Empirical Age-Earnings Profiles

Kevin M. Murphy, *University of Chicago*

Finis Welch, *University of California, Los Angeles and Unicon Research Corporation*

The "human capital earnings function," in which earnings are expressed as a quadratic in potential experience, is probably the most widely accepted empirical specification in economics. In spite of its widespread acceptance, the human capital earnings function provides a very poor approximation of the true empirical relationship between earnings and experience. The standard formulation understates early career earnings growth by about 30%–50% and overstates midcareer growth by 20%–50%. However, simple alternative specifications that fit the data are available.

There is perhaps no convention in empirical economics as standard as the specification of age-earnings profiles. "The human capital earnings function," as it has been called after Jacob Mincer's famous book (1974), is one in which the (natural) logarithm of earnings per hour, week, or year is expressed as a linear function of the number of years of school completed

This research is supported by U.S. Department of Health and Human Services, National Institute of Child Health and Human Development grant 5R01HD21713. The opinions and conclusions expressed here are solely ours and should not be construed as representing the opinions or policy of any agency of the federal government.

[*Journal of Labor Economics*, 1990, vol. 8, no. 2]

0734-306X/90/0802-0003$01.50

and as a quadratic function of years since leaving school or (potential) work experience. Although there has been some attention paid to the question of whether the logarithmic transformation is appropriate for the dependent variable (see, e.g., Heckman and Polachek 1974; Lillard, Smith, and Welch 1986), except for Mincer's original work where he reports estimates from a Gompertz curve as an alternative to the quadratic, there has been almost no attention paid to the empirical specification of the way wages vary with age. The quadratic in experience, for all practical purposes, has been universally accepted.

The quadratic specification in general and the human capital earnings function in particular have served two primary roles in labor market research. First, estimates have been used to describe cross-sectional and longitudinal patterns of career earnings growth and to describe the ways earnings patterns have varied through time and across groups at a moment in time. In addition to the direct role, the quadratic specification has been used in countless empirical studies that estimate the effects of factors such as race, training, labor market conditions, and schooling on earnings. In many of the studies of the second type, the quadratic specification serves the indirect role of controlling for confounding effects of experience on earnings. For the quadratic to serve either of these roles effectively, it must provide a reasonable approximation to actual earnings profiles.

One would like to believe that the quadratic specification is used because it "fits the data." Unfortunately, as we show, this is not the case. The quadratic approximation results in significantly biased estimates of the earnings profile. A simple comparison illustrates the magnitude of the bias. There is a fairly large collection of literature devoted to estimating the way age-income profiles shift or rotate as age and educational distributions change. Most of this literature follows the observation that wages of young workers fell relative to their more experienced contemporaries when the baby-boom generations entered the labor force. Examples of the cohort-size-effect literature include Freeman (1979), Welch (1979), Berger (1984, 1985), Murphy, Plant, and Welch (1988), and Murphy and Welch (1988). Based on the evidence in these and related papers on changes in the income gains associated with added schooling (see, in particular, Freeman 1976 and Murphy and Welch 1988), few would argue that age-income or income-schooling relationships are stable.

In fact, the evidence supporting changes in the shape of cross-sectional age-income profiles is so strong that it seems wrong for an empirical analyst to ignore it when studying the wage structure over the last 2 decades. However, the specification bias induced by ignoring changes in the shape of age-earnings profiles from year to year is actually smaller than the bias generated by the quadratic specification itself. As measured by residual sums of squared estimation errors, a fixed profile that allows only changes in the intercepts from year to year fits the data better than independent quadratics estimated for every year.

In addition to establishing average magnitudes of biases resulting from the quadratic specification, we illustrate the form of the bias. It is systematic and has profound effects on estimates of earnings growth. For example, the quadratic specification implies that wages of high school graduates grow by 35% in the first 10 years in the labor market, by 25% in the next 15 years, and decline by 16% in the subsequent 15 years. However, actual growth in the first 10 years is 54%, the rise over the next 15 years is only 18%, and the decline in the subsequent 15 years is less than 5%. Hence, the quadratic specification understates early career earnings growth by almost 50% and overstates midcareer growth by about one-third. Fully two-thirds of the late career decline implied by the quadratic is an artifact of specification. As these comparisons illustrate, the quadratic provides a poor approximation to actual earnings growth over the career.

In the process of illustrating the nature of the bias induced by the quadratic, we show that the form of the bias is remarkably stable over time and across educational groups at a given moment in time. We interpret the stability of the bias from the quadratic specification as evidence that there is a "true" specification of the earnings profile as envisioned by the Mincer earnings function but that this "true" specification is not quadratic. A search for a parsimonious specification that comes closer to approximating this true function is the final part of our study.

Section I describes the data from the March 1964–March 1987 Current Population Survey (CPS) that we use to evaluate the quadratic specification. Section II gives the estimates of bias induced by the quadratic specification and derives simple statistics that we use to evaluate alternative functional forms. Section III evaluates some simple alternative functional forms including higher-order polynomials and a more parsimonious quadratic-in-a-quadratic specification. Section IV compares the estimated earnings functions obtained from alternative functional forms with those obtained from the quadratic and shows how differences affect estimates of career earnings growth.

I. The Data

We have CPS data for 24 surveys beginning in 1964 and ending in 1987. Within each survey we select a wage sample of white men classified by age and education.[1] The education groups include categories of 8–11 years, 12 years, 13–15 years, and 16 or more years of school completed. Within each educational level, workers are divided into 40 pseudoexperience groups based on single years of age. For those with 8–11 years of schooling, experience is age less 18 years. For high school graduates, experience is age minus 19 years. Experience for men with 13–15 years of schooling is age less 20 years, and for college graduates we subtract 22 years from age. Thus, we analyze earnings for men ages 19–58, 20–59, 21–60, and 23–62, respectively, for high school dropouts, high school graduates, college dropouts, and college graduates.

Rather than analyze individual wages, we compute an average wage for each schooling-by-experience cell, so the data are compressed to 3,840 observations (24 years × 40 experience levels × 4 schooling levels). There are a total of 463,248 individual observations represented in the 3,840 cells. The number of observations is a minimum of 9,102 in 1965 and reaches a maximum of 21,934 in 1985. Numbers of observations in individual cells range from a low of 7 to a high of 439. Although the analysis is conducted using three wage measures—earnings per year, per week, and per hour—results for weekly earnings are presented.

The average log weekly wage in a year-experience-schooling cell is defined as the average of [log(annual earnings) − log(weeks worked)] among men observed in the cell. In addition to the average log weekly wage for each cell, we compute the variance of the cell average using the within-cell variation in log weekly wages. Hence, we have estimates of the average wage and the precision of the estimate for each cell.

Since the estimates of the within-cell variance are noisy, we smoothed them prior to use by first computing estimates of the variance of individual log weekly wages by cell (pooled over years). The estimates of the microvariance were then regressed on a constant term and four powers of experience within each educational group. Fitted microvariances are then combined with cell-specific sample sizes to estimate cell variances as

$$\sigma^2_{\bar{\epsilon}_{ijt}} = \frac{\hat{\sigma}^2_{ij}}{N_{ijt}}, \tag{1}$$

where $\hat{\sigma}^2_{ij}$ is the estimated variance of log weekly wages for educational level i and experience level j, and N_{ijt} is the sample size for this experience/educational cell in year t.

