings function be given by

accepted wages, specific human is negatively correlated with match quality. Lazear's (1986) raiding model also has input into the direction of bias debate. He argues that employee's wages are used by the firms' competitors as signals of their quality and predicts that the best workers are 'stolen away' so that the remaining workers tend to be relatively less able than their departing peers. This model implies therefore that the average returns to within-firm experience are biased downward and that the bias is upward for returns to pre-firm experience.

¹³ The unobservable individual effect could also be viewed as an index of an employee's stability as opposed to an index of innate ability. More stable employees naturally tend to have longer years of tenure will be - for instance due to turnover and training costs - of greater value to a firm.

$$\ln y_i = f(x_i, a_i) + \varepsilon_i \quad (i = 1, \dots, n), \quad (5)$$

where y_i represents the wage of individual i , x_i is a (1×2) vector with elements representing years of experience prior to joining the firm and years of tenure with the employer. a_i is $(1 \times k)$ vector of discrete variables further charactering the given individual. I assume that $E(\varepsilon_i / x_i, a_i) = 0$.

Firstly I note that I choose to estimate the earnings function: a) across two schooling groups; and b) for male and female employees separately. Therefore, more specifically, the earnings function (5) relates to an individual i in schooling groups, with gender g . The choice to remove schooling as an explanatory variable and to estimate the earnings function across two schooling groups – where the group an individual is placed in is conditional upon their highest educational qualification attained before entry into the labour market – is largely driven by a priori uncertainty as to the appropriate entry of schooling in the earnings function. The conventional linear entry of years of schooling - which dictates the assumption that each year of education, no matter at what level, has the same proportional effect on earnings - has been widely challenged in more recent literature.¹⁴ In addition to this, the additively separable entry of schooling – which stems from Mincer's (1974) assumption that the proportion of earnings devoted to the production of human capital is the same regardless of years of schooling attained – has also been challenged. For instance, individuals with low learning ability may naturally have less schooling and flatter earnings profiles (see, Lemieux, 2006; Heckman et al, 2003; Schady, 2003; and Hungerford and Solon, 1987 for evidence and discussion of these issues). Eliminating the entry of schooling as an explanatory variable is appealing the context and objective of this paper for it naturally leaves open the possibility of marginal differences across schooling groups in the experience profiles.

The choice to separate males and females and to estimate the earnings functions separately for both sexes is conventional. Firstly, years of pre-firm experience is not explicitly identified in the data set and is therefore proxied with a 'potential' pre-firm experience index, given the available data of age, years of service, and (proxied) years of schooling of each individual employee. As such, there is expected to be larger measurement error in the pre-firm experience variable for women compared to men (see, Bratsberg and Terrell, 1998; Light and Ureta, 1995; and Light, 1998 for demonstration of the sensitivity of estimates with regard to the assumptions made with respect to the construction of proxied measures of experience). Secondly, there exists strong logical theoretic justification for the sex separation. Under the human capital explanation for wage growth, the incentive to invest in ones-self – or equivalently the firm's incentive to invest in an individual- is a function of expectations with regard to the continuity of labour force attachment. As such, there is less incentive for women -in the aggregate- to invest, given that they are more likely to suffer an interruption to their career. There is also the well documented

¹⁴ There are two potential theoretical sources of misspecification with regard to the linear entry of years of schooling. Firstly, the mean earnings premium may be higher for some levels of education compared to others. For example the earnings premium from university education may be higher than secondary if there are differences in the cost or quality of education or if there are changes in market supply and demand across these two levels. Secondly, the earnings premium from the last year of schooling within a particular level may be higher than in any other intervening year. See (Schady, 2003, pp.191) for discussion of three possible reasons for the existence of such credentialism or sheepskin effects.

rebound effect of human capital accumulation and wage growth which occurs upon re-entry into the labour market as skills depreciated during the career interruption are restored (see, Mincer and Polachek, 1974 and Mincer and Ofek, 1980).

The question which remains is how to specify $f(\cdot)$? As already discussed, there has been a body of existing work expressing dissatisfaction with the quadratic entry of years of total labour market experience in the conventional parametric specification of the earnings function. Clearly then, the a priori uncertainty regarding the adequacy of the quadratic specification of *total* experience in the earnings function leads directly to an uncertainty regarding appropriate specification of *pre-firm* and *within-firm* experience - the two elements of total experience which I explicitly differentiate between. It is this uncertainty which motivates the nonparametric approach taken in this paper. In the nonparametric setting, the object of estimation is not a finite number of parameters in a model, but the regression function $f(\cdot)$ itself. While parametric specification of the experience variables in the earnings function will constrain the estimates of the earnings profiles to follow a specified and strict structure governed by a finite number of parameters, nonparametric estimation methods- apart from assuming the average value of the response is a smooth and continuous function of the predictors - impart no assumptions as to the functional form. As such, the uncertainty over the exact (parametric) specification to adopt is lost.

However the added flexibility of nonparametric regression comes at a cost. As there are five predictors in the earnings function, a fully nonparametric specification is infeasible for the rate of convergence of the nonparametric estimator slows down dramatically as the dimension of the model (number of predictors) increases: the so called ‘curse of dimensionality.’ Fully nonparametrically modelling of the earnings function would lead to a requirement of an impractically large sample in order to obtain estimates of acceptable levels of precision.¹⁵ In response to this, some modelling restrictions need to be imposed at the outset. Given that the motivation of this study is with respect to the estimation of the experience-earnings profiles, I choose to let years of pre-firm experience and years of within-firm experience enter the model nonparametrically, while I let the remaining (discrete) explanatory variables enter in a parametric fashion. Specifically, I augment (5) and model the earnings function as a partially-linear semiparametric model given by,¹⁶

$$\ln y_i = f(x_i) + a_i\beta + \varepsilon_i \quad (i=1, \dots, n), \quad (6)$$

where a_i is $(1 \times k)$ vector of discrete explanatory variables and β is a $(k \times 1)$ vector of parameters to be estimated. x_i is a (1×2) vector including the two forms of labour force experience of each individual i : experience prior to entry and tenure. The function $f(\cdot)$ remains unspecified. The disturbance term satisfies $E(\varepsilon_i / x_i, a_i) = 0$.

¹⁵ As well as this theoretical reason, there are also practical reasons which essentially preclude fully nonparametric modelling in this case. As the number of predictors in a nonparametric model expands, the interpretability and presentation of the estimation results becomes increasingly cumbersome. Fully nonparametric modelling of the earnings function would mean that I could only present an estimated within firm experience-earnings profile which was conditional on given values of all the other predictors held constant. Both reasons therefore render the use of a fully unrestricted nonparametric regression undesirable.

¹⁶ Other papers in the economics literature to have specified a partially-linear semiparametric model then to have followed the estimation method of Robinson (1988) include Anglin and Gencay (1996) and Schmalensee and Stoker (1999).

Robinson (1988) shows that this model can be rewritten as¹⁷

$$\ln y_i - E(\ln y_i | x_i) = (a_i - E(a_i | x_i))\beta + \varepsilon_i. \quad (7)$$

Therefore this suggests a two-step method in the estimation of the model parameters β . First, the conditional means $E(\ln y_i | x_i)$ and $E(a_i | x_i)$ are estimated by some method of nonparametric estimation. Then, second, these estimates are substituted in place of the unknown functions in (7), and then β is estimated by ordinary least squares. The resulting parameter estimates are $n^{1/2}$ consistent and asymptotically normal.

Then, since $f(x_i) = E(\ln y_i - a_i\beta | x_i)$, and given the parameter estimates $\hat{\beta}$ we can estimate the regression function $f(x_i)$ by a nonparametric regression of

$$\ln y_i - a_i\hat{\beta} \text{ on } x_i. \text{ Or, equivalently, the same estimate of } f(x_i) \text{ is given by}$$

$$\hat{E}(\ln y_i | x_i) - \hat{E}(a_i | x_i)\hat{\beta}.$$

This methodology is therefore somewhat of an intermediate strategy between the fully nonparametric and parametric approach; combining the high flexibility and robustness of the nonparametric approach and the faster convergence rate obtained in the parametric one. The model is more restrictive than the fully nonparametric as it rules out interaction between the measures of experience and any of the discrete variables in the determination of wages, but remains less restrictive than a parametric specification which brings some form of straightjacket to the entry of within firm experience and pre-firm experience (and their interaction) in the modelling of their effects on wages. However this modelling strategy has the advantage of alleviating the curse of dimensionality associated with fully nonparametric estimation and lends for easier analysis of the experience profiles- since the experience profiles are estimated with the impact of the other predictors held constant.

2.2. Nonparametric estimation and smoothing parameter selection

As well as the class of nonparametric model to estimate, there are two other important choices to be made with respect to nonparametric estimation: a) the nonparametric method (or smoother) to undertake the estimation; and b) the type of smoothing parameter (or bandwidth) and the method of choosing this smoothing parameter. With regard to the first I choose to apply local linear regression and with regard to the second, I choose to use nearest neighbour bandwidths, with the bandwidths for all the nonparametric regressions chosen by generalized cross validation (Craven and Wahba, 1979).

Local linear regression with variable bandwidth - as a nonparametric regression technique - was developed for the case of a single independent variable in Cleveland (1979) and then extended to the multivariate case in Cleveland and Devlin (1988).¹⁸ Utilisation of this method is a departure from standard approach in the applied economics literature, where the dominant estimator utilised remains the kernel or local constant estimator with fixed bandwidth (developed by Nadaraya (1964) and

¹⁷ Given that $E(\ln y_i | x_i, a_i) = f(x_i) + E(a_i | x_i)\beta$

¹⁸ The authors also present several interesting case studies in which local regression is seen to be more insightful than classical linear regression.

