Volume Title: Annals of Economic and Social Measurement, Volume 6, number 4

Volume Author/Editor: NBER

Volume Publisher:

Volume URL: http://www.nber.org/books/aesm77-4

Publication Date: October 1977

Chapter Title: The Covariance Structure of Earnings and the On-the-Job Training Hypothesis

Chapter Author: John C. Hause

Chapter URL: http://www.nber.org/chapters/c10522

Chapter pages in book: (p. 335 - 365)

Annals of Economic and Social Measurement, 6/4, 1977

THE COVARIANCE STRUCTURE OF EARNINGS AND THE ON-THE-JOB TRAINING HYPOTHESIS*

BY JOHN C. HAUSE

Human capital investment and its returns are basically intertemporal processes. However, most empirical studies of these processes have been limited to cross-sectional data that provide little or no information on the actual linkage of earnings over time as manifested in individual earnings profiles. In this study a statistical methodology is developed for analyzing cohort time series data on earnings, and it is then applied to a simple model designed to throw some light on the potential empirical importance of the hypothesis that systematic differences in on-the-job training (OJT) lead to significant differences in individual earnings profiles. The study illustrates the central role played by the covariance structure of earnings when these hypotheses are explored with time series data. The statistical analysis leads to reasonable upper-bound estimates of the dispersion of earnings profile slopes that indicate empirically relevant systematic differences in the profiles. Some evidence is also presented that indicates there is significantly less relative variance in discounted earnings profiles than in relative earnings for a single year. Both conclusions are consistent with the OJT model.

1. INTRODUCTION

The determination and explanation of statistical properties of the earnings profile are becoming focal points of research by economists interested in the distribution and life cycle history of earnings. The first-order statistics (profiles of mean earnings as a function of labor force experience or age) have been studied extensively, both because of their central role in estimates of the returns from investments in human capital and because the requisite data have been relatively easy to obtain. Synthetic profiles of earnings for groups classified by years of education and adjusted for dif-

*Initial research on this topic was funded by a National Science Foundation grant (GS-31334) to NBER for research on the determinants of the distribution of income and earnings. Additional work and completion were supported by National Institute of Education (U.S. Office of Education) Grant OEG 0-72-1569, NIE Project 2-0680. None of these institutions has officially reviewed or approved this work or its conclusions. I am greatly indebted to Chris Sims for comments that have deeply influenced my thinking about the theoretical and statistical implications of the basic model, and to Urban Norlén for resolving much of the generalized least squares (GLS) estimation procedure and carrying out the GLS and parametric calculations. L. A. Lillard and Yoram Weiss provided several very useful technical and expository comments when they read an earlier version. It is a pleasure to acknowledge the cooperation of Sten Johansson and Ingalill Eriksson, who made it possible for me to use a special data tape created by the latter for the Låginkomstutredning study in Sweden. Thanks are also due to Göran Ahrne, who assembled the income data for those born in 1936–1941. The study has also benefited from the valuable research assistance of Karlis Goppers, and from work by Eilon Amit, Lars Anlér, Edward Fagerlund, and Lars Elve Larson, who carried out computations at various stages. It has also been improved significantly by the careful editing of Ester Moskowitz. I alone am responsible for conclusions and errors, but because of the complexity of the income profiles the empirical ambiguity can be partly attributed to the data.

ferences in personal characteristics have been constructed from cross-sectional data. These profiles form the primary empirical basis for estimating internal rates of return from schooling. Economists have generally interpreted the inverted U shape of the profiles as reflecting the growth of productivity with experience (and investment), which is increasingly counterbalanced by depreciation and obsolescence. Mincer (1970, 1974) has strongly emphasized the potential significance of post-school investments in increasing an individual's earnings capacity; the primary focus of his study has been on the role of on-the-job-training (OJT).

Data on the mean earnings profile of a cohort combined with statistics of the variance about that profile (as a function of experience or age) are adequate for raising and resolving some significant questions, but leave unanswered the central issue of how the individual profiles are distributed about the group mean. That issue is important for at least two reasons: First, investments in human capital usually modify the entire subsequent earnings profile. The dispersion of earnings for a single year is due to a combination of systematic, transitory, and compensatory components that are difficult to disentangle and to relate to specific investments by the individual. The risk due to different payoffs from such investments is much more adequately reflected by the dispersion of lifetime earnings (e.g., as measured by the variance of the logarithm of the discounted earnings stream) than by the earnings dispersion of one year.[1] In principle, the variance of lifetime earnings can be partitioned into three distinct components, one containing systematic elements perceived by a person prior to an investment in human capital, another containing systematic elements not anticipated, and a third containing "random" factors. However, these questions and the question of lifetime earnings variance can be answered directly only with information on the autocovariance structure of earnings. An upper bound on the variability of lifetime earnings can be obtained consistent with knowledge of the standard deviation of earnings as a function of age, since the discounted value of the latter measure is the least upper bound of the standard deviation of (discounted) lifetime earnings.[2] Unfortunately, the lower bound of discounted earnings based

[1] This issue is discussed in greater detail in Hause (1974).

[2] The lower-bound estimate of zero is approached, for example, if working lifetime is divided into a large number of periods between which there is no correlation of earnings. The upper bound cited in the text is obtained from Hause (1974) on the basis of a proof sketched by Sims. Let x_t be earnings as a function of age, t; μ_t, the corresponding mean earnings as a function of age; and ρ the constant discount factor:

$$\left\{\int e^{-\rho t}[E(x_t - \mu_t)^2]^{1/2} dt\right\}^2$$

$$= \iint \{E[e^{-\rho t}(x_t - \mu_t)]^2\}^{1/2} \{E[e^{-\rho t'}(x_{t'} - \mu_{t'})]^2\}^{1/2} dt dt'$$

$$\geq \iint E[e^{-\rho t}(x_t - \mu_t) e^{-\rho t'}(x_{t'} - \mu_{t'})] dt dt'$$

on cross-sectional earnings data is zero. and these bounds are far too wide to be of much interest. Fragmentary data reported in past studies reveal that earnings over time have substantial autocorrelation, but it is considerably less than unity; so neither bound is satisfactory for our needs.

Second, data on individual earnings profiles can also help unravel the effect of human capital investments on the life cycle of earnings. At present, "optimal programs" of human capital investment are not well understood. partly because only very limited information is at hand on the substitutability and complementarity of the different investments. In studies of American data (e.g., Mincer 1974, Chap. 4) a strong positive association has been established between the level of schooling attainment and the slope of the mean earnings profile. as well as with the number of years over which the profile continues to have a substantial positive slope. The mean profiles do not reveal the extent to which schooling and on-the-job training may be partial substitutes: this information would be useful for measuring the importance of differences in investment opportunities in determining total human capital investment.