As will be clear below, the variance estimates play a key role in our ability to measure bias in the quadratic and other specifications of earnings profiles. In addition to providing the repeated observations within each cell required to calculate the variance estimates, the CPS data are ideally suited for our analysis for at least three reasons. First, the data cover a

[1] Those included in the wage sample were civilian and not enrolled in school (major activity). In the previous year (to which income data refer), they worked at least 40 weeks and were full-time employees, were not self-employed, and did not work without pay. Their weekly wage (defined as wage and salary earnings divided by weeks worked the previous year) must be at least $10. Because the 1964–67 survey questionnaires differ from those beginning in 1968, criteria for inclusion differ slightly for the two periods. For the 1968–86 surveys, men were excluded if they lived in group quarters or if they worked part of the previous year due to being in school, retired, or in the armed forces. For the 1964–67 surveys, those whose industry and occupations were coded as 99 (in the armed forces or did not work) were excluded.

longer period of time than any other comparably sized data set. This allows us to evaluate the appropriateness of the alternative specifications for permanent and transitory components of bias. Second, the large samples allow us to evaluate specifications for individual schooling groups and to examine differences and similarities in the specification bias across them. Finally, the large samples allow us to obtain precise estimates of the actual average earnings profile over the sample and reasonably precise estimates of individual year profiles that can then be used to describe the specification biases against an unrestricted alternative.

II. The Quadratic Specification

Our analysis of the quadratic specification is based on 96 independent regressions,[2] one for each of the four schooling classes in each of the 24 years.[3] We estimate the parameters for each of the 96 education-year pairs by weighted least squares using the inverse of the variance estimates from equation (1) as the weights. We then compute an average of the estimated profiles (weighted by sample size in each year) to obtain estimates of the average quadratic earnings profile over the sample. The average profiles are given in figure 1, where the profiles for the four educational groups are stacked on the same graph.

The results are unsurprising and resemble those obtained in previous research. The earnings profiles are increasing over the early career, are declining slightly approaching retirement, and are strongly ordered by education. However, some disagreement with the standard human capital earnings function is already apparent in figure 1 since the profiles are not vertically parallel across schooling groups, as the standard specification would imply. However, the profiles are much closer to being vertically parallel when they are stacked on the basis of experience as in figure 1 than they would be if stacked on the basis of age. Mincer's emphasis on experience rather than age seems on target.[4]

Figure 2 presents estimates of the "true" average earnings profile using the sample size weighted average of the yearly cell means rather than the

[2] Those familiar with the rotation group structure of the CPS survey will note that our samples are not strictly independent over years due to an overlap of somewhat less than one-half of the survey from March to March.

[3] While throughout this article we estimate cross-sectional profiles, the same results we state here apply equally well to longitudinal profiles that allow for year effects. To see this, one simply has to note that an annual quadratic specification pooled over years has the same form as a longitudinal specification pooled over cohorts, as long as one allows for the same mix of cohort and year effects in the two specifications. While there are some issues of identification in such models, they are irrelevant for our purposes since the least squares residuals are well defined in any case.

[4] Formal statistical analyses that search for the offset that most correctly fits the data confirm our chosen offset, except that college graduates should begin at age 24 rather than 23.

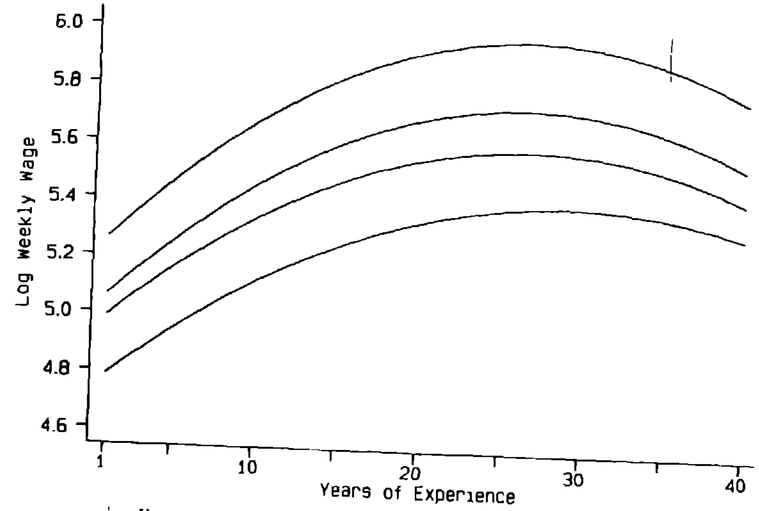

FIG. 1.—Average earnings profiles from quadratic, 1963–86

quadratic predictions. The general shapes seem roughly similar to those implied by the quadratic. The earnings profiles are concave with rapid initial earnings growth; they peak around 30 years of experience and decline slightly toward the end of the career retirement as in the quadratic specification.

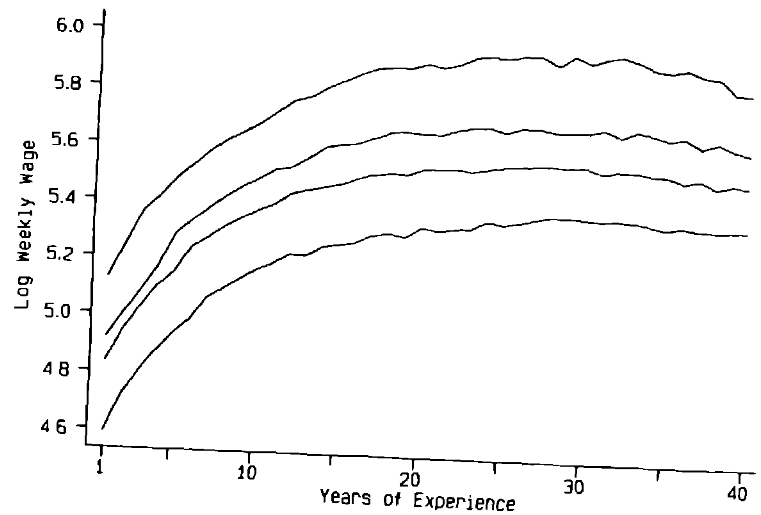

FIG. 2.—Actual average earnings profiles, 1963–86

In fact, in our opinion, it is the superficial similarity of the actual and quadratic profiles, together with the simplicity of the quadratic form, that has led to its widespread acceptance. But general appearances can be deceiving. Figure 3 overlays the quadratic and actual average profiles for each of the four schooling groups. The discrepancies between the profiles are apparent in these figures. The quadratic overstates initial earnings for all schooling groups and understates earnings at 10 years of experience for all groups. The quadratic also overstates earnings at midcareer and understates actual earnings at retirement. As we describe in Section IV, this causes the quadratic to understate earnings growth in the early career, to overestimate midcareer growth, and to greatly exaggerate the decline in earnings in the late career. These same patterns are also evident in comparisons of the individual year profiles.

The patterns of errors in the quadratic specification are shown clearly in figure 4, where the residuals (actual minus quadratic prediction) are plotted against experience for the four schooling groups. The similarity of the specification errors across schooling groups is striking, as is the magnitude of these errors. The similarity across schooling groups (which are independently estimated) and the smoothness of these specification errors across years of experience show that the estimates represent largely systematic biases and are not simply artifacts of sampling error. The specification errors are largest in the early career, where differences of up to 18% between the actual values and the quadratic predictions are observed for those with 8–11 years of schooling. The errors in the early career are smaller in percentage levels for higher schooling levels but are still significant (over 10%). For all schooling groups, the errors decline over the early career until they cross zero at between 5 and 7 years of experience. From 7 to about 17 years of experience, the quadratic understates earnings with a maximum error of about 6% at 10 years of experience. From about 18 until about 35 years of experience, the errors are negative (indicating that the quadratic overstates earnings) with a maximum absolute error of around 6% at about 25 years of experience. After 35 years of experience the errors become positive, so that by the end of the career the quadratic understates earnings for all schooling groups with a maximum error of about 8% for high school graduates. Based on the magnitude and pattern of residuals, it seems fair to say that the quadratic fits the actual empirical earnings profiles poorly, especially at low levels of experience.