Watson (1964)).¹⁹ The motivation for this departure comes from Fan and Gijbels (1992) and Ruppert and Wand (1994)²⁰ who illustrate asymptotic properties of the local linear estimator with variable bandwidth and demonstrate its superiority over the kernel estimator. Specifically, they illustrate that the bias and variance near the boundary of the local linear estimator are of the same order of magnitude as the interior, whilst local mean estimation suffers boundary bias near the edges of the region over which the data have been collected.

To intuitively describe the nonparametric methodology utilised here, consider the fully nonparametric model:²¹

$$y_i = f(x_i) + \varepsilon_i, \quad (8)$$

where y_i is a quantitative dependent variable and $x_i = (x_{i1}, x_{i2})$ is a vector of two quantitative explanatory variables. $E(\varepsilon_i / x_i) = 0$. Given the data $\{y_i, x_i\}$, $i = 1, \dots, n$, the aim of nonparametric estimation is to estimate $f(x_i) = E(y_i | x_i)$ - the conditional mean of y_i given x_i - without the specification of any parametric form for $f(x_i)$.

Consider the plot of observations y_i against the observations x_i . Though there is likely to be a lot of noise in such a plot, the shape of the plot may still roughly reveal how the mean of y_i changes with x_i . Nonparametric estimation proceeds by evaluating the data at a series of focal points $x_0 = (x_{01}, x_{02})$ - which are typically taken to be the observations x_i themselves. The estimation of the regression function then takes place at each focal point x_0 . Central to the methodology is the intuitively sensible assertion that the observations at or around a given focal point are more informative to the conditional mean relationship at that particular point than those observations which are distant. Given this, estimation proceeds by defining a 'neighbourhood' around each focal point and then 'averaging' the response values of the observations which fall in each neighbourhood. This clearly then leads to two logical questions: 1) how do we 'average' response values in a given neighbourhood? 2) How do we define the given 'neighbourhood' and how large do we make it?²²

For each focal point x_0 , let us, for now, arbitrarily define a bandwidth $h(x_0)$ and a corresponding smoothing window, which captures the neighbourhood around each focal point $((x_0 - h(x_0), x_0 + h(x_0)))$. Let $z_i = \frac{x_i - x_0}{h(x_0)}$ give the scaled distance between the predictor value for the i^{th} observation and the focal value. We then need a kernel (weight) function that assigns greater weight to those observations that are closer to x_0 . I choose the tricube weight function proposed by Cleveland (1979), which is given by $w(z_i) = (1 - |z_i|^3)^3$ if $|z_i| < 1$, or 0 if otherwise. Therefore observations x_i falling outside the smoothing window are given zero weight, while the weight given to observations falling inside the window falls off symmetrically and

¹⁹ Among various other nonparametric methods (or smoothers) there includes spline smoothers and orthogonal series. Simonoff (1996), Hastie and Tibshirani (1990), Hardle (1990) and Hardle et al (2004) all outline these and various other methods.

²⁰ Fan and Gijbels (1992) look at uni-variate local regression while Ruppert and Wand (1994) extend to multivariate.

²¹ Some excellent review papers of nonparametric regression methods include Cleveland and Loader (1996), DiNardo and Tobias (2001), Yatchew (1998) and Blundell and Duncan (1998).

²² This intuitive description implicitly assumed that the two explanatory variables were measured in the same metric. The question of how to define a neighbourhood around x_0 when predictors are in different scales is explored in Fox (2000, pp.15-17).

smoothly as $|x_i - x_0|$ grows.²³ Given the choice of weight function, we now perform a weighted linear least squares regression at the given focal point x_0 to minimise the weighted sum of square residuals:

$$\min_{a_0, b_0, b_1} \sum_{i=1}^n (y_i - (a_0 + b_0(x_i - x_{01}) - b_1(x_i - x_{02})))^2 w(z_i). \quad (9)$$

The local linear estimate at x_0 is $\hat{f}(x_0) = \hat{a}_0$ as this defines the position of the local regression line at the point x_0 . The method proceeds in the same manor for each focal point x_0 and the fitted values are connected. Higher order polynomials can of course be utilised and can be shown to decrease the bias of the estimate, while increasing the variance. There is theoretical advantage to odd ordered polynomials. Namely, the polynomial of odd order $p + 1$ has the same asymptotic variance as the polynomial of even order p , but lower bias (see, Fan and Gijbels, 1996 and Simonoff, 1996). Nevertheless it is generally agreed that the linear approach is sufficiently flexible for most purposes.²⁴

A constant bandwidth which looks a fixed distance to the left and right of each focal point is clearly the simplest specification. However this can potentially lead to empty neighbourhoods when the independent variables have non-uniform distributions. However data sparsity problems can be reduced by ensuring the neighbourhoods contain sufficient data. Nearest neighbour bandwidths choose $h(x_0)$ so that the local neighbourhood always contains a specified number of points. Therefore the distance of the bandwidth conforms to the density of the data around that particular focal point. Specifically, nearest neighbour bandwidths involve choosing a fixed number k of the nearest observations around each focal value. Therefore the neighbourhoods are set according to the equality $k = n\alpha$, where n is the total number of observations and $\alpha \in (0, 1)$ is the fixed ratio of nearest observations relative to total observations to be used as the neighbourhood for estimation at each focal point, also called the span. The degree of smoothing applied to the data can therefore be expressed by the span of the estimator.

The question that remains is how large to make the span of the local linear estimator? The expressions for the mean and variance of the multivariate local linear estimate at a focal point is given in Ruppert and Wand (1994), where it is shown as the smoothing parameter increases the bias increases, but conversely the variance decreases. As such the chosen smoothing parameter is of central importance for it

²³ There are many other weighting functions with the same properties (see, DiNardo and Tobias (2001), table 1). Though it has been generally established that the particular choice of weighting function does not seem to matter a great deal (Ullah, 1988, pp.643)

²⁴ Local constant least squares regression attempts to minimise the weighted sum of square residuals of the following criterion:

$$\sum_{i=1}^n w(z_i)(y_i - a_0)^2.$$

The minimizer of which is obviously $\hat{f}(x_0) = \hat{a}_0 = \frac{\sum_{i=1}^n w(z_i)y_i}{\sum_{i=1}^n w(z_i)}$. This is the popular kernel estimator

proposed by Nadaraya (1964) and Watson (1964).

dictates a statistical compromise between bias and variance. To assess the overall performance of the fit however we need to create a measure of performance which represents the global behaviour of the estimator. One such measure is the prediction mean-squared error for future observations:

$$PMSE_h = \frac{1}{n} \sum_{i=1}^n E[y_i^* - \hat{f}_h(x_i)]^2, \quad (10)$$

where y_i^* is a new observation at x_i and $\hat{f}_h(x_i)$ is its local linear estimate conditional upon the nearest neighbour bandwidth h . This motivates a cross-validated estimate of the prediction mean-squared error, given by

$$CV_h = \frac{1}{n} \sum_{i=1}^n [y_i^* - \hat{f}_{-i,h}(x_i)]^2, \quad (11)$$

where $\hat{f}_{-i,h}(x_i)$ denotes the local linear estimate at x_i - conditional upon nearest neighbour bandwidth h - computed by leaving out the i^{th} data point.²⁵ The optimal smoothing parameter can then be found iteratively, with the chosen parameter being the one which minimises the cross validation score. Justification for this selection procedure is that $E[CV_h] \approx PMSE_h$ (Hastie and Tibshirani, 1990).

As local regression solves a least squares problem, the fitted values \hat{f}_i can be expressed as a weighed sum of the observed y_i values and there exists an $(n \times n)$ smoother matrix \mathbf{S} , which maps the data to the fitted values:

$$\hat{\mathbf{f}}_h = \mathbf{S}\mathbf{y} \quad (12)$$

where $\hat{\mathbf{f}}_h = (\hat{f}_h(x_1), \dots, \hat{f}_h(x_n))'$ is the column vector of fitted values and $\mathbf{y} = (y_1, \dots, y_n)'$ the column vector of observed response values. Each row of \mathbf{S} consists of the weights appropriate to estimation at each x_i and the diagonal elements of the smoother matrix measure the sensitivity of the fitted curve \hat{f}_h to the individual data points x_i - the leverage values of each observation.

It can be shown that CV_h can be written as a function of these fitted values,

$$CV_h = \frac{1}{n} \sum_{i=1}^n \left[\frac{y_i - \hat{f}_h(x_i)}{1 - s_{i,h}} \right]^2 \quad (13)$$

where $s_{i,h}$ is the i^{th} diagonal element of the smoother matrix.

²⁵ This is also known as the jackknifed fit at x_i . Note that if the i^{th} observation was not omitted then the smoothing parameter which minimises the cross validation measure will be zero and the estimator will simply interpolate the observed data $\hat{f}_h(x_i) = y_i$: the fitted and observed values will be equal (if x_i values are distinct).