Direct empirical investigation of these topics is possible with individual data covering a long segment of the lifetime earnings profile. from which autocovariances can be computed. Although scattered pieces of evidence on autocorrelations of earnings and of income have been reported, e.g., Friedman and Kuznets (1954). Hanna (1948). Mendershausen (1946), Thatcher (1971), the populations on which the results are based are usually heterogeneous in age and education. Hence, these correlations are weighted averages of different segments of earnings (or income) profiles from cohorts with differing schooling attainments. and are consequently difficult to interpret. Finally, little seems to have been done systematically to exploit the covariance structure to estimate or explain lifetime earnings. aside from some work on the random walk hypothesis, e.g., by Fase (1970).

In the next section, the OJT hypothesis is considered in detail. and a statistical specification is given for its study. In the third section, a model is developed to provide possible tests for the existence of OJT effects and estimates of the differences in earnings profile slopes. Next, the parametric model is applied to time series data on income for a cohort of Swedish men with various levels of schooling attainment. A few results are also reported on the possible compensatory effects of OJT for a sample

(by Schwartz's inequality).

$$= E\iint e^{-\rho t}(x_t - \mu_t)e^{-\rho t'}(x_{t'} - \mu_{t'})\,dt\,dt'$$

$$= E\left[\int e^{-\rho t}(x_t - \mu_t)\,dt\right]^2$$

Taking square roots of the first and last steps yields the upper bound asserted in the text.

of American men. In a final section, future work that could be carried out within this framework is briefly discussed.

2. The On-the-Job Training Hypothesis

The OJT hypothesis asserts that systematic differences in OJT lead to significant differences in earnings profiles. The hypothesis has been discussed extensively by Mincer (1970), whose empirical work has been based exclusively on first-order statistics on earnings profiles. Two important empirical implications follow readily from the hypothesis. First, differences in OJT investment will generate dispersion in the slopes of earnings profiles. Second, part of the dispersion in earnings at a given age is a consequence of different investments in OJT: those who invest heavily in OJT are paid less at first than others of similar economic ability, but the earnings of the former subsequently rise rapidly enough and long enough to compensate them for their lower initial earnings. Hence, OJT is a systematic compensatory mechanism that tends to reduce the relative variation in lifetime earnings compared with the relative variation in annual earnings. This study is concerned primarily with the first implication. The time series data on which most of this study is based are too short to evaluate the net extent to which OJT plays a compensatory role in reducing the variance of lifetime earnings. However, the main sample, described in detail below, does provide data that yield empirically significant upper-bound estimates of OJT effects for low levels of schooling attainment; the findings indicate there are systematic and important differences in the speed with which individual earnings increase with experience. Those findings are compatible with the OJT model, but not with a model in which the evolution of personal earnings is generated primarily by random walk deviations from a common trend.

Mincer (1970, 1974) has suggested that a significant source of the differences in human capital investment arise from differences in OJT, which in turn lead to differences in earnings profiles. In some jobs, it takes a relatively long time for new workers to acquire normal levels of job skill. Workers entering such jobs may initially receive low earnings, corresponding to their low net productivity. As they acquire more experience and skill, their earnings rise. Since capital market exchange opportunities make current dollars more valuable than future dollars, it can be shown by means of a simple model that when the labor market is in equilibrium, future earnings must be high enough to offset the low initial earnings in those jobs (in the sense of equalizing present values). Apprenticeships for certain crafts or the establishment of a professional reputation in medicine or law are examples of the kinds of significant post-school investment required in some jobs.

Consider a model in which there is a perfect capital market, perfect

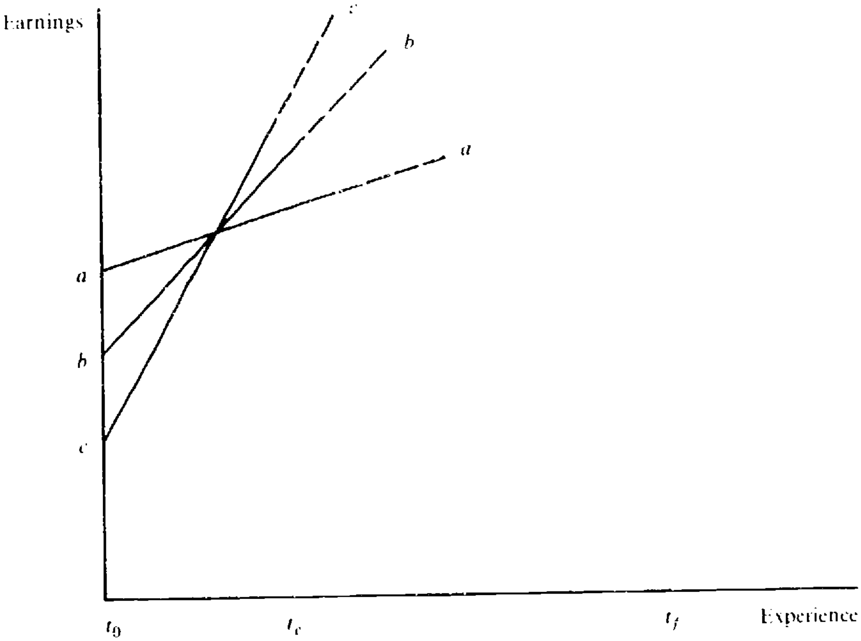

Figure 1 Linearized Rising Sector of Earnings Profile of Equal-Ability Individuals

foresight, no nonpecuniary occupational preferences, and equal earnings potential for a cohort of entrants into the labor force, whose members have the same schooling attainment. In Figure 1, earnings profiles are illustrated for three jobs, *a*, *b*, and *c*, requiring increasing levels of post-school OJT investments. If it is assumed that no additional out-of-pocket investments are required for any of these jobs, then the optimizing behavior of individuals would generate a labor market equilibrium in which the earnings profiles have the same present value; i.e., $\int_{t_0}^{t_f} e^{-\rho\tau} y_i(\tau)\, d\tau$ is the same for $i = a, b, c$ where $y_i(t)$ is the earnings profile for job i, ρ is the market interest rate, and t_0 and t_f are the respective dates of entry into and exit from the labor force.