While formal statistical tests are not required to evaluate the quadratic, given the evidence in figure 4, we now derive some simple statistical measures that provide a benchmark for our analysis of other functional forms where the specification errors are less obvious. Consider our basic equation for schooling group i in year t, where we estimate

$$y_{ixt} = b_0 + b_1 x + b_2 x^2 + v \qquad (2)$$

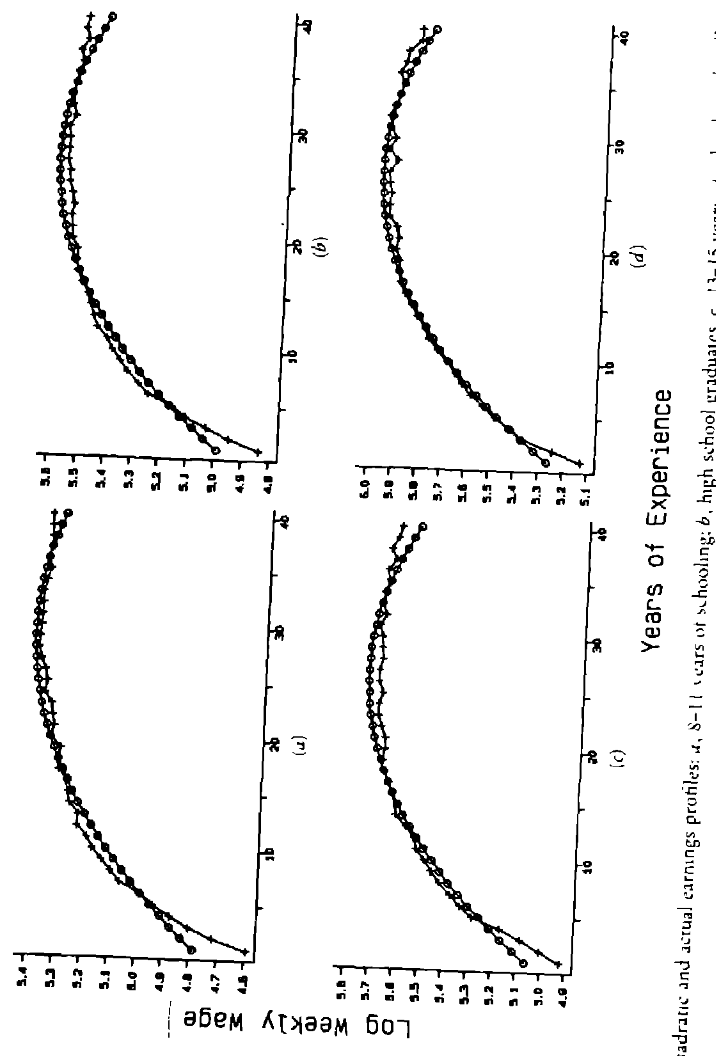

FIG. 3.—Quadratic and actual earnings profiles: *a*, 8–11 years of schooling; *b*, high school graduates; *c*, 13–15 years of schooling; *d*, college graduates

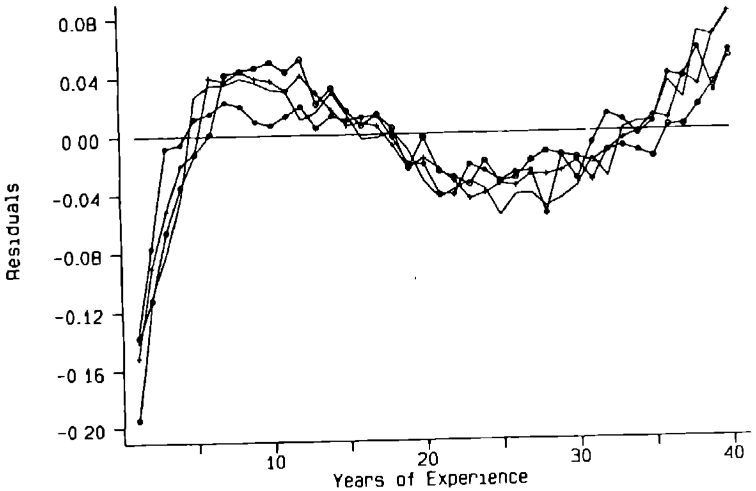

FIG. 4.—Residuals from quadratic specification: ○ = 8–11 years of schooling; + = high school graduates; □ = 13–15 years of schooling; plain line = college graduates

via weighted least squares. Letting $y^*_{ixt} = y_{ixt}(\sqrt{N_{ixt}}/\hat{\sigma}_{it})$, $x^*_{ixt} = x(\sqrt{N_{ixt}}/\hat{\sigma}_{it})$, $(x^2)^*_{ixt} = x^2(\sqrt{N_{ixt}}/\hat{\sigma}_{it})$, and $w^*_{ixt} = \sqrt{N_{ixt}}/\hat{\sigma}_{it}$, we can write this equation in the weighted data as

$$y^*_{ixt} = b_0 w^*_{ixt} + b_1 x^*_{ixt} + b_2 (x^2)^*_{ixt} + v^*_{ixt}, \tag{3}$$

where the variance of $v^*_{ixt} \approx 1$ for all i, x, t. However, unless the true earnings profiles are quadratic, v^*_{ixt} will contain both a bias component and a random error component. Hence, we have

$$v^*_{ixt} = \theta^*_{ixt} + \varepsilon_{ixt}, \tag{4}$$

where $\theta_{ixt} = E(v^*_{ixt})$, $E(\varepsilon_{ixt}) = 0$, and $E(\varepsilon^2_{ixt}) = 1$ for all i, x, t. A natural test statistic to measure the amount of bias is to use

$$\hat{\lambda}_{it} = \sum_x (\hat{v}^*_{ixt})^2, \tag{5}$$

where $\hat{v}^*_{ixt}$ is the weighted least squares residual from estimating equation (3). Under the null hypothesis of no specification bias we have $E(\hat{\lambda}_{it}) = 37$, and $\hat{\lambda}_{it}$ is approximately χ^2 with 37 degrees of freedom. In addition, when $\theta_{ixt} \neq 0$, we have

$$E(\hat{\lambda}_{it}) = \sum_x \theta^2_{ixt} + 37, \tag{6}$$

so that if we define our estimate of the bias as

$$\hat{\beta}_{it} = \frac{1}{40}(\hat{\lambda}_{it} - 37), \tag{7}$$

we have

$$E(\hat{\beta}_{it}) = \frac{1}{40}\sum_x \theta^{*2}_{ixt}, \tag{8}$$

so that $\hat{\beta}_{it}$ gives us an unbiased estimate of the mean squared bias (in the weighted data) for the quadratic specification for schooling group i in year t. Similarly, aggregates such as

$$\hat{\beta} = \frac{1}{96}\sum_i \sum_t \hat{\beta}_{it} \tag{9}$$

that estimate the average squared bias over all years and schooling groups can be defined to obtain measures of bias aggregated over time and/or groups.

For some comparisons it is useful to look at the unweighted data. For the unweighted data we can rewrite equation (4) as

$$v_{ixt} = \theta_{ixt} + \frac{\hat{\sigma}_{it}}{\sqrt{N_{ixt}}}\varepsilon_{ixt}, \tag{10}$$

where v_{ixt} is the actual residual from the quadratic specification and θ_{ixt} is the bias term. A natural decomposition is to write equation (10) as

$$v_{ixt} = \bar{\theta}_{ix} + \tilde{\theta}_{ixt} + \frac{\hat{\sigma}_{it}}{\sqrt{N_{ixt}}}\varepsilon_{ixt}, \tag{11}$$

where $\bar{\theta}_{ix}$ is the permanent bias component and $\tilde{\theta}_{ixt}$ is a year-specific transitory bias component. For this purpose we define

$$\bar{\theta}_{ix} = \frac{\sum_t N_{ixt}\theta_{ixt}}{\sum_t N_{ixt}} \tag{12}$$

as an observation weighted average of the annual bias terms, which leaves $\tilde{\theta}_{ixt} = \theta_{ixt} - \bar{\theta}_{ix}$. In the weighted data we can then rewrite equation (4) as