Degrees of freedom for the nonparametric regression model can therefore be defined in a way with analogy to linear least squares regression.²⁶ That is, the *approximate* number of degrees-of-freedom is the number of parameters is the trace of the smoother matrix:

$$DF = \text{tr}(\mathbf{S}). \quad (14)$$

This motivates the generalized cross validation criterion for smoothing parameter selection. This criterion provides an approximation to cross validation and is easier to compute for it replaces each value $1 - s_{iih}$ with their average

$1 - \left(\frac{\text{tr}(\mathbf{S})}{n}\right)^2$. Therefore the two criteria are similar except the generalised cross validation approach effectively down-weights, or penalizes the effects of high leverage points (Simonoff, 1996). Hence the generalised cross validation score for the nonparametric estimate is given by:

$$GCV_h = \frac{1}{n} \sum_{i=1}^n \frac{[y_i - \hat{f}_h(x_i)]^2}{\left[1 - \left(\frac{\text{tr}(\mathbf{S})}{n}\right)^2\right]}. \quad (15)$$

Minimisation of the criterion therefore involves a trade-off between the goodness of fit of the model and model parsimony. Model parsimony being expressed as a function of the equivalent degrees of freedom $\text{tr}(\mathbf{S})$. (See, Hastie and Tibshirani, 1990 and Loader, 1999a for a discussion and motivation for the use of this method)

These automatic or classical methods for choosing the smoothing parameter have come under attack in recent literature as their performance in practice is sometimes questionable. An alternative approach which has been proposed is to construct a ‘plug-in’ estimator of the optimal smoothing parameter from solving to

minimise the average mean square error of the estimator $\frac{1}{n} \sum_{i=1}^n E[\hat{f}(x_i) - f(x_i)]^2$ and

replacing the unknown parameters in such an expression with estimates. Fan and Gijbels (1995, 1996) illustrate this approach for local linear regression. However, the classical versus plug-in smoothing parameter selection debate is still somewhat in its infancy and Loader (1999b) surveys the issue and existing evidence at some length and argues that the variability of the classical approaches is not a problem, but simply an underlying symptom of the difficulty of smoothing parameter choice.

In the first stage of the Robinson (1988) estimation approach outlined in the previous section the method of local linear logistic regression - a method which belongs to a much wider class of generalized nonparametric regression models - is used to nonparametrically estimate the conditional means $E(a_i | x_i)$ when the dependent variable is binary. The nonparametric augmentation to generalised linear models is conceptually identical as that to parametric ordinary linear models. Namely, the nonparametric models retain the random component and link function of the generalized linear model, but the predictors are entered as a smooth function of the response. Fan et al (1995) analyse nonparametric estimates in the context of local

²⁶ That is, in a standard linear model of the form $y_i = X_i\beta + \varepsilon_i$, the matrix \mathbf{S} is analogous to the projection matrix or hat matrix $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ where $\mathbf{X} = (\mathbf{X}'_1, \dots, \mathbf{X}'_k)$ and the degrees of freedom of the model is equal to the number of parameters, given by $\text{tr}(\mathbf{H})$.

generalised linear models and derive expressions for bias and variance in that setting. They show that the attractive properties of the local polynomial approach as opposed to the traditional kernel approach also carry over in generalized nonparametric models.

3. Data Analysis

A cross-section of data is drawn from the personnel records of a large major UK based financial sector employer. The data covers all UK based employees and gives a snapshot of their status on the 1st December 1999. I choose to focus upon full-time employees only, with full-time being defined as a working week of 30 or more hours.²⁷ As the birth dates of all employees are available, the age of employees at this point in time is calculated straightforwardly. Likewise, as an exact entry date into the firm is available, length of within-firm experience (or service) is also measured exactly. Years of formal schooling are not explicitly listed in the data set but, instead information regarding the highest qualification upon entry into the firm is available. Those employees for which this information is not available are excluded. This information is then used to construct a proxy for years of experience in the labour market attained out-with the firm. I let the achievement of a degree translate to 16 years of formal schooling; further education to 14 years; A-level (or equivalent) to 13 years and GCSE (or equivalent) to 11 years. I then use this mapping to proxy years of pre-firm experience by age - years of service-years of schooling-5. Therefore measurement error in the proxied years of pre-firm experience variable will stem from both the assumptions made by the mapping of years of schooling and from time spent out-with the firm in which the employees were not in full time employment. Instead of entering a set of schooling dummies as explanatory variables in the earnings function, I choose to estimate the earnings function across two schooling groups. I calculate a (gross) hourly wage for each individual using information on weekly contracted hours and by summing annual salary, annual allowance, 'London' allowance, and an annual company profit related bonus which was received on the 1st of May 1999. Over 60 percent of the staff observed received a bonus. The dependent variable is the natural logarithm of this calculated hourly wage. In addition to these key variables of interest I generate three dummy variables further characterising each individual. They relate to region of employment, marital status, and whether the employee has children.²⁸

The earnings function is therefore estimated across four cells encompassing the two genders and two schooling groups which are: 1) Higher education: the set of employees who have achieved a degree or who have attained some form of further education prior to entry into the labour market; and 2) secondary education: the employees who have achieved up to secondary-school level qualifications prior to entry. I choose to specify only two schooling groups – even though the data permits up to four- in order to increase the number of observations falling in each gender/schooling-level cell. In total there are 21,702 observations. Table 1 provides a summary of the data.

²⁷ The motivation for doing so is to eliminate labour supply effects upon the determination of earnings, so that hours worked need not appear as an explanatory variable in the earnings function.

²⁸ I note that these variables may be considered as endogenous and consequently the estimated effects of these variables upon wages may be biased.

Table 1
Summary of Data

	Mean	St.dev	Min	Max	Mean	St.dev	Min	Max
Males/ Secondary Education ^a				Males/Higher Education ^b				
Inhrwage	2.47	.53	1.40	5.41	2.93	.62	1.41	5.29
pexpernet	7.27	7.84	0	57.27	5.85	6.32	0	48.66
service	8.37	7.28	0	41.91	9.07	8.60	0	39.25
ms	.44	.50	0	1	.52	.50	0	1
child	.32	.49	0	1	.40	.49	0	1
region	.38	.49	0	1	.48	.50	0	1
Females/ Secondary Education ^c				Females/Higher Education ^d				
Inhrwage	2.17	.31	1.40	4.03	2.49	.52	1.41	4.56
pexpernet	7.32	7.70	0	45.60	4.68	5.43	0	33.93
service	7.39	6.23	0	37.30	6.74	6.84	0	36.28
ms	.42	.49	0	1	.39	.49	0	1
child	.24	.42	0	1	.18	.39	0	1
region	.23	.42	0	1	.36	.48	0	1

note: the number of observations in each cell is: a) 5,584; b) 5,115; c) 8,318; d) 2,685

The variables in table 1 are defined as follows:

Inhrwage is the natural logarithm of the hourly wage

pexpernet is the proxied number of years of pre-firm labour market experience

service is years of within-firm labour market experience (or tenure)

ms= 1 if the employee is married

child=1 if the employee has at least one child

region=1 if the employee works within greater London

The summary statistics presented in table 1 somewhat cloud the distributions of the key variables. Therefore I present in figure 1 the nonparametrically estimated univariate density functions for Inhrwage, pexpernet and service using the all the data. In the interest of parsimony, the plots for each of the four cells are not presented.

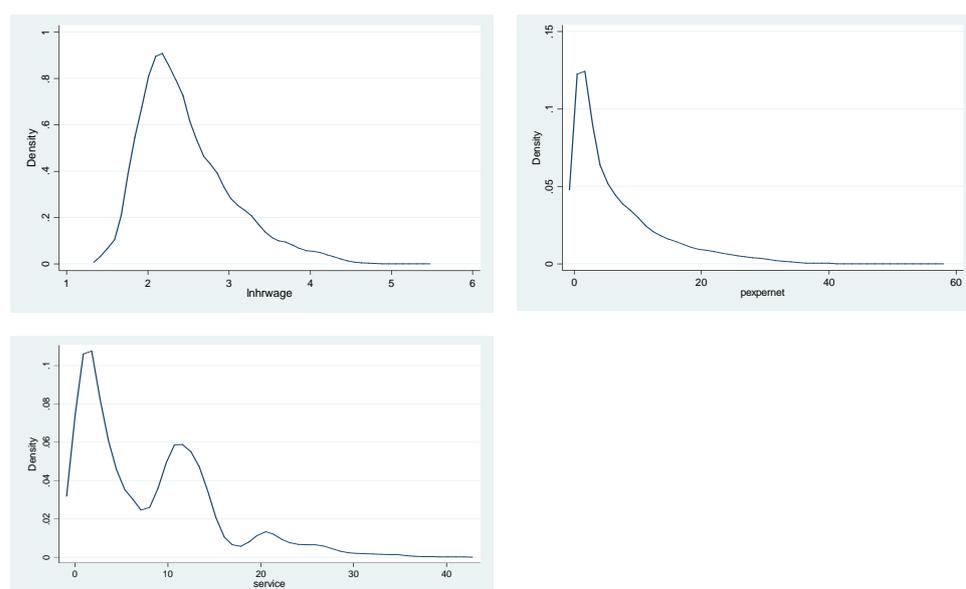


Figure 1. Kernel Density Estimates of Inhrwage, pexpernet and service

Figure 1 shows that the log earnings density function is positively skewed. Those at the top of the earnings distribution are earning considerably more than the

rest. The employees in the firm are predominantly hired when relatively new to the labour market. In fact over 20 percent of the observed workers have less than 1 year of labour market experience out-with the firm, over 56 percent having less than 10, and nearly 93 percent of the employees have less than 20. Just fewer than 16 percent of staff have less than one year of service, over 61 percent having less than ten and over 91 percent of the employees have less than 20 years of service.