There is little theory to guide us in selecting a particular function, $y_i(t)$, for the earnings profile, and there is no theoretical necessity for the family of profiles to be linear in the rising portion. Like cost functions, the earnings profiles presumably depend on technology and prices. Mincer's own work (1974, p. 17) with simple, analytically convenient earnings functions often leads to a "crossover" (or "overtaking") point t_c (or to a relatively narrow crossover interval) for given schooling attainment, and I retain this assumption in the following development.[3] In principle,

[3]The crossover concept is particularly relevant for discussion of the compensatory role of OJT, in which investment costs are subsequently offset by higher earnings.

there is little formal justification for assuming that earnings profiles of people with equal ability cross each other after the same number of years of job experience, nor is there much reason for assuming that this same amount of time would be appropriate for different equal-ability cohorts. If we have a set of equal-ability people with earnings profiles $y_i(t)$ that have the same present value but a large dispersion in initial earnings $y_i(t_0)$, reflecting differences in initial OJT, there is no mathematical necessity for all the $y_i(t)$'s to intersect at the same time, t_c (although it might be a reasonable simplification to assume that each pair of profiles has only one intersection). If data were available for measuring earnings profiles of individuals possessing similar economic ability, it would be possible statistically to characterize the strength of the central tendency of profiles of equal-ability people to cross at the same point. In the statistical model I develop below, I make the restrictive assumption of a single crossover point because the model is then easier to manipulate and interpret. Subsequently, I discuss the precise role played by that assumption in my statistical tests.

The assumption of a linear profile is a convenient fiction which is probably not unduly restrictive. Since we are concerned primarily with the extent to which earnings profile slopes differ, for most of the statistical work we can regard the linear profile segments as deviations from the mean earnings profile for the cohort (which has the usual skewed, inverted-U shape with the ascent in the early years considerably steeper than the descent in the later years). Furthermore, the time series data I analyze here span only six or seven years rather than the entire earnings profile.

The empirical relevance of the model clearly depends on the extent to which there are systematic differences in the earnings profiles that are due to economic choices made by workers. If the profiles merely reflect inevitable and unalterable increases in productivity that accompany job experience and physical and psychological maturation and aging, there is little theoretical gain from describing the profile as if people were making investment decisions in OJT once they entered the labor force. The scenario that accompanies Mincer's formal OJT model (1974) might be interpreted as if this form of human capital investment required a worker to make optimal period-by-period investment decisions; if the worker decided to hold his stock of human capital constant, his observed earnings would equal potential earnings and his earnings profile would be horizontal. It is not possible to infer the range of OJT investment choices available to or actually chosen by workers from the mean earnings profile of a cohort with given schooling attainment. However, the finding of a crossover of mean profiles of earnings for different occupations, given schooling, with the initially lower profile remaining higher after the crossover, would suggest that the choice of a job carries with it an implied choice of earnings profile. In that case, competition for jobs should tend to equalize returns at the margin, with the result that there will be no net advantage of one profile over the other for workers with similar tastes and economic ability. Even this information would not indicate what amount of on-the-job investment workers can elect once they have chosen a job.

Much of the remainder of this paper is devoted to developing and testing a model that can indicate the maximum economic significance that OJT could have by determining the extent to which there are systematic differences in the growth rate of earnings with experience, and to developing an indirect test of the possible net effect on OJT, based on partial correlations of earnings. There may, of course, be other reasons besides OJT for differences in the slopes of earnings profiles, for example, differences in ability or unanticipated excess demand for a particular occupation. Time series data are not available, however, for a more complete specification of the model– by including personal characteristics, for example, which may themselves be determined in part by previous investments in human capital– so that it could be used to disentangle OJT effects from other distinct factors related to differences in the slopes of earnings profiles. Hence, the dispersion in slopes described in this paper provides an upper bound on OJT effects for the sample cohorts studied.

The simple model of earnings profiles illustrated in Figure 1 has several implications for correlations of earnings from different points along the profile. In the following discussion it is assumed that time periods i and j both lie in the time interval where the earnings profiles are rising and that $i < j$. If $i, j \gtrless t_c$, then $r_{ij} > 0$. If either i or j is equal to t_c, $r_{ij} = 0$ [in a degenerate sense, since the variance of $y(t_c) = 0$ in this model], and if $i < t_c < j$, $r_{ij} < 0$.

The model in this form is clearly inadequate for empirical work. First, there is no t_c such that var $y(t_c) = 0$ and steadily increases as it moves in either direction in time away from t_c. Second, fragmentary evidence suggests that cohorts of individuals who have the same schooling attainment and who entered the labor force simultaneously have positive earnings correlations for all relevant i's and j's. These empirical features may arise because there are substantial differences in the potential economic capacity of people at the time they enter the labor force, whereas in the model, it is assumed that everyone has the same initial economic ability. Variation in initial economic capacity can be formally incorporated into the model by assuming a distribution of earnings at crossover time and by assuming further that those with the highest earnings at that time have the highest (discounted) *lifetime* earnings. If the variance of this earnings distribution is large enough, it could mask the simple patterns of earnings correlations implied by the initial OJT model. Last, there is presumably substantial residual variability in earnings that must be taken into account.

These considerations lead to the following statistical specification of the (linearized) earnings profiles:

$$x_t = m + tw + u_t \tag{1}$$

In this equation, x_t represents earnings in period t, and time has been translated so that the crossover period $t_c = 0$. m, w, and u_t are all random variables. u_t, the earnings residual in period t, is assumed to be uncorrelated with m and w for all t. m is assumed to yield the distribution of earnings at the crossover period t_c ($= 0$) in the complete absence of u_t, and represents the distribution of earnings differentials due to differences in economic ability (or earnings potential) of individuals before they enter the labor force; m is supposed to capture genetic and environmental differences in background as well as earlier investments in human capital. w determines the slope of the rising portion of the linearized earnings profile, and is the key element in this model, linking it with the OJT hypothesis: w is determined by individuals by their occupation and the amount of OJT they decide to acquire (to the extent that OJT is a choice variable, given the occupation an individual enters). If there are large, systematic differences in the OJT obtained by different workers in a sample, these differences should be reflected by a correspondingly large dispersion of the distribution of w for the sample members. m and w have no time subscript (t), since they are assumed to be specific to the individual, and do not vary along the early segment of the profile.