$$v^*_{ixt} = \frac{\sqrt{N_{ixt}}}{\hat{\sigma}_{ix}}\,\bar{\theta}_{ix} + \frac{\sqrt{N_{ixt}}}{\hat{\sigma}_{ix}}\,\tilde{\theta}_{ixt} + \varepsilon_{ixt}, \tag{13}$$

so that our estimate of the bias over all years for schooling group i as

$$\hat{\beta}_i = \frac{1}{960}\left(\sum_x \sum_t (\hat{v}^*_{ixt})^2 - (37)(24)\right) = \frac{1}{960}\sum_x \sum_t (v^*_{ixt})^2 - \frac{37}{40}, \tag{14}$$

where

$$E(\hat{\beta}_i) = \frac{1}{960}\left[\sum_x \frac{N_{ix}}{\sigma^2_{ix}}\,\bar{\theta}^2_{ix}\right] + \frac{1}{960}\left[\sum_x \sum_t \frac{N_{ixt}}{\sigma^2_{ix}}\,\tilde{\theta}^2_{ixt}\right] = \beta_{ip} + \beta_{it}, \tag{15}$$

where the first term, β_{ip}, is the permanent bias component and β_{it} is defined to be the total transitory component of bias. The permanent component is estimated as

$$\hat{\beta}_{ip} = \frac{1}{960}\sum_x\left(\sum_t \frac{\sqrt{N_{ixt}}}{\sigma_{ix}}\,\hat{v}^*_{ixt}\right)^2 - \frac{1}{960}\sum_x \sum_t \left[E\left(\hat{\varepsilon}^2_{ixt}\,\frac{N_{ixt}}{\sigma^2_{ix}}\right)\right], \tag{16}$$

where the expected values, $E(\hat{\varepsilon}^2_{ixt})$, are calculated from the diagonal elements of the covariance matrix of the regression residuals. Table 1 presents

Table 1
Bias Calculations for Weighted Residuals from Quadratic Specification

Educational Group	Mean Squared Residual (1)	Expected Mean Squared Residual (2)	Estimated Total Bias Component* (3)	Estimated Permanent Bias Component† (4)	χ^2 (5)
8–11	1.556	.925	.631	.538	1,494 (888)
12	2.615	.925	1.69	1.56	2,510 (888)
13–15	1.714	.925	.789	.678	1,645 (888)
16+	1.282	.925	.357	.293	1,231 (888)
Overall	1.792	.925	.867	.767	6,881 (3,552)

NOTE. – Degrees of freedom are in parentheses.
* Calculated as mean squared residual (col. 1) minus expected mean squared residual (col. 2).
† Calculated as in eq. (16).

the results of these calculations for each of the four schooling groups and for the four groups combined. As is clear from the table, about one-half of the mean squared errors in the year-by-year specifications are due to bias (.867 out of 1.792 overall), and of this bias, over 88% (.767/.867) represents a permanent specification bias that is common across years. The bias is largest for high school graduates, the group with the most data (about 40% of our sample), mostly due to the fact that bias is measured relative to the error variance (see below for estimates of the absolute bias). For high school graduates, the bias accounts for 65% (1.69/2.615) of the sum of squared residuals for the individual year regressions. Other calculations show that bias accounts for over 95% of the sum of squared residuals for the average profile for high school graduates, reflecting the fact that the permanent bias is nearly as large as the total bias, but the residual variance for the average profile is significantly smaller than the residual variance for the individual year profiles. For college graduates, the bias terms are smaller (relative to residuals) but are still significant since the χ^2 statistic of 1,231 is still 8 standard errors above its expected value of 888. For high school graduates, the χ^2 statistic of 2,510 is 38 standard deviations above its expected value of 888. When the four schooling groups are combined in the overall calculation, the bias has a χ^2 statistic of 6,881, which is 39 standard deviations above its expected value (3,552).

Table 2 recomputes a similar decomposition to that used in table 1 for the actual regression residuals rather than the weighted residuals used for table 1. The disadvantage of the unweighted calculations is that the error sum of squares is no longer distributed as a χ^2, but these calculations have the advantage that the magnitudes are directly interpretable and comparable across educational classes. To facilitate interpretation we translate the total and permanent bias calculations into root mean squared format in the final two columns. Since the overall numbers in the first four columns tell a similar story to those in table 1, we focus on the group comparisons and overall calculations using the root mean squared error calculations. Comparing groups, we find that the root mean squared bias is highest for those with 13–15 years of schooling (.0507) and lowest for college graduates (.0325). The overall root mean squared bias implies that the quadratic specification leads to an average specification error (in the root mean squared sense) of over 3.25%. (The average absolute error is 3.15%.) In the aggregate we find that permanent bias accounts for nearly all of the total bias. In fact, for college graduates the point estimates imply that the permanent bias exceeds the total bias (which reflects sampling error). The specifications also confirm what we saw in figures 3 and 4, that the magnitude of the specification bias is largest for those with 13–15 years of schooling and for those with 8–11 years of schooling. It is smallest for college graduates. Nevertheless, the bias components account for over 42% (.0019/.0045) of the regression residuals for college graduates.

Table 2
Root Mean Squared Bias Calculations for Unweighted Residuals from Quadratic Specification

Educational Group	Mean Squared Residual (1)	Expected Mean Squared Residual (2)	Total Mean Squared Bias (3)	Mean Squared Permanent Bias (4)	Total Root Mean Squared Bias (5)	Root Mean Squared Permanent Bias (6)
8–11	.0047	.0025	.0022	.0020	.0469	.0443
12	.0029	.0010	.0019	.0017	.0436	.0418
13–15	.0058	.0032	.0026	.0020	.0507	.0453
16+	.0046	.0036	.0011	.0011	.0325	.0330
Overall	.0045	.0026	.0019	.0017	.0439	.0414

Table 3
Correlations of Regression Residuals* across Schooling Levels for Quadratic Specification

Educational Group	8–11	12	13–15
12	.93	.	. .
13–15	.87	.94	. .
16+	.84	.89	.90

* Unweighted

So far, we have derived statistics for measuring the magnitude of the specification bias and for decomposing bias into permanent and transitory components. The concern is not simply with the magnitude of the bias but with its systematic nature. Tables 3 and 4 address this issue. Table 3 gives the correlations of averaged (over years) regression residuals between schooling levels. Since the regressions are computed independently, the expected value of these correlations is zero under the hypothesis of no specification error. The fact that the lowest correlation is .84 confirms what we knew from figure 4, that the bias pattern is very common to all educational levels. Table 4 presents statistics based on the covariance of residuals at adjacent experience levels and overall. Since these residuals come from the same regression, the expected covariance is not zero and is actually negative. The large positive covariances given in table 4, ranging from .235 to 1.176, effectively pick up the smooth nature of the bias illustrated in figure 4.

The final column in table 4 uses the estimated covariance of the residuals (col. 1) less the expected covariance of the residuals (col. 2) divided by the estimated mean squared total bias (col. 3 from table 1) to obtain an estimate of the first-order autocorrelation of the permanent bias component. These estimates range from a low of .70 for those with 8–11 years of schooling to a high of .85 for college graduates. The overall autocorrelation is .75, which reflects the smooth residual patterns seen in figure 4.

Where do these calculations leave us? In our opinion, the bias in the quadratic represents both good and bad news. First, the fact that the bias

Table 4
Covariances of Experience-adjacent Residuals from Quadratic Specification

Educational Group	Expected Covariance	Covariance	Autocorrelation of Bias
8–11	−.069	.385	.70
12	−.070	1.176	.74
13–15	−.069	.531	.76
16+	−.070	.235	.85
Overall	−.070	.582	.75

is so large is bad news for the quadratic specification. However, the systematic nature of this bias across schooling and experience levels (within a schooling group) and the stability of the bias component through time indicate that simple alternative specifications may solve the problem. It is on this basis that we stated in the introduction that there appears to be a "true" earnings function specification but that the specification is not a quadratic. The next section represents our attempt to find a reasonably parsimonious approximation to this true specification.