The firm has an explicit hierarchical structure in which workers can be assigned to one of 14 grades.²⁹ Two of the grades seem not to be a part of the hierarchy, but are unclassified states. Excluding these grades, the rest of the hierarchical structure is simple and can be divided into three or four significant levels, consisting of training grades T1-T3, clerical grades C1 – C2, middle management grades M1-M2 and senior management grades SM1-SM5. In table 2, I present the relationship between the grades in the hierarchy, the number of workers in each grade, the average years of both pre-firm and within-firm experience of the workers observed in each grade as well as the average hourly wage within each grade. I do this for each of the four gender/schooling group classes of employed workers. In the aggregate, 19 percent of all the employees are in training grades, 36 percent in clerical, 24 percent in middle management, 11 percent in senior management, the residual being unclassified. Such aggregation however clouds the observed heterogeneity across the gender/schooling classes. Of male employees with a higher level of education, 67 percent are in either middle or senior management grades. For males with a secondary level of education, the ratio is 37 percent, and the ratio for females with higher education and those with a secondary education is 42 percent and 12 percent respectively. Moreover, with regard to senior management grades, over 85 percent of employees in this category are male, with 65 percent of these males having a higher level of education. The conclusion to be drawn from these summary statistics is that, regardless of gender, it appears that the accumulation of a higher level of education results in a much higher probability of being employed in a managerial grade and that male employees tend to fair much better than their female colleagues in terms of placement into higher parts of the hierarchy.³⁰ Such heterogeneity with regard to placement in the hierarchy across schooling groups and gender adds further justification in the decision to separate the data across the four cells in the estimation of the earnings function.

²⁹ Treble et al (2001) analyse this firm's personnel data from January 1989 to March 1997 and present a detailed descriptive analysis of the structure of this firm's hierarchy, promotion policies, the ports of entry and exit, pay policies and their relationship with the hierarchical levels.

³⁰ Lazear and Rosen (1990) argue that firms will set a higher promotion standard for women relative to men, as job leaving among those who are promoted imposes a cost on the firm and that women's greater non-market abilities and opportunities lead to a higher likelihood of departure.

Table 2
Structure of Firm Hierarchy and Relationship between Years of Experience and Grade in Hierarchy

a. Males/Secondary Education						b. Males/Higher Education					
Grade	Number of Workers	Percent	Mean of Pexpernet	Mean of Service	Mean of Hourly Wage	Grade	Number of Workers	Percent	Mean of Pexpernet	Mean of Service	Mean of Hourly Wage
T1	8	0.14	2.51 (.48)	.75 (.45)	5.19 (.70)	T1	1	0.02	27.72 (.00)	2.84 (.00)	5.77 (.00)
T2	180	3.22	6.02 (.56)	1.52 (.27)	6.35 (.18)	T2	21	0.41	3.90 (1.21)	2.09 (.58)	7.80 (.86)
T3	885	15.85	6.75 (.23)	2.61 (.12)	6.95 (.04)	T3	247	4.83	4.14 (.36)	1.96 (.21)	6.68 (.07)
C1	1,129	20.22	6.61 (.21)	6.66 (.16)	9.31 (.08)	C1	368	7.19	4.47 (.27)	4.57 (.26)	9.47 (.21)
C2	824	14.76	5.87 (.25)	9.40 (.22)	11.01 (.08)	C2	437	8.54	3.74 (.23)	5.55 (.34)	10.64 (.08)
M1	959	17.17	6.42 (.24)	10.46 (.20)	14.50 (.11)	M1	905	17.69	5.19 (.21)	7.80(.25)	14.81 (.11)
M2	609	10.91	8.23 (.36)	12.52 (.30)	19.63 (.12)	M2	991	19.37	5.58 (.20)	11.13 (.28)	20.79 (.16)
SM1	331	5.93	10.29 (.52)	14.11 (.49)	29.37 (.45)	SM1	797	15.58	5.60 (.22)	13.50 (.32)	29.48 (.27)
SM2	119	2.13	12.22 (.81)	11.94 (.88)	48.07 (1.09)	SM2	491	9.60	8.67 (.34)	12.01 (.39)	46.21 (.58)
SM3	33	0.59	13.17 (2.14)	15.29 (1.71)	66.13 (2.48)	SM3	209	4.09	9.93 (.48)	12.74 (.61)	62.83 (1.08)
SM4	3	0.05	7.59 (5.21)	23.32 (4.47)	97.63 (14.56)	SM4	31	0.61	10.88 (1.48)	14.11 (1.63)	93.92 (4.94)
SM5	-	-	-	-	-	SM5	1	0.02	6.86 (.00)	20.34 (.00)	120.88 (.00)
	5,080	90.97					4,499	87.96			

c. Females/Secondary Education						d. Females/Higher Education					
Grade	Number of Workers	Percent	Mean of Pexpernet	Mean of Service	Mean of Hourly Wage	Grade	Number of Workers	Percent	Mean of Pexpernet	Mean of Service	Mean of Hourly Wage
T1	8	0.10	4.46 (3.43)	1.56 (1.18)	5.26 (.56)	T1	-	-	-	-	-
T2	171	2.06	4.96 (.56)	.98 (.18)	5.54 (.10)	T2	13	0.48	7.20 (2.54)	1.17 (.46)	6.19 (.18)
T3	2,207	26.53	7.57 (.16)	3.18 (.09)	6.86 (.02)	T3	364	13.56	3.80 (.22)	1.88 (.13)	6.69 (.08)
C1	2,755	33.12	7.17 (.15)	8.41 (.11)	8.56 (.03)	C1	448	16.69	4.09 (.23)	5.38 (.25)	8.43 (.09)
C2	1,446	17.38	6.54 (.19)	10.47 (.15)	10.54 (.04)	C2	370	13.78	4.03 (.28)	7.77 (.38)	10.60 (.11)
M1	758	9.11	6.05 (.26)	11.12 (.19)	13.24 (.08)	M1	502	18.70	3.912 (.23)	8.82 (.33)	13.41 (.11)
M2	209	2.51	7.95 (.56)	11.78 (.48)	18.11 (.26)	M2	357	13.30	5.04 (.30)	9.78 (.39)	19.50 (.25)
SM1	58	0.70	8.48 (1.05)	13.97 (.98)	25.94 (.76)	SM1	171	6.37	6.03 (.43)	9.44 (.57)	29.57 (.64)
SM2	9	0.11	9.72 (2.44)	12.58 (2.45)	43.46 (1.94)	SM2	83	3.09	7.89 (.81)	10.11 (.86)	43.87 (1.36)
SM3	2	0.02	8.74 (2.80)	22.47 (2.91)	41.45 (4.19)	SM3	15	0.56	12.03 (1.53)	9.53 (1.31)	56.13 (4.09)
SM4	-	-	-	-	-	SM4	3	0.11	10.97 (2.03)	10.91 (1.73)	77.85 (3.38)
SM5	-	-	-	-	-	SM5	-	-	-	-	-
	7,623	91.64					2,326	86.63			

Note: Grade in the hierarchy begins with lowest and ends with the highest. T1-T3 are training grades, C1 – C2 are clerical grades, M1-M2 are middle management grades and SM1-SM5 are senior management grades. Workers who are in unclassified grades in the hierarchy are not represented in the table. Pexpernet refers to the proxied number of years of pre-firm labour market experience, while Service refers to years of within-firm labour market experience (or tenure)

The nature of the data utilised here presents a valuable opportunity to assess the relationship between expected match-quality and individual quality with years of pre-firm and within-firm experience, as modelled in equations 3 and 4. It would be expected that the nature of the jobs required skill and responsibility increases as we move up the hierarchy. Consequently ‘better’ workers should be assigned by the firm to jobs which are higher up the hierarchy. Therefore the observation of the position filled by an employee in the firm hierarchy is arguably informative to some extent of both of their quality and of the quality of the match.³¹ It was previously hypothesised that the expected value of the employees’ match quality should be an increasing function of years of pre-firm experience as individuals with more previous labour market experience have sampled a larger number of job-offers and have sequentially moved to better matches within the labour market before joining this particular firm (Burdett, 1978). Column 4 of table 2 therefore suggestively indicates this is true as average years of previous labour market for the employees within each grade of the hierarchy increases as the level of the hierarchy increases. To gain a better perspective on this issue, I present in table 3, the distribution of employees in each grade, conditional upon the number of years of pre-firm labour market experience the employees have before entry and then also conditional upon the number of years of experience they have within the firm.³²

The table indicates that, across all four cells, employees entering the firm with a greater number of years of pre-firm experience tend to be increasing relatively more

³¹ Brunello and Ariga (1997) argue the same point.

³² Presenting the distribution of grade class jointly conditional on given values of both the covariates would generate problems with presentation. The alert reader may point out, however, that in general a worker with a worker with long tenure will tend to be expected to have shorter years of pre-firm experience, compared to a given worker with an average spell of tenure.

concentrated in senior management grades. For example, for the male employees with a higher level of education who enter the firm with less than one year of pre-firm experience, less than 29 percent of them are seen to operate within senior management grades in the firm. However for those enter the firm with between 15 to 20 years of pre-firm experience nearly 50 percent occupy senior management. As such, it does appear that expected match quality of an employee is greater – if grade in the hierarchy is a valid indicator of match quality- the greater the years of previous labour market experience that employee has.