There is little theory to guide us in specifying the statistical properties of the earnings residual, u_t, and in the following, I will assume that $u_t = y_t + z_t$; y_t is a random walk component:

$$y_t = \sum_{i=t_0}^{t} \epsilon_i,$$

where ϵ_i represents the independent (nontransitory) random shock: $\sigma_{\epsilon_i\epsilon_j} = 0$ if $i \neq j$, and $= \sigma_{\epsilon_i\epsilon_i}$ if $i = j$. z_t represents a transitory earnings residual in period t, and I assume $\sigma_{z_iz_j} = 0$ if $i \neq j$, and $= \sigma_{z_iz_i}$ if $i = j$. Finally, $\sigma_{\epsilon_iz_j} = \sigma_{\epsilon_jz_i} = 0$ for all i and j.

Some students of earnings profiles have assumed that changes in earnings along the profile can be represented by a random walk, which correponds to random but nontransitory factors that permanently change the expected level of earnings (e.g., see Fase 1970 for extensive work with this model). The element y_t captures this aspect. The concept of "transitory" variations ("white noise") in earnings (and income) has long been used in the theoretical and econometric specification of earnings profiles, and z_t represents this component. It incorporates minor accidents, incidental unemployment, and the like, which are assumed to exert an effect on earnings for only a single period.

No assumptions have yet been made about σ_{mw}, the covariance between the distribution of earnings at crossover time t_c (in the absence of residual earnings variations arising from u_t) and the distribution of slopes of the earnings profiles. It seems plausible that high-ability people will have higher earnings at t_c than people of low ability *and* may also have an earnings profile with a steeper slope because of their capacity to acquire certain job-related skills more rapidly. Previous work by Hause (1972) provides some empirical evidence that a direct measure of cognitive ability (from test scores) is positively correlated with the slope of the earnings profile. At the very least, it seems reasonable to assume $\sigma_{mw} \geq 0$, and I maintain this assumption in the following discussion. If we had precise knowledge of the crossover time, it would be possible to estimate this covariance directly from longitudinal data. However, it is impossible in specification 1 to identify σ_{mw} and t_c simultaneously.

The consequences of misspecifying t_c can be determined from the covariance structure implied by (1). If the age at which crossover is assumed to take place is underestimated, then the estimate of σ_{mw} would be too small or even negative, and vice versa if the crossover age is overestimated.[4]

Although there is no direct way to observe t_c, Mincer (1974, p. 17) has shown, for one particular specification of the on-the-job investment

[4] I wish to thank L. A. Lillard for simplifying the notation used in this derivation and pointing out an earlier lapse:

$$x_{t^*-\tau} = m + (t^* - \tau)w + u$$

$$= (m - \tau w) + t^* w + u$$

(2) $$= m^* + t^* w^* + u^*$$

Hence

(3a) $$\text{var}\, u^* u^{*\prime} \equiv [\sigma^*_{uu}] = [\sigma_{uu}]$$

(3b) $$\sigma_{m^*m^*} \equiv \sigma^*_{mm} = \text{var}\,(m - \tau w) = \sigma_{mm} - 2\tau\sigma_{mw} + \tau^2\sigma_{ww}$$

(3c) $$\sigma_{w^*w^*} \equiv \sigma^*_{ww} = \sigma_{ww}$$

(3d) $$\sigma_{m^*w^*} \equiv \sigma^*_{mw} = \sigma_{mw} - \tau\sigma_{ww}$$

and

(3'b) $$\sigma_{mm} = \sigma^*_{mm} + 2\tau\sigma^*_{mw} + \tau^2\sigma^*_{mw}$$

(3'd) $$\sigma_{mw} = \sigma^*_{mw} + \tau\sigma^*_{ww}$$

If the crossover is misspecified by underestimating the age at which it is assumed to take place, τ is positive. Equations 3 then imply that the covariance matrix of residuals is unchanged, as is the scalar σ^*_{ww}. However, the apparent crossover variance σ^*_{mm} is too large if $\tau^2\sigma_{ww} > 2\tau\sigma_{mw}$. Given our assumption that $\sigma_{mw} \geq 0$, this occurs if $\tau > 2\sigma_{mw}/\sigma_{ww}$ (for $\tau > 0$), and is always true if $\tau < 0$. Furthermore, the apparent covariance of m and w, σ^*_{mw}, is too small (from 3b). Indeed, a sufficiently large underestimate of the crossover age could make the apparent σ^*_{mw} negative, even if the true $\sigma_{mw} < 0$, as we assumed earlier. Conversely, if the crossover age is overestimated (τ negative), the apparent σ^*_{mw} is too large.

function, that $1/r$ is an upper bound for t_c, where r is the rate of return on investment in OJT. The key assumption in his derivation is that OJT investment is a decreasing function of time, and that current earnings are equal to the sum of earnings obtainable without OJT plus the product of the interest rate multiplied by accumulated human capital and minus current investment; the sum of the last two components is then set equal to zero.

If the sample data include a direct measure of ability as well as time series data on earnings, correlations of estimated m's with the direct measure should be interpreted with caution because an error in specifying t_c will lead to an error in estimating the "true" crossover m's. However, since misspecification of t_c will not affect w, neither will it affect correlations between w and directly measured ability.

3. Statistical Testing of the OJT Hypothesis

In this section two procedures are considered for determining the possible existence and empirical significance of investment in OJT. The first procedure is an extension of the test used for determining the pattern of simple correlations of earnings when there are no residual variations in earnings or differences in ability. The second is an attempt to estimate the dispersion of the w's from individual time series of earnings.

3.1 *Testing OJT Effects by Partial Correlations of Earnings*

In section 2, we considered the distribution of m as reflecting differences in economic ability that would be observed at t_c in the absence of residual earnings variation, u_t. With that condition in mind, we look at the partial correlation of earnings $r_{ik \cdot t_c}$, where $i < t_c < k$. If we think of observed earnings at t_c as measuring (with error) economic ability, then we would expect this partial correlation to be negative, in analogy with the simple correlation of earnings, $r_{ik} < 0$, for a cohort of equal-ability people, if $i < t_c < k$. Since we have no direct information on the precise year of experience at which t_c occurs, we are led to consider the sign of the partial correlation of earnings $r_{ik \cdot j}$, where $i < j < k$. This problem is equivalent to determining the sign of the regression coefficient of x_i in a regression of x_k on x_j and x_i, which in turn depends on the sign of the determinant:

$$D = (\sigma_{x_i x_k} \sigma_{x_j x_j} - \sigma_{x_i x_j} \sigma_{x_j x_k})$$

The earnings variance and covariances in D can be decomposed into variance and covariance components of m (economic ability), w (slope of the earnings profile), and u (the random disturbance in earnings) through equation 1. If we assume the random term is not present, then, as shown in Appendix A, $D = -(\sigma_{mm}\sigma_{ww} - \sigma_{mw}^2)(j - i)(k - j)$. D (and thus $r_{ik \cdot j}$)