III. Alternative Specifications

The major virtues of the quadratic specification are that it is easily estimated and has relatively few parameters, thus conserving degrees of freedom when researchers have many fewer observations than we have in the CPS. Our goal in this section will be to find some simple alternatives to the quadratic that eliminate the majority of the bias and then to try to simplify the alternatives to obtain a functional form with as few parameters as the quadratic.

Since the ease of estimation for the quadratic owes largely to the fact that it is linear in the parameters, we restrict alternatives to specifications that maintain linearity. A simple extension of the quadratic is to add higher-order terms in experience. Figures 5 and 6 plot the regression residuals for cubic and quartic specifications, respectively. The scale used in both plots is the same as is used for the quadratic in figure 4 for ease of comparison.

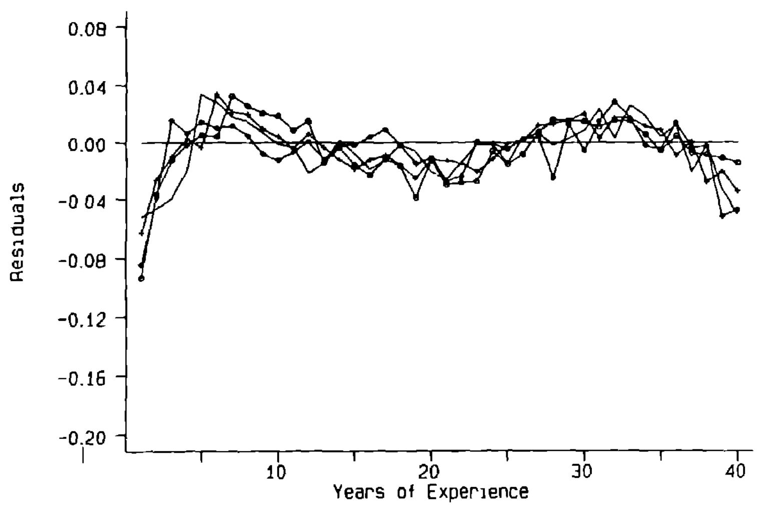

FIG. 5.—Residuals from cubic specification: O = 8–11 years of schooling; + = high school graduates; □ = 13–15 years of schooling; plain line = college graduates

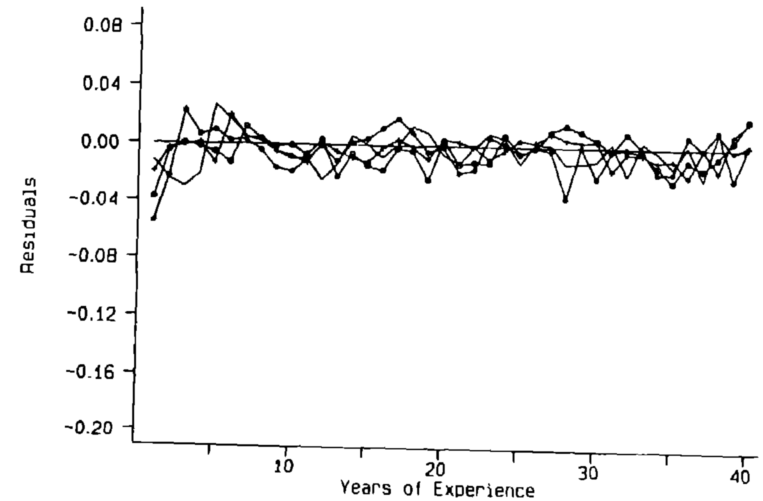

FIG. 6.—Residuals from quartic specification: O = 8–11 years of schooling; + = high school graduates; □ = 13–15 years of schooling; plain line = college graduates.

The residuals from the cubic show a noticeable pattern similar but less extreme than for the quadratic, while the quartic residuals show very little pattern.

For the cubic, the early career residuals are reduced significantly relative to the quadratic specification (5%–9% in the first year vs. 14%–20%). However, the general pattern remains much the same over the first 10–12 years. Residuals are negative in the early years and then positive from years 4–12 with a maximum positive residual of between 2%–3.5%. Things are different in the later career with negative residuals of 2%–4% around 30 years of experience and negative residuals again at the end of the career. Clearly the cubic is superior to the quadratic, but a significant systematic bias component still remains (particularly in the early career).

For the quartic residuals in figure 6, there appears to be a slight pattern to the residuals, and initial earnings are still overstated by between 1%–6%, versus 14%–20% for the quadratic, and 5%–9% for the cubic. While there appears to be some bias left in the quartic specification, it seems clear that it represents a major improvement over the cubic and especially over the quadratic.

In terms of our statistical tests, tables 5 and 6 present the same statistics for the cubic and quartic that are presented for the quadratic in tables 1 and 2. As shown in table 5, the addition of a third-order term reduces the overall mean squared bias by 75% (from .867 to .214) relative to the quadratic, while the permanent bias component is reduced by 80% (from .767 to .151). Similar improvements are seen for each of the individual edu-

Table 5
Bias Calculations for Weighted Residuals from Cubic and Quartic Specifications

Educational Group	Mean Squared Residual (1)	Expected Mean Squared Residual (2)	Estimated Total Bias (3)	Estimated Permanent Bias (4)	χ^2 (5)
Cubic:					
8–11	1 078	900	.178	129	1,035 (864)
12	1.267	.900	.367	281	1,216 (864)
13–15	1.108	.900	.208	118	1,063 (864)
16+	1.003	.900	103	076	963 (864)
Overall	1.114	900	.214	151	4,276 (3,456)
Quartic:					
8–11	.922	.875	047	016	885 (840)
12	.997	.875	122	.036	957 (840)
13–15	.975	.875	.010	.029	936 (840)
16+	.929	875	054	036	892 (840)
Overall	.956	.875	.081	.029	3,672 (3,360)

NOTE —Degrees of freedom are in parentheses

Table 6
Root Mean Squared Bias Calculations for Unweighted Residuals from Cubic and Quartic Specifications

Educational Group	Mean Squared Residual (1)	Expected Mean Squared Residual (2)	Total Mean Squared Bias (3)	Mean Squared Permanent Bias (4)	Root Mean Squared Total Bias (5)	Root Mean Squared Permanent Bias (6)
Cubic:						
8–11	00299	.00243	024	.021	.024	021
12	00138	.00099	.020	.018	.020	018
13–15	.00389	00311	.028	.018	028	018
16+	00360	.00347	011	018	011	018
Overall	00297	00250	.022	.019	022	.019
Quartic:						
8–11	.0025	.00235	00026	.00005	.013	007
12	0011	00096	.00014	00005	012	007
13–15	.0034	.0030	00041	.00004	02	.006
16+	.0033	.0034	.00009	.00001	.	01
Overall	.0025	0024	.00015	00006	012	008

cational classes as well. The largest improvement is for high school graduates, where the total bias declines by 78% (from 1.69 to .367) and the permanent bias declines by 82%. The smallest decline is for those with 8–11 years of schooling, where total bias drops by "only" 72% (from .631 to .178) and permanent bias declines by 75% (from .538 to .129).

The quartic specification makes significant progress relative to the cubic, reducing the overall total bias component by 62% relative to the cubic, thereby reducing the total bias component to only 9% of the level found in the quadratic specification.

In terms of permanent bias, the improvements with the quartic are even greater. Overall, permanent bias is reduced by 81% (from .151 to .029) relative to the cubic and 96% (from .767 to .029) relative to the quadratic. Hence, in terms of the systematic component of bias across years, the quartic makes a major improvement relative to the cubic, eliminating 80% of the mean squared bias, and an enormous improvement relative to the quadratic, eliminating all but 4% of the mean squared bias. In terms of the individual schooling groups, moving from the cubic to the quartic reduces permanent bias most for the two lowest educational groups (by 88% for 8–11 years of schooling and 87% for high school graduates) and least for college graduates, where permanent bias is reduced by 53%.