The trend between years of service and the grade in the hierarchy is far more pronounced. The highest proportion of the employees with less than one year of service is seen to operate in the training grades. However this proportion alters smoothly in favour of higher grades as we consider the groups of employees with increasingly higher observed years of service. This is initially suggestive that expected individual or match quality (two time-invariant characteristics) is increasing with years of service for it was hypothesised that ‘better’ workers should be assigned to higher levels in the hierarchy. However this suggestion ignores the dynamic nature of the firms’ learning about individuals’ quality or their match quality, as suggested in Jovanovic’s (1979) ‘experience’ good model of matching and mobility. To explain, equations (3) and (4) indicated that each individuals actual firm match and individual quality were time invariant. Even so, it will take time for the employee and firm to learn about the value of these parameters and uncertainty about the true quality declines with tenure. Therefore *perceived* match quality (which determines wages and mobility) is in fact time-variant. As such the stylised model presented is more applicable to perfect information setting and the search good model of job mobility (Burdett, 1978) in which match quality is known *ex ante*. In Jovanovic’s model there is a sequential updating of the expected value of their match quality and individual quality which then – if position held in the firm is indicative of the firm’s expectation of these qualities –this may lead to an evolutionary position change for the employees in the firm hierarchy. These positional changes are optimal responses from further learning about the workers’ quality and arguably ‘better’ workers should climb the job ladder faster.³³ As such this is not definitive evidence that expected match quality or individual quality is increasing in years of service. To explain, consider a cohort of employees who enter the firm and are scattered across various grades in the hierarchy. After the tenth year we may observe that the positions occupied by the remaining employees tend to, on average, be higher. This does not necessarily indicate, by itself, that the average value of (the time-invariant) match quality or individual quality of employees at the firm has increased after ten years of service for we do not observe what would occur if those employees who left the firm actually remained. It may be that the positions occupied by the entire cohort after ten years would have been, on average, even higher.

³³ Baker, Gibbs and Holmstrom (1994) in their analysis of the wage policy of the management employees in a single medium sized US based service firm found that workers who receive larger wage increases early in their stay at one level of the job ladder are promoted more quickly.

Table 3
Distribution of Employees in Grade Class Conditional on Years of Experience

a. Males/Secondary Education					b. Males/Higher Education						
pexpnet, x	Training Grades	Clerical Grades	Middle Management Grades	Senior Management Grades	pexpnet, x	Training Grades	Clerical Grades	Middle Management Grades	Senior Management Grades		
	(T1 – T3)	(C1 – C2)	(M1 – M2)	(SM1 –SM5)		(T1 – T3)	(C1 – C2)		(SM1 –SM5)		
x<1	1,110	11.08	36.4	39.91	8.38	x<1	1,245	4.66	15.83	44.9	28.75
1≥x<2	681	17.47	44.34	27.76	5.43	1≥x<2	573	8.9	21.64	34.73	21.46
2≥x<5	1,197	31	35.42	20.97	3.84	2≥x<5	1,163	8.43	23.91	35.25	16.42
5≥x<10	1,090	22.76	38.62	20.74	6.24	5≥x<10	1,093	3.57	11.89	35.23	34.67
10≥x<15	667	15.59	28.78	28.78	13.05	10≥x<15	553	2.35	8.13	31.1	46.47
15≥x<20	369	11.65	25.2	30.9	19.24	15≥x<20	268	0.75	4.48	33.96	48.51
20≥x<25	214	10.75	21.97	38.79	16.82	20≥x<25	130	2.31	6.16	38.46	43.08
x≥25	256	16.41	27.35	27.34	18.36	x≥25	90	4.44	12.23	33.34	38.89
service, y						service, y					
y<1	878	56.38	21.87	11.73	2.05	y<1	693	19.33	26.7	27.13	9.96
1≥y<2	548	38.68	32.48	15.14	4.74	1≥y<2	579	10.37	23.49	35.76	13.64
2≥y<5	886	25.51	44.24	13.77	5.75	2≥y<5	1,045	5.07	21.72	34.93	22.48
5≥y<10	894	6.48	45.42	29.42	8.72	5≥y<10	811	1.73	14.92	41.43	31.2
10≥y<15	1,718	4.3	39.53	42.55	8.55	10≥y<15	863	0.46	8.46	43.8	38.48
15≥y<20	200	1.5	15	40.5	18.5	15≥y<20	330	0.30	5.45	38.49	46.97
20≥y<25	253	0.40	17	39.92	26.48	20≥y<25	433	0.46	5.78	35.1	54.42
y≥25	207	1.94	15.94	40.58	29.96	y≥25	361	0.28	5.54	39.61	49.32
c. Females/Secondary Education					d. Females/Higher Education						
pexpnet, x	Training Grades	Clerical Grades	Middle Management Grades	Senior Management Grades	pexpnet, x	Training Grades	Clerical Grades	Middle Management Grades	Senior Management Grades		
	(T1 – T3)	(C1 – C2)	(C1 – C2)	(SM1 –SM5)		(T1 – T3)	(C1 – C2)	(M1 – M2)	(SM1 –SM5)		
x<1	1,446	15.63	61.75	18.12	0.97	x<1	709	9.87	34.13	42.03	7.05
1≥x<2	1,088	25.27	57.72	12.14	0.83	1≥x<2	388	20.62	35.06	26.03	6.18
2≥x<5	1,890	39.42	45.66	8.73	0.32	2≥x<5	711	20.11	31.79	25.18	6.75
5≥x<10	1,681	34.5	43.07	9.52	0.78	5≥x<10	504	10.91	24.6	31.74	16.07
10≥x<15	880	23.07	49.54	12.5	1.02	10≥x<15	206	5.83	24.76	35.92	17.48
15≥x<20	564	22.34	52.3	12.76	1.95	15≥x<20	95	9.47	22.11	24.21	24.21
20≥x<25	409	28.11	46.21	9.05	1.46	20≥x<25	44	13.63	20.45	40.9	13.64
x≥25	360	31.12	48.06	8.05	0.28	x≥25	28	7.14	32.14	21.43	14.28
service, y						service, y					
y<1	1,396	69.63	17.76	2.65	0.14	y<1	450	37.11	25.33	13.55	4
1≥y<2	911	55.55	30.63	4.06	0.22	1≥y<2	384	26.04	31.51	22.66	4.95
2≥y<5	1,463	33.36	49.08	3.89	0.34	2≥y<5	660	13.33	36.97	25.91	8.49
5≥y<10	1,324	10.81	64.05	14.95	0.83	5≥y<10	417	2.88	27.1	41.25	13.91
10≥y<15	2,650	9.28	67.32	18.9	0.86	10≥y<15	464	1.72	32.55	43.96	14.22
15≥y<20	213	6.57	55.87	23	5.17	15≥y<20	96	1.04	19.8	60.42	13.55
20≥y<25	239	4.6	55.65	25.94	2.93	20≥y<25	153	0.65	24.84	47.71	21.57
y≥25	122	4.92	59.02	21.32	6.56	y≥25	61		29.51	54.1	14.76

Note: Workers who are in unclassified grades in the hierarchy are not represented in the table. pexpnet refers to the proxied number of years of pre-firm labour market experience, while service refers to years of within-firm labour market experience (or tenure)

4. Empirical results

4.1. Nonparametric Estimates

The specification of the semi-parametric partially linear model is

$$\ln hrwage_i = \beta_0 + f(\text{service}_i, \text{pexpnet}_i) + \beta_1 ms_i + \beta_2 \text{child}_i + \beta_3 \text{region}_i + \varepsilon_i \quad (16)$$

where the mean of ε_i , conditional on the explanatory variables, is zero. The definition of the variables in the model was presented in the previous section, and estimation was discussed in section two.

I present in table 4, the parametric estimates $\hat{\beta}_i$ of equation 16. Given the interest of this paper is with regard to the within firm experience- earnings profiles, I focus discussion upon the estimation results of these. However as the parametric

estimates add further explanatory power to understanding the earnings differentials of the firm's workers, some discussion is merited. Firstly I note that all the estimates of the parametric components of the model are in line with prior expectations. A male employee with a secondary level of education who is based in London can expect to earn over 35 percent more than a similarly characterised individual employed elsewhere in Britain, all else equal. For female and male employees with a higher level of education, they can expect around 50 percent higher wages than those employees with a higher education upon labour market entry who are based elsewhere. The personnel records indicate that firm employs in 13 regions across Britain, with over 34 percent of the total sample employed in Greater London. Such estimates indicate that the London based employees are either of significantly higher quality and /or operate in jobs involving far greater complexity than elsewhere. Those who are married tend to enjoy an earnings premium relative to those not. Either marital status is acting as a proxy for personal traits relating to success in the work environment, or the relative stability associated with married individuals is being rewarded with a premium. The effect of having children upon earnings tends to operate in different directions across genders. A man with children can expect to earn around 9 percent more than a man without, while a woman with children can expect to earn nearly 5 percent less than a woman without, all else equal. Women with children have presumably taken time out of the labour market prior to entering the firm, which has correspondingly interrupted their human capital investment process, hence they earn less.

Table 4
Parameter Estimates of Semi-Parametric Model^a

Parameter	a.Males/Secondary Education	b.Males/Higher Education	c.Females/Secondary Education	d.Females/Higher Education
ms	.102 (.0135)	.126 (.0175)	.007* (.006)	.0352 (.0173)
child	.0925 (.0142)	.0873 (.018)	-.0466 (.007)	-.050 (.023)
region	.353 (.010)	.500 (.0120)	.284 (.006)	.510 (.015)
number of observations	5584	5115	8318	2685

^a Standard errors in parenthesis. Dependent variable is $\ln hrwage$. Nonparametric estimates of models displayed in figures 2 and 3. * indicates not significant at any reasonable level of confidence.