is negative, unless m and w are perfectly correlated, in which case $D = 0$. This result follows from the assumption $i < j < k$, and from the relation $\sigma_{mm}\sigma_{ww} - \sigma^2_{mw} = \sigma_{mm}\sigma_{ww}(1 - r^2_{mw}) > 0$ (for $r_{mw} < 1$). It verifies the intuitive argument that $r_{ik \cdot j}$ should be negative by analogy with the condition $r_{ik} < 0$ if $i < t_c < k$, when the variances of m and of residual earnings are zero, a conclusion that follows from the simplest OJT model illustrated in Figure 1. The demonstration that $r_{ik \cdot j}$ must be negative does *not* depend on the existence of a central crossover time. Basically, it implies a systematic difference in the parameters of the earnings profiles from individual to individual, which participants in the labor force are plausibly aware of when they make career choices. Nevertheless, we shall see that the crossover feature of the OJT model is very useful for determining conditions under which $r_{ik \cdot j} < 0$ when random variations are present, which is certainly the case in earnings data.[5]

Some restrictions on the autocovariance structure of the earnings residual, u_t, are required to draw further conclusions about the determinants of the sign of $r_{ik \cdot j}$ when we allow for residual variations in earnings. In the remainder of this section (3.1), I retain the assumption of section 2 that the earnings residual is the sum of a random walk, y_t, and a "white noise" component, z_t. In Appendix A I carry out the straightforward but slightly tedious calculation of determinant D under this assumption, and reach the following conclusions: (i) If there is no systematic dispersion in the slopes of the earnings profiles in this model, then σ_{ww} and σ_{mw} are both zero, and $r_{ik \cdot j}$ is positive. (ii) The presence of residual earnings variation tends to make it more difficult to observe a negative $r_{ik \cdot j}$, and may overwhelm this negative component of the OJT mechanism. (iii) Given whatever residual variation is present in the earnings data, a negative value for $r_{ik \cdot j}$ is more likely to be observed if earnings data can be obtained at the crossover time, t_c, and for a given number of years on each side of it.

In summary, I present a simple qualitative test for statistical evidence consistent with the OJT hypothesis that systematic differences in earnings profile slopes exist. The test is to determine whether the partial correlation of earnings at three different experience levels ($r_{ik \cdot j}$) is negative. The only component in our model that can generate a negative $r_{ik \cdot j}$ is σ_{ww}, the systematic variance of earnings profile slopes. Since the negative partial correlation may be reduced or outweighed by variance components resulting from random changes in earnings, attention is paid to earnings data char-

[5]The relevance of the crossover assumption when random earnings variations are present becomes apparent if it is disregarded, and $x_t = m + tw + u_t$ is interpreted as a linearized approximation of deviations from the mean earnings profile, with $t = 0$ for the time individuals enter the labor force, and the m's regarded as deviations of the intercepts of the individual profiles from mean initial earnings. In this alternative framework, very little can be said a priori about the sign or magnitude of σ_{mw}, and there is no hint of how i, j, and k might be chosen to increase the likelihood of observing a negative $r_{ik \cdot j}$.

acteristics required to circumvent this problem as far as possible. Time series data on individual earnings or income are not abundant, and a major reason for developing this test is to provide a way of verifying whether systematic differences in earnings profile slopes are present when only fragmentary time series are available (specifically, earnings at three points in time).

3.2 *Testing for OJT Effects by Estimating* σ_{ww}

If sufficiently rich time series data on individual earnings are available, both the slope parameter of the individual earnings profile, w, and the constant, m, can be estimated from the simple regression:

$$x_t = m + wt + u_t \tag{1}$$

and the importance of one or two standard deviations of w in creating earnings differentials across individual earnings profiles is readily determined. Furthermore, quantitative estimates of the potential empirical significance of the OJT-generated systematic differences in earnings profile slopes can be made, instead of only the qualitative test for slope differences proposed in section 3.1 above. The latter procedure leads to an upper-bound estimate of the true standard deviation of w. The individual w's are estimated from short time series of earnings (in our work, seven years), and thus contain significant sampling variability. The variance of these estimated w's is the sum of the true variance of w and the variance of the sampling error. Since the sampling variance depends on the length of the individual time series of earnings, its size relative to the true variance of w does not decrease as the number of individuals in the sample increases.

Two alternative methods are considered for obtaining reasonable estimates of σ_{ww}, the variance of the earnings profile slope parameter. In the first method, sample information on the autocovariance structure of the earnings residual (u_t) is exploited without imposing prior restrictions on that structure. Since economic theory provides very little guidance in the selection of such restrictions, the main advantage of this approach is that we avoid setting arbitrary ones. The main disadvantage is that it provides no basis for distinguishing between the sampling variance and true variance of the w's. In the second method a priori restrictions are imposed on the autocovariance structure of u_t that are strong enough to identify the true variance of the w's.

If the residual u_t's had constant variance and no autocorrelation, OLS (ordinary least squares) and GLS (generalized least squares) estimates of the w's would be identical. But it is highly unlikely that the u_t's would have no autocorrelation. Hence, GLS estimates should be obtained, since they have less sampling variance than other linear unbiased estimates of the w's, and this property is important in getting a plausible upper-bound estimate of the true σ_{ww}.

Regression equation 1 is rewritten in vector form for the ith individual, with earnings time series of length T:

$$(1)' \qquad \underset{T\times 1}{x_i} = \underset{T\times 2}{Z}\ \underset{2\times 1}{\beta_i} + \underset{T\times 1}{u_i}$$

where x_i is the time series of a person's earnings, Z is the nonrandom matrix

$$Z' = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \tau_0 & \tau_1 & \ldots & \tau_{T-1} \end{bmatrix}$$

and $\beta_i' = [m_i, w_i]$, the constant and slope parameters of the ith individual's earnings profile. It is assumed that the covariance matrix $E(u_i u_i') = \Omega$ is identical for all individuals in the sample. From a theorem by Rao, approximate GLS estimates of the individual β_i's can be obtained by using the empirical covariance matrix of earnings

$$\hat{\Omega}^* = (n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})'$$

in lieu of the true (unknown) covariance matrix of the disturbances, Ω, in the GLS calculation.[6] Indeed, it is is not even necessary to estimate the

[6]Rao's theorem. (1967: see also Norlén, 1975 and Appendix B to the present study) states the following: Let the covariance matrix Ω^* have the form $\Omega^* = \Omega + Z\Phi Z'$, where Φ is an arbitrary symmetric positive definite matrix, and Ω is the true covariance matrix of the disturbances. Then GLS estimators with Ω and Ω^* are the same. A short proof and other details are given in Appendix B. It follows from Rao's theorem that the empirical covariance matrix of earnings $\hat{\Omega}^*$ may be used for approximate GLS estimates, since

$$E(\hat{\Omega}^*) = \Omega + Z[(n-1)^{-1}\sum_{i=1}^{n}(\beta_i - \bar{\beta})(\beta_i - \bar{\beta})']Z'$$

where $E(\hat{\Omega}^*)$ is the expected value of $\hat{\Omega}^*$ and $\bar{\beta}$ is the mean of the β_i's. Thus $E(\hat{\Omega}^*)$ satisfies the conditions of Rao's theorem.