In terms of the χ^2 statistics, the critical level for the cubic specifications is 947 for the individual schooling groups and is 3,619 overall. Hence, the null hypothesis of no specification error can be rejected for all schooling groups individually and overall. The only "close call" is for college graduates, where the χ^2 statistic is 2.34 standard deviations above its expected value. For the quartic specification, the critical values of the χ^2 statistic are 920 for the individual education groups and 3,521 overall. In this case, the χ^2 statistics are 1.08, 2.80, 2.30, and 1.24 standard errors above their expected value for the individual groups (from lowest to highest educational levels) and 3.8 standard errors above for the overall calculations (these compare with 4.04, 8.32, 4.70, 2.34, and 9.88 standard errors for the cubic specification). Based on these statistics, we cannot reject the hypothesis of no specification errors for the lowest and highest educational groups but can reject it for the middle groups and all groups combined. However, given the large numbers of degrees of freedom, the tests are extremely powerful and such marginal rejections of the null hypothesis should be interpreted with caution. Put differently, the rejection for the overall specification of the quartic is based on the residual sum of squares exceeding its expected value by about 8%, and results for the permanent component imply that the persistent bias accounts for less than 3% of the residuals generating figure 6, which are themselves small relative to the bias from the quadratic. A more appropriate test is to examine the absolute size of the bias, as we do in the next table.

Table 6 provides the analogous calculations for the unweighted residuals for both the cubic and the quartic specifications. As in table 2 for the quadratic, the last two columns give estimated root mean squared total and permanent biases. Whereas the root mean squared total biases ranged from 3.2% to 5.1% for the quadratic, they are only 1.1%–2.8% for the cubic (2.2% overall). For the quartic the estimated total root mean squared bias is actually negative for college graduates (i.e., the residual sum of squares is less than expected) and has a maximum of 1.3% for those with 8–11 years of schooling (1.2% overall). For the root mean squared permanent bias, the cubic yields biases of from 1.8% to about 2.1% (1.9% overall), and the quartic yields biases of between .7% and 1% (.8% overall). These compare favorably with the permanent bias of between 3.3% and 4.5% (4.1% overall) obtained for the quadratic.

The quartic specification also reduces the systematic nature of the bias in addition to its magnitude. The correlations across schooling groups for the quartic range from −.05 to .56 with half of the correlations being less than .1. These compared favorably with correlations of .66–.85 for the cubic and the correlations of .84–.94 listed for the quadratic in table 3.

The autocorrelation of the permanent bias component also shows a substantial improvement for the quartic over either the cubic or the quadratic. The estimated autocorrelations of the bias terms range from −.09 for those with 8–11 years of schooling to .39 for college graduates, with the autocorrelations being negligible except for college graduates. These compare with correlations ranging from .46 to .91 for the cubic and .70 to .85 for the quadratic.

Based on these results, the statistics in tables 5 and 6, and the residual plots in figure 6, our general perception is that the quartic fits the data reasonably well and eliminates nearly all of the bias generated by the quadratic (over 96% of the permanent bias in the quadratic specification is captured by the quartic). However, moving to the unrestricted quartics used to generate the residuals in figure 6 uses up two additional degrees of freedom per year per educational group (for a total of 192 degrees of freedom), while the stability of the bias over time (as implied by the large permanent bias component from the quadratic) and the stability of the bias across educational levels (as implied by the high correlation and the similarity of the residual plots for the quadratic in fig. 4) would seem to imply that the added flexibility of two additional parameters for each year and schooling group pair is not necessary. Rather, it would seem that a time and group invariant correction may go a long way toward eliminating biases found in the quadratic specification.

While there are many ways to search for such a correction, we decided to examine a quadratic in a quadratic as an alternative. Hence, we examine equations of the form

$$y_{ixt} = a_{0it} + a_{1it} z_{ixt} + a_{2it} z_{ixt}^2 \tag{17a}$$

and

$$z_{ixt} = c_0 + c_1 x + c_2 x^2, \tag{17b}$$

where the parameters in (17a) are allowed to vary over time and across schooling groups as in the quadratic specification, and the parameters in (17b) are constant over time and across groups. All but one of the parameters in equation (17b) are redundant, so that (17b) can be simplified to

$$z_{ixt} = x + \gamma x^2. \tag{17b$'$}$$

Unlike the unrestricted quartic that added 192 parameters to be estimated, this specification adds only one. Moreover, once γ has been estimated, the specification in (17a) is as parsimonious as the quadratic. The only difficulty is that after substituting for z in equation (17a) we have

$$y_{ixt} = a_0 + a_1 x + (a_2 + a_1\gamma)x^2 + 2\gamma a_2 x^3 + \gamma^2 a_2 x^4, \tag{18}$$

which is a quartic with a nonlinear parameter restriction. Using nonlinear least squares we select $\gamma = 1/60$ as a reasonable value to impose across schooling groups and over time.[5] Using $\gamma = 1/60$, we then estimate equation (17a) by weighted least squares. These specifications have the same functional form as the quadratic estimates presented in Section II, except the quadratic is in terms of the transformed measure of experience as given in (17b) rather than years of experience per se.

The average residuals from the nested quadratic specification are given in figure 7. While not quite as good as the quartic, the residuals from the nested quadratic look significantly better than the residuals from the cubic, which uses significantly more (95) degrees of freedom. The test statistics used to evaluate the quadratic, cubic, and quartic specifications are presented in tables 7 and 8 for the nested quadratic. In terms of mean squared total bias, the overall level is .155 compared to .214 for the cubic (a reduction of 28%). In terms of permanent bias (our major concern) the nested quadratic does much better with a mean squared bias of .063 versus .151 for the cubic (a reduction of 58%). Looking at the results for the unweighted residuals in table 8, we see that the overall root mean squared bias for the

[5] As should be clear from eq. (2), the peak of the earnings function occurs at $x = 1/2\gamma$, so with $\gamma = 1/60$ a peak in earnings occurs at 30 years of experience, which seems roughly consistent with the unrestricted average profiles given in fig. 2.

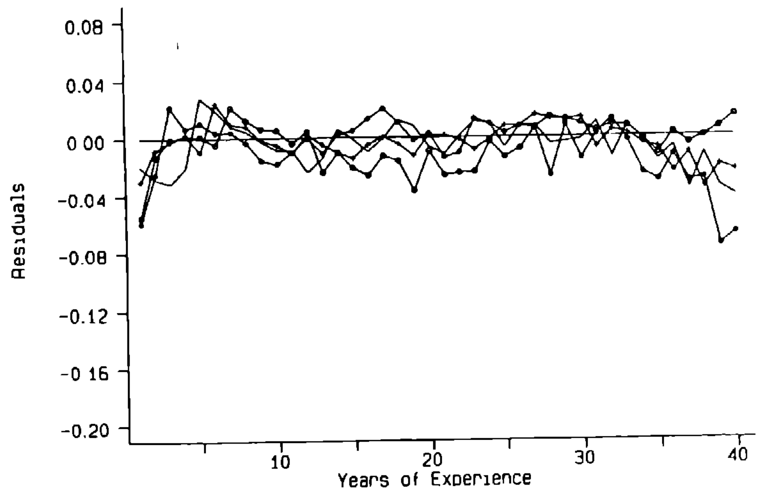

FIG. 7.—Residuals from nested quadratic specification: 8-11 years of schooling, + = high school graduates; □ = 13-15 years of schooling, plain line = college graduates

nested quadratic is 1.9% versus 2.2% for the cubic. Once again, the improvement is greater for the permanent bias component, with a root mean squared permanent bias of 1.4% for the nested quadratic versus 1.9% for the cubic, an improvement of over 25%. Of course, both the cubic and the nested quadratic represent an enormous improvement over the 4.1% bias component for the quadratic.