As two variables - years of within-firm experience and years of pre-firm experience - are being modelled nonparametrically the resulting estimates $\hat{f}(\cdot)$ can be graphed as a surface over p_{exp} and service coordinates. However it is hard to see details of the structure of $\hat{f}(\cdot)$ in such a plot. Therefore to aid interpretability I choose instead to display cross sections of the surface.

Figure 2 displays the estimated within-firm experience-earnings profiles with various years of pre-firm experience held constant.³⁴ In line with prior expectations, the expected starting wage at the firm is shown to be increasing in years of pre-firm experience. From a human capital perspective, an individual with greater years of previous labour market experience will have accumulated more general human capital upon entry to the firm and is therefore rewarded with a higher starting salary than a worker with less years of previous labour market experience prior to entry.³⁵ In

³⁴ The chosen years of pre-firm experience are 1, 5, 10, 15, and 25. Displaying more within-firm experience-earnings profiles would yield presentational difficulties. A full set of estimation results are available on request.

³⁵ This trend does slow down and reverse as years of previous experience increases further.

addition, the estimated profiles for those with a higher education are higher than that for those with a secondary education. This again is in line with prior expectations. The other immediately obvious trend is that the within-firm earnings profiles are not parallel across different years of previous years of labour market experience, therefore indicating that the function $f(\cdot)$ is not additive in functions of $pexpemet$ and service. Specifically, the estimated wage growth for a given number of years of service becomes lower the higher are the number of years of experience accumulated out-with the firm before entry. By the same token, regardless of the number of years of outside experience, earnings growth diminishes with years of tenure, hence the within-firm earnings profiles display concavity. This finding can of course be reconciled with basic predictions of rational human capital investment: the incentive to invest in human capital declines as the worker ages, as investment undertaken at older ages can only be rented out for shorter periods due to an investment horizon which naturally declines with age (Ben-Porath, 1967) and wages follow the productivity path created by such human capital investment.

In table 5, I present estimates of cumulative within-firm wage growth rewarded for the completion of various years of tenure. The table is divided into four panels, representing each of the cells over which estimation took place. Within each panel, each row refers to a given number of years of previous experience held constant and the remaining columns display expected cumulative wage growth from entry to after various years of tenure. Consider first the estimated within-firm earnings growth for male employees who have attained only up to a secondary level of education prior to entry into the labour market. The worker falling into this category who enters the firm after one year of previous labour market experience can expect cumulative wage growth of over 55 percent after five years of service and over 77 percent after ten years of service. While a worker entering after fifteen years of previous labour market experience can however expect cumulative wage growth of only 49 and 54 percent after the same number of years of service.

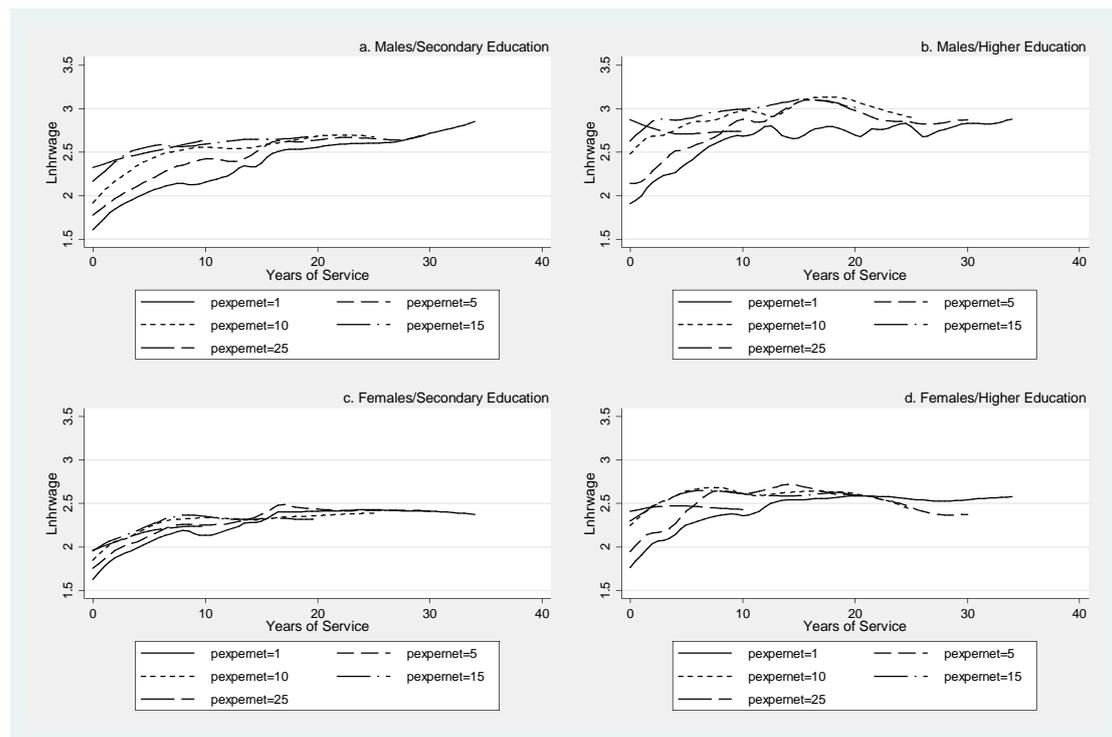


Figure 2: Nonparametrically Estimated Within-Firm Experience Earnings Profiles
 Table 3
 Estimates of Within-Firm Wage Growth*

service	1	2	3	4	5	10	15	20	25
pexpernet									
a.Males/Secondary Education:									
1	0.146	0.278	0.375	0.463	0.554	0.773	1.158	1.600	1.724
5	0.104	0.213	0.308	0.400	0.495	0.907	1.147	1.371	1.409
10	0.156	0.286	0.406	0.548	0.646	0.901	0.937	1.149	1.133
15	0.123	0.253	0.381	0.451	0.492	0.537	0.626	0.686	-
20	0.070	0.131	0.178	0.204	0.219	0.404	0.526	-	-
25	0.047	0.089	0.128	0.163	0.195	0.375	-	-	-
b.Males/Higher Education:									
1	0.097	0.270	0.381	0.433	0.582	1.171	1.130	1.199	1.437
5	0.017	0.145	0.296	0.447	0.486	1.082	1.543	1.304	1.246
10	0.155	0.230	0.249	0.318	0.415	0.648	0.830	0.837	0.675
15	0.130	0.248	0.294	0.266	0.445	0.610	0.467	-	-
20	0.036	0.100	0.172	0.187	0.121	0.121	0.121	-	-
25	-0.054	-0.098	-0.130	-0.149	-0.152	-0.126	-	-	-
c.Females/Secondary Education:									
1	0.157	0.284	0.361	0.411	0.528	0.656	0.955	1.195	1.232
5	0.110	0.228	0.307	0.365	0.441	0.644	0.858	0.976	0.949
10	0.132	0.241	0.287	0.367	0.467	0.634	0.595	0.671	0.720
15	0.091	0.156	0.208	0.273	0.346	0.497	0.451	0.445	-
20	0.059	0.120	0.188	0.251	0.316	0.408	0.398	-	-
25	0.050	0.102	0.154	0.202	0.248	0.320	-	-	-
d.Females/Higher Education:									
1	0.154	0.308	0.366	0.459	0.627	0.817	1.180	1.272	1.197
5	0.153	0.251	0.276	0.383	0.589	0.968	1.147	0.935	0.667
10	0.125	0.252	0.331	0.397	0.488	0.441	0.480	0.445	0.200
15	0.090	0.184	0.266	0.338	0.402	0.378	0.348	0.336	-
20	0.097	0.185	0.260	0.319	0.361	0.335	0.285	-	-
25	0.032	0.052	0.063	0.067	0.064	0.023	-	-	-

* Results refer to the estimated cumulative wage growth from zero to various years of service dependent upon a given number of years of pre-firm experience accumulated before entry to the firm.

To gain a different perspective into the estimated lifecycle earnings dynamics of the workers in the firm, in figure 3 I display the nonparametrically estimated log wage against *total years of labour market experience*. Consequently these within-firm earnings profiles begin at p years of labour market experience, with p being equal to the number of years of pre-firm labour market experience. Therefore, joining the starting points of each of the curves in the figures is in fact the estimated ‘starting earnings profile’: expected log wages for a given number of years of pre-firm experience with years of service held constant at zero.