For the formula for $\hat{\Omega}^*$ in the text, it is assumed that all observational vectors are complete. Since the data base in most available time series on earnings or income is modest, the computations cannot be restricted to complete observations. When there are missing observations, each element of $\hat{\Omega}^*$ has the form

$$\frac{\sum_{i=1}^{n_{jk}}(x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{(n_{jk} - 1)}$$

where the sum is only over the n_{jk} observations for which x_{ij} and x_{ik} are present (i is the subscript for an individual), and the means $\bar{x}_j$ and $\bar{x}_k$ are also based on this same set of observations. For simplicity, the slightly misleading but less cumbersome notation is retained in the text even when there are missing observational elements. However, the empirical matrix constructed from the incomplete observations is not necessarily positive definite.

individual parameters and then to compute their empirical covariance matrix. Instead, we may calculate this last matrix directly from the simple expression $(Z'\hat{\Omega}^{*-1}Z)^{-1}$. The lower right-hand corner of this 2×2 matrix is our estimate of σ_{ww}, the variance of the profile slopes across individuals, while the upper left-hand corner gives an estimate of σ_{mm}, the variance of the individual regression constants at the crossover time.[7]

In the absence of restrictions on Ω, it is not possible to say how much of the estimated variance of σ_{ww} is due to the true variance of w and how much to sampling variation in Ω; all that can be said is that this procedure provides a reasonable upper-bound estimate of the true variance of w. The lower-bound estimate of σ_{ww} in the absence of restrictions is zero and is attained if all the estimated variance is due to sampling variation.

In the alternative procedure, a priori restrictions are imposed on Ω. I have already shown that the expected value of the empirical covariance of earnings $E(\hat{\Omega}^*) = \Omega + Z\Phi Z'$, where the components of the 2×2 matrix Φ are associated with the covariance matrix of the β_i's, i.e., $\Phi = (n-1)^{-1} \sum_{i=1}^{n} \beta_i \beta_i'$. Since Ω is a $T \times T$ covariance matrix, and Φ is a 2×2 covariance matrix, Ω has at most $(T+1)T/2$ distinct parameters, the same number as in the empirical covariance matrix, and Φ has at most three distinct parameters. Thus, if enough restrictions are placed on the parameters of Ω, it may be possible to identify the parameters of Φ, and our estimate of Φ would then be a reasonable estimate of the true covariance matrix of the earnings profile parameters. Our primary interest is in the term in the lower right-hand corner of the Φ matrix, since that parameter is identified as the true variance σ_{ww}.

In particular, if the hypothesis is maintained that the disturbance u_t can be decomposed into a random walk with equal-sized increments ϵ_t and a constant-variance transitory disturbance z_t, then Ω depends on two parameters, $\sigma_{\epsilon\epsilon}$ and σ_{zz}. We then have

(7) $$\underset{T \times T}{\Omega} = \sigma_{\epsilon\epsilon} R_T + \sigma_{zz} I_T$$

[7]In the empirical work in the following section, the dependent variable is the deviation of an individual's earnings in period t from mean earnings of sample members at period t. In that case, the means of the individual parameters, w_i and m_i, equal zero. The GLS estimate of the parameters for the ith individual is

$$\hat{\beta}_i = (Z'\Omega^{-1}Z)^{-1} Z'\Omega^{-1}(x_i - \bar{x}).$$

In my approximate GLS estimate, true Ω is replaced by $\hat{\Omega}^*$ in this formula. The empirical covariance matrix of the β's is $(n-1)^{-1} \sum_{i=1}^{n} \hat{\beta}_i \hat{\beta}_i'$. Since

$$\hat{\Omega}^* = (n-1)^{-1} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})'$$

by definition, upon substitution of the GLS estimate for β_i, $(n-1)^{-1} \Sigma \hat{\beta}_i \hat{\beta}_i' = (Z'\hat{\Omega}^{*-1}Z)^{-1}$ is obtained as asserted in the text.

where R_T is the "random walk matrix" whose ijth element r_{ij} is the smaller of the numbers (i, j) and I_T is the $T \times T$ identity matrix, with a diagonal of 1's. Thus the covariance matrix of earnings depends on five distinct parameters (the two from Ω and the three from Φ), and it is possible to identify σ_{ww} if $T \geq 3$. In section 4, I consider several other simple restrictions on the parameters of Ω that make it possible to estimate σ_{ww}.

4. Results from Statistical Calculations for OJT Effects

The empirical results discussed here are based on cohorts of men. Within each cohort, there is little variance in age (and presumably, in years of post-school employment experience), and all the members have the same level of schooling attainment. Most of the calculations reported here utilize subsamples of the Swedish Low-Income Commission Study (Låginkomstutredningen; LIU), for individuals born between 1936 and 1941. The schooling attainment subsamples include graduates of the Swedish elementary school (*folkskola*) who terminated their formal education approximately at age 14, folkskola graduates who then had some additional vocational training, graduates of secondary school (*realskola*), and realskola graduates with additional vocational training. Samples for *gymnasium* graduates or those with academic or professional degrees were too small for study. Taxable income appears to be an adequate surrogate for earnings for the segment of the earnings profile I analyzed. The taxable incomes are divided by the Swedish consumer price index, and the time index on which the earnings regressions are calculated is based on *age* rather than calendar year in which income was received. This procedure neglects the possible effect of annual increases in labor productivity for the Swedish economy. It is assumed this factor can be disregarded for seven-year cohort data. It is also assumed that within schooling levels, age is a good proxy for post-school employment experience. Taxable income data for 1951–1966 (except 1959) were obtained from official Swedish records. The regression calculations were begun at age 22. Earlier than that, there is enormous noise in the taxable income data because individuals with schooling below university level usually fulfill their compulsory military service when they are between 19 and 20 years old.