Table 7
Bias Calculations for Weighted Residuals from Nested Quadratic Specification

Educational Group	Mean Squared Residual (1)	Expected Mean Squared Residual (2)	Estimated Total Bias Component* (3)	Estimated Permanent Bias Component† (4)	χ^2 (5)
8–11	1.10	.925	.177	043	1,056 (888)
12	1.12	.925	.198	.095	1,075 (888)
13–15	1.07	.925	.145	048	1,027 (888)
16+	1.03	925	100	065	989 (888)
Overall	1.08	.925	155	.063	4,147 (3,552)

NOTE.—Degrees of freedom are in parentheses
* Calculated as mean squared residual (1) minus expected mean squared residual (2)
† Calculated as in eq. (16)

Table 8
Root Mean Squared Bias Calculations for Unweighted Residual from Nested Quadratic Specification

Educational Group	Mean Squared Residual (1)	Expected Mean Squared Residual (2)	Total Mean Squared Bias (3)	Mean Squared Permanent Bias (4)	Total Root Mean Squared Bias (5)	Root Mean Squared Permanent Bias (6)
8–11	.0030	0025	0005	0002	.023	013
12	.0012	0010	0002	.0001	015	011
13-15	.0038	.0033	.0005	0001	023	012
16+	0038	.0037	.0002	0004	0123	.020
Overall	.0029	0026	.0004	.0002	019	014

The nested quadratic also does better than the cubic in terms of the systematic nature of the bias. The correlations across schooling groups range from −.01 to .59 versus .66 to .85 for the cubic, and the estimated autocorrelations of the bias terms across experience levels range from .28 to .66 (.45 overall) versus .46 to .91 for the cubic (.57 overall). Hence, it appears that the nested quadratic does significantly better than the cubic in terms of all our measures but uses only three rather than four degrees of freedom per year/schooling category. Relative to the quadratic, which uses the same number of degrees of freedom, the nested quadratic looks excellent and is actually much closer in performance to the free-form quartic than to the quadratic.

Based on these results, our conclusion is that the quartic provides a reasonably good approximation to the "true" earnings function and that the nested quadratic provides a suitable alternative when parsimony is required or degrees of freedom are at a premium. In the next section, we describe some of the differences that these alternative specifications can make for inference.

IV. How Much Difference Does It Make?

As we stated in the introduction, the quadratic specification has served two distinct roles in recent labor market research. First, the quadratic is used as a framework for analyzing patterns of career earnings growth and the human capital investment process that underlies the earnings growth process. Second, the quadratic specification has gained widespread acceptance when the researcher wishes to "control" for career earnings growth when examining other factors that affect labor earnings.

In terms of career earnings growth, our basic finding is that the quadratic specification is unacceptable. Figures 8, 9, and 10 illustrate why. In figure 8, we plot actual earnings growth and predicted earnings growth from the quadratic specification for each of the four schooling groups. Figure 9 gives the corresponding plot for the quartic specification. Finally, figure

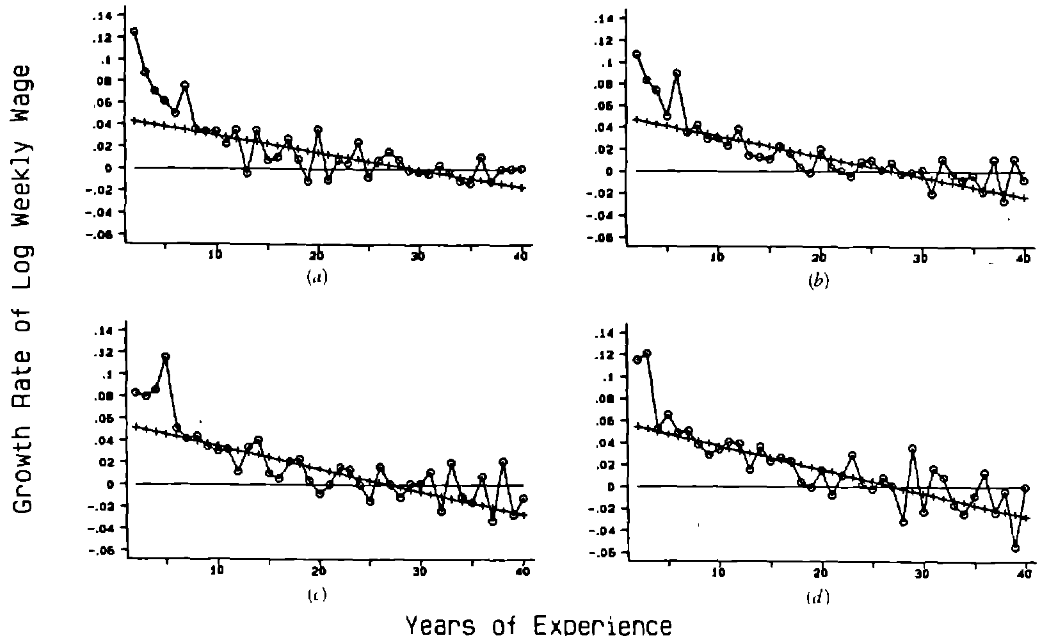

FIG. 8 —Quadratic and actual earnings growth profiles. *a*, 8–11 years of schooling, *b*, high school graduates, *c*, 13–15 years of schooling, *d*, college graduates.

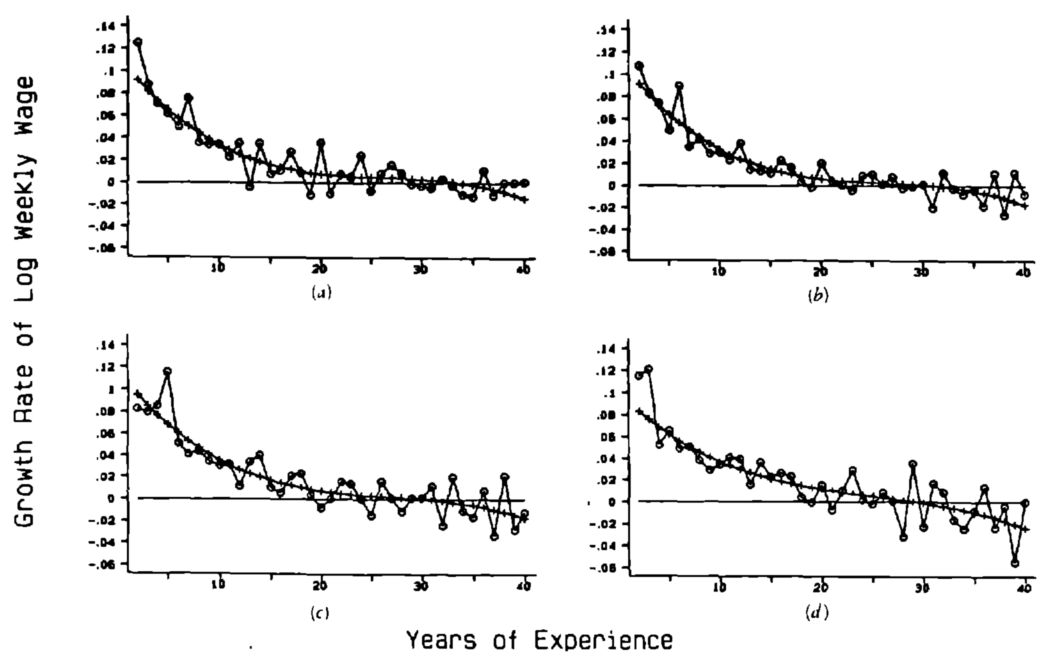

FIG. 9 —Quartic and actual earnings growth profiles: *a*, 8–11 years of schooling; *b*, high school graduates, *c*, 13–15 years of schooling, *d*, college graduates