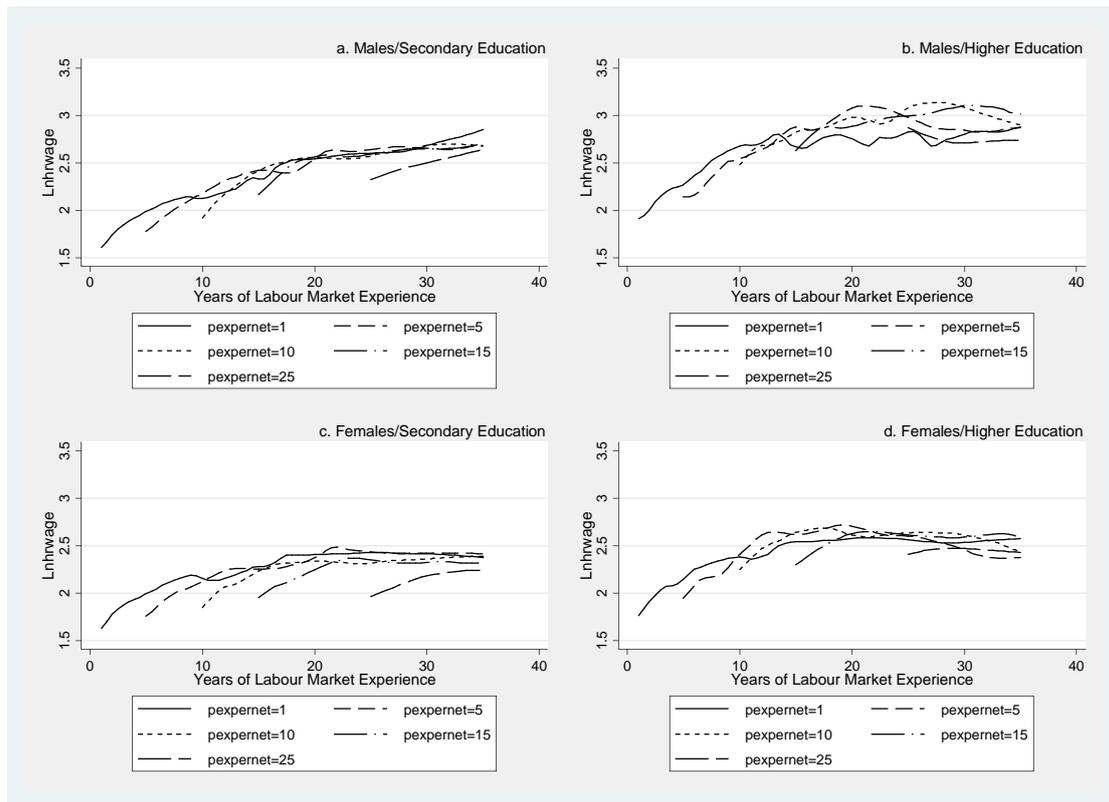


Figure 3: Nonparametrically Estimated Experience Earnings Profiles

Consistent with the notion that experience accumulated prior to entry into the firm will have involved both general and specific training, but that only general training (by definition) is reflected in entry wages at the firm, expected wages are less on entry -regardless of years of pre-firm experience –when compared to the wages of those with equal years of total labour market experience, but less years of pre-firm experience (so more years of service). However, the initial gap in estimated wages of an entrant to the firm, relative to the incumbents appears to diminish with years of service, as shown by the converge of the various within firm experience earnings profiles. A potential human capital explanation is simply that the accumulation of firm-specific human capital happens only fairly rapidly upon entry, so that once accumulated, expected earnings within the firm at a given number of years of total labour market experience tend to converge, regardless of the number of years of previous labour market experience accumulated before entry.³⁶

There are also clear differences between the profiles for those coming into the labour market with a higher education and those with a secondary education. Firstly, the estimated entry wages after a given number of years of previous experience compared to the estimated wages of the already incumbent employees is much less

³⁶ This is in fact perhaps consistent with Altonji and Shakotko's (1987) findings. They included a dummy variable equal to 1 if $tenure_{ijt} > 0$ in their earnings function to capture early career growth with tenure that is otherwise restricted by their parametric specification. In both their least squares and IV corrected estimates they find that much of the wage increase with tenure occurs in the first year of the job, even more so for the corrected (true) estimates in which the wage-tenure profile is essentially flat after zero years on the job (table 1, pp.444). Also, Bartel (1995, pp.408, table 4) using data of professional employees from the personnel records of a large manufacturing company with direct information on formal training confirmed that both the probability of receiving training and the amount of training are highest for the newly hired employees.

drastic than that displayed for those with a secondary school level of education. A potential explanation is that individuals with a higher level of education tend to be joining from firms within the same industry to a greater extent than those employees with a secondary level of education. Given that industry specific human capital is not lost when moving across firms within the same industry, then the gap in starting wages will be less for entrants compared to incumbents for those with a higher level of education than those with a secondary level of education.³⁷ However information on where previous experience was accumulated is unobserved in the personnel dataset and this precludes investigation of the hypothesis.

What seems at odds with human capital theory - on face value at least - is that the earnings profiles for individuals with a higher level of education holding various years of previous experience constant tend to overtake each other as years of service accumulates. In the absence of specific human capital we would expect that $E(\ln hrwage | service = 15) = E(\ln hrwage | service = 10, pexpnet = 5)$, but not that $E(\ln hrwage | service = 15) < E(\ln hrwage | service = 10, pexpnet = 5)$. However, this is only true if with reference to a given worker. The presented estimates are conditional on observable variables and unobserved heterogeneity across workers is not explicitly controlled for. In equation 3, it was hypothesised that the expected match quality of workers entering the firm is increasing in years of previous labour market experience and support for this hypothesis was given in table 3, where it was found that a given worker is far more likely to be situated in the top senior management level of the hierarchy if he enters the firm with more years of previous labour market experience. As such, a rational hypothesis is that the wage level is increasing in previous labour market experience and that it is increasing enough, so that the within-firm earnings profiles overtake as service increases. However it must be noted that another influence which is not controlled for is the influence of mobility behaviour of the employees after entry into the firm and again this will influence the estimated within-firm earnings profiles. Whether match-quality and individual quality is an increasing or decreasing function of years of tenure is unknown. Moreover, although the concave structure of the profiles are consistent with the predications of human capital theory, the influence of agency themed models with lifecycle implications, such as Lazear's (1979, 1981) delayed payment model, which assert that earnings will increase with service regardless of productivity enhancing human capital accumulation, cannot be ruled out.

5. Conclusion

The aim of this paper was to 1) respond to the specification uncertainty with regard to parametric earnings functions by specifying the earnings function as a

³⁷ Parent (2000) and Neal (1995) present evidence suggesting the importance of industry-specific human capital as opposed to purely firm-specific human capital for the interpretation of the earnings-tenure profile. Using data from both the National Longitudinal Survey of Youth (NLSY) and the Panel Study of Income Dynamics (PSID), Parent (2000) finds that when controlling for total experience in the current industry that the effect of tenure on wages is markedly reduced, almost fully so when using the methodological approach of Altonji and Shakotko (1987). Likewise, Neal (1995) finds, using data from the Displaced Worker Surveys, that tenure with the pre-displacement employer is positively correlated with the wage earned in the post-displacement job only for workers who stay in the same industry. Also, Neal (1999) argues, in a stylized model of job search, that more educated workers are less likely to make a career change (which he defines as a switch of industry of employment) after committing to full-time employment and finds support for this hypothesis using the National Longitudinal Survey of Youth (NLSY).

partially-linear semiparametric model and by estimating the earnings function across schooling groups as well as genders and to 2) investigate the interpretability issues surrounding experience-earnings profiles.

The nonparametrically estimated within-firm experience profiles confirmed the influence of heterogeneity across employees and the influence which this has upon the expected wages of individuals within the firm. The estimated within-firm earnings profiles were conditional upon the number of years of previous labour market experience accumulated before entry and it was shown that the profiles tend to converge. The rapidity of convergence was much faster for employees who entered the labour market with a higher level of education and in fact, for these employees, the profiles actually overtook each other. As such, the separation of schooling groups in estimation of the earnings profiles was merited. The fact that an individual entering the firm with more years of previous labour market experience is estimated to earn more than an incumbent employee with the same years of total experience, but more years of tenure, is un-reconcilable with the notion of homogenous employees, but can be explained with heterogeneity in expected match quality and individual quality, both of which are a function of years of service and previous experience. The hypothesis that expected firm-match quality is an increasing function of previous labour market experience was given support in the data analysis section, where it was shown the likelihood of an employee occupying a senior management grade in the hierarchy was increasing in years of previous experience.

The primary merit and value added in this paper was in the econometric methodology utilised. The approach is somewhat novel in this area of economics and the estimation results have formed a valuable tool in the exploration and understanding of earnings evolution within the firm.