4.1 *Differences in Earnings Profiles within Schooling Attainment Classes*

We consider first a series of alternative estimates of σ_{ww}, which is the variance of the slope of the earnings (taxable income) profiles about the empirical mean profile for the cohort of elementary school graduates in the regression equation:

$$(1)'' \qquad x_{it} - \bar{x}_t = m_i + w_i t + u_{it}$$

where $\bar{x}_t$ is the empirical mean at time t.

For set A of Table 1, a simple OLS (ordinary least squares) regression was run for each individual, and the means and standard deviations of w were calculated for all individuals from the estimated regression parameters. Only observations for which taxable income in the period exceeded 5,000 Swedish kronor (Skr) were included. The major defect of OLS is that it takes no account of the covariance structure of u_t, and that leads to unnecessarily large errors in the estimates of m and w in the individual regressions. As a result, the calculated value of σ_{ww} tends to be exaggerated.

Set B consists of three approximate GLS regressions based on Rao's theorem, using the empirical covariance of earnings

$$\hat{\Omega}^* = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})'$$

In section 3 it was shown that $E(\hat{\Omega}^*)$ satisfies Rao's theorem. It was argued there that these estimates provide the basis for an upper-bound estimate for the standard deviation of w across individuals, since it is not possible to distinguish the sampling variability of the estimated w's from the variation of the true individual w's. However, these estimates should be superior to the OLS estimates, since we expect the sampling variance of the estimated w's to be less with GLS. In the three regressions in this set it is required that individuals have respective taxable incomes $\geq$ 5,000 Skr, $\geq$ 3,000 Skr, and $>$ 0 Skr, in order to test whether a truncation point for taxable income seriously modifies $\sqrt{\sigma_{ww}}$.

In set C, four alternative restrictions are imposed on the covariance matrix of the disturbances, Ω, in an attempt to distinguish between the true variance of the individual profile parameters and the sampling variance. In the first three cases it is assumed that u_t is a simple moving average of independent disturbances ϵ_t: in case $C1$, $u_t = \lambda\epsilon_t$; in case $C2$, $u_t = \lambda\epsilon_t + \theta\epsilon_{t-1}$; and in case $C3$, $u_t = \lambda\epsilon_t + \theta\epsilon_{t-1} + \psi\epsilon_{t-2}$. In each case, ϵ is assumed to be normally and independently distributed, with zero mean and unit variance. In the last case, $C4$, $u_t = y_t + \lambda z_t$, where y_t is a random walk: $y_t = \sum_{t=T_0}^{T-1} \gamma\epsilon_t$; both ϵ_t and z_t are normally distributed, independent, and have zero mean and unit variance. The explicit form of Ω for $C4$ was given in equation 7, with $\lambda^2 = \sigma_{zz}$ and $\gamma^2 = \sigma_{\epsilon\epsilon}$. The form of the covariance matrix of earnings is assumed to be $Z\Phi Z' + \Omega$, where the $2 \times T$ matrix Z has already been given, following equation 1'. The lower right-hand element of the Φ matrix, Φ_{22}, is taken as the estimate of σ_{ww}, conditional on the a priori restrictions on Ω.[8]

[8]The parameters of Φ and Ω are estimated in this study by weighted OLS by regressing the 28 elements of the triangular portion of the matrix (excluding those above the principal diagonal) on the corresponding elements of the parameterized alternatives of $Z\Phi Z' + \Omega$. The weights are the number of observations available for estimating such covariance term of $\hat{\Omega}^*$. This procedure for estimating the set C parameters is extremely crude, and future research should attempt to develop maximum likelihood estimates of the restricted model.

TABLE 1

ALTERNATIVE ESTIMATES OF STANDARD DEVIATION OF EARNINGS PROFILE SLOPES AND RELATED PARAMETERS FOR ELEMENTARY SCHOOL GRADUATES

Description of Sample	Number of Cases[a]	Standard Deviation				Correlations			Disturbance Parameters			
		w	m	m^*	m^{**}	r_{wm}	r_{wm^*}	$r_{wm^{**}}$	λ	θ	ψ	γ
Set *A*: OLS income ≥ 5,000 Skr	111	.60	2.11	1.80	2.40	−.60	−.05	+.74				
Set *B*: GLS												
Income > 5,000 Skr	111	.57	2.11	1.73	2.39	−.58	−.04	+.69				
Income ≥ 3,000 Skr	128	.59	2.30	1.98	2.62	−.53	−.03	+.65				
Income > 0 Skr	135	.60	2.53	2.25	2.82	−.46	+.015	+.63				
Set *C*: parameterized estimates: income ≥ 5,000 Skr												
*C*1 White noise (pure transitory)	111	.53	1.47	1.43	2.52	−.37	+.50	+.84	1.45			
*C*2 First-order moving average	111	.51	1.38	1.42	2.48	−.33	+.62	+.86	1.62	.325		
*C*3 Second-order moving average	111	.50	1.31	1.38	2.45	−.32	+.71	+.82	1.74	.437	.191	
*C*4 White noise and random walk	111	.48	1.35	1.14	2.00	−.56	+.22	+.84	1.11			.433

Note: w is the slope parameter of earnings profile deviations from the mean profile of earnings. For explanation of other terms and symbols, see text.

Source: The sample is from the Swedish Low-Income Commission Study (Laginkomstutredningen, LIU) and is composed of men who were born during 1936–1941 and who had terminated their formal schooling upon completion of the elementary (*folkskola*) level. The time series data on taxable income used here were obtained by Göran Ahrne from official statistics. In the time vector $t' = (0, 1, 2, \ldots, 6)$ zero corresponds to age 22 and six to age 28. The corresponding mean profile for taxable income is $x' = (17.42, 8.08, 8.65, 9.24, 10.14, 10.27, 10.81) \times 1000$ Skr (Swedish crowns). The sample is further restricted to those having valid taxable income statistics for four years or more.

[a]The number of cases (N) is 111, except in set *B*, where $N = 128$ for income ≥ 3,000 Skr and 135 for income > 0 Skr.

Table 1 contains the results obtained from the different estimating procedures. The most important estimates for this study are for the standard deviation of the earnings profile slope, w, over individuals. The difference between the OLS and corresponding GLS estimates (first two entries in Table 1) is small, and it is unlikely that it is statistically significant at conventional significance levels for this cohort. However, the GLS estimate is the smaller of the two, as expected, since GLS is more efficient than OLS. The level at which the income series is truncated (set B estimates) makes little difference in the result.