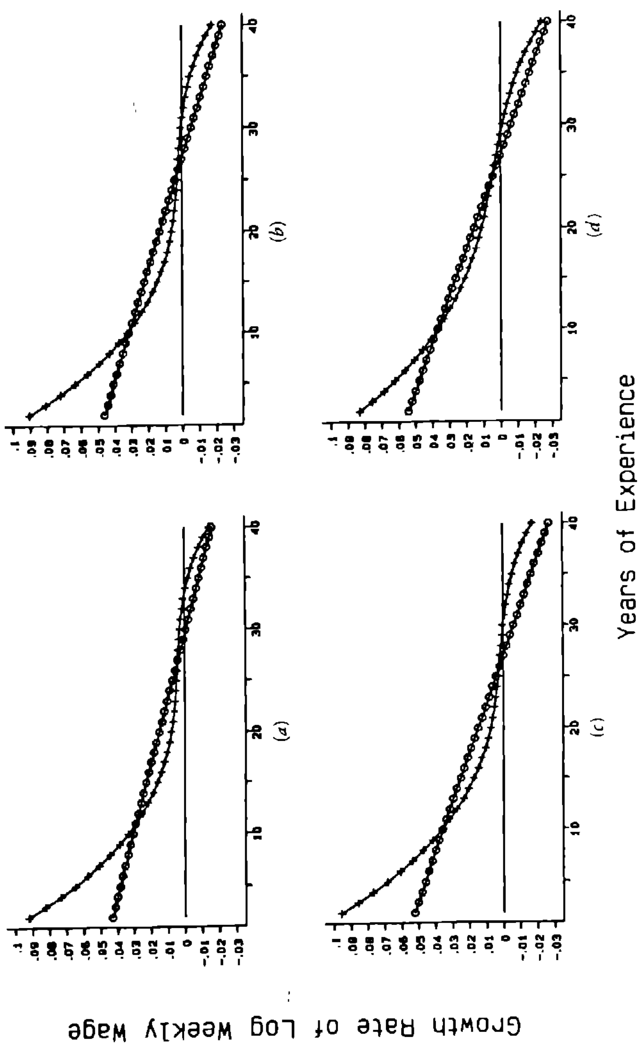

FIG. 12.—Quartic (+) and quadratic (O) earnings growth profiles: *a*, 8–11 years of schooling; *b*, high school graduates; *c*, 13–15 years of schooling; *d*, college graduates

10 compares predicted growth from the quadratic and quartic specifications. Figures 8 and 9 illustrate that the quartic fits the earnings growth pattern well, while the quadratic misses significantly. Figure 10 compares the growth profiles for the quadratic and quartic specifications and shows the quartic performs much better. The quadratic specification constrains earnings growth to decline linearly over the career. However, as is clear from figure 10 (and fig. 8), earnings growth declines much more rapidly in the early career than in either the mid- or late career. It is not the assumption that earnings growth declines over the career that is rejected by the data; rather, the assumption of a *constant* rate of decrease is what is at odds with the evidence. Since the human capital investment process predicts only a declining rate of earnings growth, the quartic and other specifications are consistent with the basic components of Mincer's analysis. Only the arbitrary assumptions of constant rates of return and linearly declining investment are rejected.

The picture of life-cycle earnings growth emerging from the quadratic profiles in figures 8 and 10 is modified considerably by the more flexible specifications. For example, the quartic predicts earnings growth of 8%–10% per year at 1 year of experience, while the quadratic predicts career growth of only 4%–5.5%. In fact, for all schooling levels except college graduates, actual earnings growth *and* quartic predictions of growth are higher for each of the first 5 years than the quadratic predicts for year 1.

Table 9 illustrates the impact of specification errors on estimates of career earnings growth by dividing the life cycle into three phases, the early career (years 1–10 in the first panel), midcareer (years 10–25 in the

Table 9
Estimates of Career Earnings Growth from Average Profiles

Educational Group	Actual Growth (1)	Quadratic Prediction (2)	Quartic Prediction (3)	Nested Quadratic Prediction (4)	Quadratic Prediction Relative to Actual Growth* (5)
Years 1–10:					
8–11	775	391	710	671	51
12	713	419	692	.670	59
13–15	761	482	.758	.742	63
16+	.743	.509	.681	.673	.68
Years 10–25:					
8–11	.206	311	.211	229	1 51
12	198	.287	183	182	1.45
13–15	213	325	215	211	1 53
16+	293	345	275	.266	1 18
Years 25–40:					
8–11	003	−.076	022	024	−23.3
12	− 045	− 148	− 058	- 014	3.32
13–15	− 051	− 171	−065	- 018	3.36
16+	−.098	−.172	− 104	132	1.76

* Column 5 is col. 2/col. 1

second panel) and late career (years 25–40 in the third panel). The rows in each panel refer to the educational classes. The first column gives actual earning growth calculated from the average profile, while columns 2–4 give predicted growth from the quadratic, quartic, and nested quadratic, respectively. The ratio of growth predicted by the quadratic to actual growth is given in column 5. The discrepancies between the quadratic and either the actual or quartic predictions are striking. Growth predicted by the quadratic for the first 10 years varies from a low of 51% of actual growth for those with 8–11 years of schooling to a high of 68% of actual growth for college graduates. For the midcareer, predicted growth exceeds actual growth for all schooling groups with a maximum discrepancy of 52% for those with 13–15 years of schooling. Finally, the decline in earnings predicted by the quadratic over the last 15 years is only an artifact of the quadratic specification. For high school graduates, the prediction of a 15% decline is more than three times the actual 4.5% decline. To illustrate the distortion, we note that a researcher armed with the estimates from the quadratic would conclude that just under 60% of the earnings growth in the first 25 years comes in the first 10 years. In contrast, calculations based on the raw data imply that almost 80% of the growth in earnings comes in the first 10 years. Based on results like these, one can only conclude that the quadratic gives a completely unsatisfactory picture of earnings growth and that the use of an alternative specification is definitely required.

While it seems clear that the quadratic must be scrapped for purposes of estimating career earning patterns, it is unclear whether the quadratic can still be used to effectively "control" for life-cycle wage effects when other factors affecting wages are of primary interest. On these matters we can provide no clear answers, only some words of caution. First, the biases in the quadratic cannot help and can only create problems. Second, how severe the problems are will depend crucially on how much the variables of interest vary within experience levels. One should keep in mind that the difference in the bias for the quadratic between those with 1–5 years of experience and those with 5–15 years of experience is about 10%. Since we often look for wage effects on the order of a few percentage points, if even a fraction of this misspecification showed up in a bias term for a regression coefficient, it could significantly bias the results. Moreover, the results on wage growth would seem to imply things look even more ominous if the quadratic is used to control for wage growth. In our opinion, prudence would require the use of one of the alternative specifications we recommend in either case.

References

Berger, Mark C. "Cohort Size and the Earnings Growth of Young Workers." *Industrial and Labor Relations Review* 37 (1984): 582–91.

———. "The Effect of Cohort Size on Earnings Growth: A Reexamination of the Evidence." *Journal of Political Economy* 93 (1985): 561–73.

Freeman, Richard B. *The Overeducated American.* New York: Academic Press, 1976.

———. "The Effect of Demographic Factors on Age-Earnings Profiles." *Journal of Human Resources* 14 (1979): 289–318.

Heckman, J. J., and Polachek, S. "Empirical Evidence on the Function Form of the Earnings-Schooling Relationship." *Journal of the American Statistical Association* 69, no. 346 (1974): 350–54.

Lillard, L.; Smith, J.; Welch, F. "What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation." *Journal of Political Economy* 94 (1986): 489–506.

Mincer, Jacob. *Schooling, Experience, and Earnings.* New York: National Bureau of Economic Research, 1974.

Murphy, K.; Plant, M.; and Welch, F. "Cohort Size and Earnings in the United States." In *Economics of Changing Age Distributions in Developed Countries,* edited by R. Lee et al. Oxford: Clarendon, 1988.

Murphy, K., and Welch, F. "The Structure of Wages." Unpublished manuscript. Los Angeles: Unicon Research Corporation, 1988.

Welch, Finis. "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust." *Journal of Political Economy* 87 (1979): S65–S97.