References

- Abraham, K.G., Farber, H.S., 1987. Job duration, seniority, and earnings, *American Economic Review* 77 (3), 278-297
- Altonji, J.G., Shakotko, R.A., 1987. Do wages rise with job seniority? *Review of Economic Studies* 65 (3), 437-459
- Altonji, J.G., Williams, N., 1997. Do wages rise with job seniority? A reassessment, National Bureau of Economic Research Working Paper 6010, Cambridge, MA
- Anglin, P.M., Gencay, R. 1996. Semiparametric estimation of a hedonic price function, *Journal of Applied Econometrics* 11 (6), 633-648
- Barron, J.M., Black, D.A., Loewenstein, M.A., 1989. Job matching and on-the-job training, *Journal of Labor Economics* 7 (1), 1-19
- Bartel, A.P., 1995. Training, wage growth, and job performance: evidence from a company database, *Journal of Labor Economics* 13 (3), 401-425
- Basu, R., Ullah, A., 1992. Chinese earnings-age profile: a nonparametric analysis, *Journal of International Trade and Economic Development* 1, 151-165
- Becker, G.S., 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, first edition, Columbia University Press, New York
- Becker, G.S., Stigler, G., 1974. Law enforcement, malfeasance and compensation, *Journal of Legal Studies* 3, 1-18
- Ben-Porath, Y., 1967. The production of human capital and the life-cycle of earnings, *Journal of Political Economy* 75, 352-365
- Blinder, A.S., 1976. On dogmatism in human capital theory, *Journal of Human Resources* 11 (1), 8-22
- Blundell, R., Duncan, A., 1998. Kernel regression in empirical microeconomics, *Journal of Human Resources* 33 (1), 62-87
- Bratsberg, B., Terrell, D., 1998. Experience, tenure and wage growth of young black and white men, *Journal of Human Resources* 33 (3), 658-682
- Brown, J.N., 1989., Why do wages increase with tenure? On-the-job training and life-cycle wage growth observed within firms, *American Economic Review* 79 (5), 971-991
- Brown, S., Sessions, J.G., 2006. Some evidence on the relationship between performance-related pay and the shape of the experience-earnings profile, *Southern Economic Journal* 72 (3), 660-676
- Brunello, G., Ariga, K., 1997. Earnings and seniority in Japan: a re-appraisal of the existing evidence and a comparison with the UK, *Labour Economics* 4, 47-69
- Burdett, K., 1978. A theory of employee search and quit rates, *American Economic Review* 68 (1), 212-220
- Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association* 74 (368), 829- 836
- Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting, *Journal of the American Statistical Association* 83 (403), 596-610
- Cleveland, W.S., Loader, C., 1996. Smoothing by local regression: principles and methods, *Statistical Theory and Computational Aspects of Smoothing*, AT&T Bell Laboratories
- Craven, P., Wahba, G., 1979. Smoothing noisy data with spline functions. *Numer. Math* 31, 377-403
- DiNardo, J., Tobias, J.L., 2001. Nonparametric density and regression estimation, *Journal of Economic Perspectives* 15 (4), 11-28
- Fan, J., Gijbels, I., 1992. Variable bandwidth and local linear regression smoothers, *Annals of Statistics* 20 (4), 2008-2036
- Fan, J., Gijbels, I., 1995. Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation, *Journal of the Royal Statistical Society, Series B*, 57 (2), 371-394
- Fan, J., Heckman, N.E., Wand, M.P., 1995. Local polynomial kernel regression for generalized linear models and quasi-likelihood functions, *Journal of the American Statistical Association* 90 (429), 141-150
- Farber, H.S., 1999. Mobility and stability: the dynamics of job change in labor markets, in: Ashenfelter, O., Card, D., eds., *Handbook of Labor Economics*, vol. 3B, North-Holland, Amsterdam, Chapter 37
- Flabbi, L., Ichino, A., 2001. Productivity, seniority and wages: new evidence from personnel data, *Labour Economics* 8, 359-387
- Fox, J., 2000. *Multiple and Generalized Nonparametric Regression*. Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-131. Thousand Oaks, CA
- Härdle, W., 1990. *Applied Nonparametric Regression*, first edition, Cambridge University Press, Cambridge

- Härdle, W., Müller, M., Sperlich, S., Werwatz, A., 2000. *Nonparametric and Semiparametric Models*, Springer Series in Statistics
- Harris, M., Holmstrom, B., 1982. A theory of wage dynamics, *Review of Economic Studies* 49, 315-33
- Hashimoto, M., 1981. Firm-specific human capital as a shared investment, *American Economic Review* 71 (3), 475-482
- Hashimoto, M., Raisian, J., 1985. Employment tenure and earnings profiles in Japan and the United States, *American Economic Review* 75 (4), 721-735
- Hastie, T.J., Tibshirani, R.J., 1990. *Generalized Additive Models*, first edition, Chapman and Hall
- Heckman, J.J., Lochner, L.J., Todd, P.E., 2003. Fifty years of mincer earnings regressions, National Bureau of Economic Research Working Paper 9732
- Hellerstein, J.K., Neumark, D., 1995. Are earnings profiles steeper than productivity profiles? Evidence from Israeli firm-level data, *Journal of Human Resources* 30 (1), 89-112
- Hungerford, T., Solon, G., 1987. Sheepskin effects in the returns to education, *Review of Economics and Statistics* 69 (1), 175-177
- Hutchens, R.M., 1987. A test of Lazear's theory of delayed payment contracts, *Journal of Labour Economics* 5 (4), part 2: The new economics of personnel, S153-S170
- Jovanovic, B., 1979. Job matching and the theory of turnover, *Journal of Political Economy* 87, 927-990
- Lazear, E.O., 1979. Why is there mandatory retirement? *Journal of Political Economy* 87 (6), 1261-84
- Lazear, E.P., 1981. Agency, earnings profiles, productivity, and hours restrictions, *American Economic Review* 71 (4), 606-620
- Lazear, E.P., Moore, R.L., 1984. Incentives, productivity, and labor contracts, *Quarterly Journal of Economics* 99 (2), 275-296
- Lazear, E.P., 1986. Raids and offer matching. In: Ehrenberg, R.G., ed., *Research in Labor Economics*, Greenwich, CT, pp.141-65
- Lazear, E.P., Rosen, S., 1990. Male-female wage differentials in job ladders, *Journal of Labor Economics* 8 (1), part 2: essays in honor of Albert Rees, S106-S123
- Lemieux, T., 2006. The "mincer equation" thirty years after schooling, experience, and earnings. In: Gossbard-Shechtman, S. ed., *Jacob Mincer a Pioneer of Modern Labor Economics*, Springer Verlag
- Light, A., 1998. Estimating returns to schooling: when does the career begin? *Economics of Education Review* 17 (1), 31-45
- Light, A., Ureta, M., 1995. Early-career work experience and gender wage differentials, *Journal of Labor Economics* 13 (1), 121-154
- Loader, C.R., 1999a. *Local Regression and Likelihood*, Springer-Verlag, New York
- Loader, C.R., 1999b. Bandwidth selection: classical or plug-in? *Annals of Statistics* 27 (2), 415-438
- Medoff, J.L., Abraham, K.G., 1980. Experience, performance, and earnings, *Quarterly Journal of Economics* 95 (4), 703-736
- Medoff, J.L., Abraham, K.G., 1981. Are those paid more really more productive? The case of experience, *Journal of Human Resources* 16 (2), 186-216
- Mincer, J., 1974. *Schooling, Experience, and Earnings*, University of Columbia Press for the National Bureau of Economic Research, New York
- Mincer, J., Polachek, S. 1974. Family investments in human capital: earnings of women, *Journal of Political Economy* 82 (2), S76-S108
- Mincer, J., Ofek, H., 1980. Interrupted work careers: depreciation and restoration of human capital. National Bureau of Economic Research, Working paper 479
- Mincer, J. and Jovanovic, B., 1981. Labor mobility and wages. In: Rosen, S., ed., *Studies in Labor Markets*, University of Chicago Press, Chicago, pp. 21-63
- Murphy, K.M., Welch, F., 1990. Empirical age-earnings profiles, *Journal of Labor Economics* 8 (2), 202-229
- Neal, D., 1995. Industry specific human capital: evidence from displaced workers, *Journal of Labor Economics* 13 (4), 653-677
- Neal, D., 1997. The complexity of job mobility among young men, *Journal of Labor Economics* 17 (2), 237-261
- Parent, D., 2000. Industry-specific capital and the wage profile: evidence from the national longitudinal survey of youth and the panel study of income dynamics, *Journal of Labor Economics* 18 (2), 306-323
- Polachek, S.W., 1995. Earnings over the life cycle: what do human capital models explain? *Scottish Journal of Political Economy* 42 (3), 267-289
- Parsons, D.O., 1986. The employment relationship: Job attachment, work effort, and the nature of contracts. In: Ashenfelter, O., Layard, R., eds., *Handbook of Labor Economics*, 789-848

- Robinson, H., 2000. Are you experienced? Recent British evidence on experience-earnings profiles, *Applied Economics* 35, 1101-1115
- Robinson, P.M., 1988. Root-n-consistent semiparametric regression, *Econometrica* 56 (4), 931-954
- Ruppert, D., Wand, M.P., 1994. Multivariate locally weighted least squares regression, *Annals of Statistics* 22 (3), 1346-1370
- Salop, S., Salop, J., 1976. Self-selection and turnover in the labor market, *Quarterly Journal of Economics* 91, 619-628
- Schady, N.R., 2003. Convexity and sheepskin effects in the human capital earnings function: recent evidence for Filipino men, *Oxford Bulletin of Economics and Statistics*, 65 (2), 171-196
- Schmelensee, R., Stoker, T.M. Household gasoline demand in the United States, *Econometrica* 67 (3), 645-662
- Simonoff, J.S., 1996. *Smoothing Methods in Statistics*, first edition, Springer Series in Statistics
- Stevens, M., 2003. Earnings functions, specific human capital, and job matching: tenure bias is negative, *Journal of Labor Economics* 21(4), 783-805
- Topel, R., 1991. Specific capital, mobility, and wages: wages rise with job seniority. *Journal of Political Economy* 99 (1), 145-176
- Treble, J., Van Gameren, E., Bridges, S., Barnby, T., 2001. The internal economics of the firm: further evidence from personnel data, *Labour Economics* 8 (5), 531-552
- Ullah, A., 1988. Non-parametric estimation of econometric functionals. *Canadian Journal of Economics* 21, 3, 625-658
- Williams, N., 1991, Reexamining the wage, tenure and experience relationship, *Review of Economics and Statistics* 73 (3), 512-517
- Willis, R.J., 1986. Wage determinants: a survey and reinterpretation of human capital earnings functions. In: Ashenfelter, O., Layard, R., eds. *Handbook of Labor Economics* 1, Elsevier
- Yatchew, A., 1998. Nonparametric regression techniques in economics, *Journal of Economic Literature* 36 (2), 669-721
- Zheng, J.X., 2000. Specification testing and nonparametric estimation of the human capital model, *Advances in Econometrics* 14, 129-154