The set C estimates for $\sqrt{\sigma_{ww}}$ are all smaller than those for sets A and B, again as expected, since the method by which the set C estimates are obtained is intended to result in a direct estimate of the true variance of the w's by explicitly estimating the residual process parameters.

If, in set B, we take one standard deviation of difference in the profile slope for five years, for only those with taxable income $\geq$ 5,000 Skr, we obtain a change in relative incomes of 2,750 Skr — empirically a substantial difference. However, this estimate is a reasonable *upper bound*, because it incorporates the sampling variation in the $\hat{w}$'s. The random walk estimate, $C4$, provides a corresponding five-year effect of 2,400 Skr, which is still a fairly large empirical difference. Since the set C estimates attempt to allow for sampling variation, the reader may find them more persuasive.

Three sets of estimates of the standard deviation of the regression constant are provided, corresponding to three different crossover ages, since, as was pointed out in section 3, the value of the constant depends on the origin of the time vector associated with taxable income. $\sqrt{\sigma_{mm}}$ corresponds to $t = t_c = 0$ at age 22, the first term in the income profile data used in these calculations. There is no basis for assuming that crossover t_c occurs at this young age. Indeed, the estimated covariance of w and m, $\hat{\sigma}_{wm}$, is substantial and negative, contrary to other theoretical considerations and direct empirical evidence that the correlation between ability and the slope of the earnings profile is positive (see, e.g., Hause 1972).

In Table 1, m^* corresponds to crossover at age 24, and m^{**} to age 27. Even for $t_c = 0$ at 24, the covariance term is, with one exception, negative in sets A and B, although it is very small (and probably not significantly different from zero). This result may be inferred from the magnitude of r_{wm^*} for sets A and B in the table. However, for crossover at age 27, all the $r_{wm^{**}}$ coefficients are positive, and exceed 0.6.

Set C estimates of the standard deviations of m, m^*, and m^{**} are substantially different from the corresponding A and B estimates. This difference may reflect identification problems in an adequate parametric specification of Ω.[9] In that case, it is very doubtful that the set C estimates

[9] One complication arises in the case in which Ω is assumed to be the combination of a random walk and a transitory disturbance. Since the data do not pick up the taxable

of $\sqrt{\sigma_{mm}}$ are relevant. It is interesting to note that the set C estimates of r_{wm}, r_{wm^*}, and $r_{wm^{**}}$ change from negative values to large positive ones, as t_c is assumed to occur at ages 22, 24, and 27, respectively.[10]

Since the method used to estimate the set C parameters is very crude (see footnote 6), it is desirable to determine whether the estimated parameters of the restricted residual covariance matrix Ω are reasonable. The square roots of the estimated variance components of Ω may be interpreted as standard deviations of the random variables generating the income disturbances. The contemporaneous transitory deviation varies from 1,320 Skr (second-order moving average model, C3 to 1,060 Skr (model with random walk and white noise, $C4$.[11] These two standard deviations are 13.0 percent and 10.4 percent, respectively, of the mean income of elementary school graduates at age 26. Transitory income disturbances of this magnitude seem possible for young male workers in this age interval. $\sqrt{\sigma_{uu}}$, the standard deviation of the increment of the random walk in $C4$, is 660 Skr, which is 6.5 percent of mean income at age 26. Reasonably enough, the transitory disturbance has a substantially larger standard deviation. That estimate of $\sqrt{\sigma_{uu}}$ can be compared with one obtained by Fase (1970), who reports a standard deviation of about 3.5 percent of mean salary for elementary school graduates in his statistical study of salaried white collar workers. Fase's estimates are based on data for two adjacent years for white collar workers of all ages, and he assumes only a random walk disturbance. Given his sample, it is not surprising that he obtained a smaller estimate for $\sqrt{\sigma_{uu}}$ for Dutch workers than mine. Still, the Dutch and Swedish estimates are sufficiently similar and the size plausible enough to suggest that the crude technique used for the set C estimates yields reasonable results.[12]

---

income profile at the beginning of labor force participation, it seems plausible that the random walk for the initial data point will have already progressed to a level where the variance from the first (and subsequent) years should augment Ω by $\mu 11'$. (In this formula 1 is to be interpreted as a T-dimensional column vector of 1's.) But μ cannot be identified, since $\mu 11'$ cannot be distinguished from $\Phi_{11} 11'$. Hence the estimated Φ_{11} cannot be assumed to be a good estimate across individuals of the variance of m, the constant parameter in the regressions of (1'), in specification $C4$.

[10] It should be noted in passing that the estimated r_{wm} from $C4$ is +0.98 if the true crossover is assumed to occur at age 32. This value seems unreasonably high for the correlation of the ability level with slope. If crossover is assumed to occur at ages 25 or 26, more modest positive correlations are obtained that seem more plausible for indicating an estimate of crossover age.

[11] The corresponding deviation for the white noise model ($C1$) is 1,200 Skr, and for the first-order moving average model ($C2$), it is 1,270 Skr. The first-order lagged deviation for model $C2$ is 570 Skr; for $C3$, it is 660 Skr. The second-order lagged deviation for $C3$ is 440 Skr.

[12] An unsuccessful attempt was made to compute similar estimates for two additional cohorts of individuals with higher levels of schooling attainment. The samples were very small, and the GLS procedure yielded *larger* estimates of the profile slope variance than the OLS estimates. This anomalous result is due to incomplete observations vectors for computing the empirical covariance matrix of earnings. Appendix B shows that the differ-

Finally, we consider a sprinkling of estimates of partial correlation coefficients of income or earnings, $r_{ik.j}$, where the time indexes satisfy the condition $i < j < k$. A major purpose of section 3.1 in the discussion of statistical testing of the OJT hypothesis was to develop a procedure that could be used with fragmentary earnings profile data (consisting of earnings for at least three different points along the profile). The purpose of the test was to determine whether any systematic OJT effects that might be present in the data were strong enough to overcome the effect of the stochastic residual earnings variable, u_t, thereby yielding a significant negative sign for the partial correlation. In the discussion there, it was concluded that under several specifications, the best opportunity for observing a net OJT effect (i.e., a negative $r_{ik.j}$) occurs if i and k are on opposite sides of t_c and $j \sim t_c$. Several sets of partial correlations computed from a few alternative samples are shown in Table 2.

The partial correlations for the first two samples are in most cases based on very small numbers of observations, and are erratic. Most of the coefficients are positive, and they do not, as a whole, provide support for the supposition that gross OJT effects are strong enough to outweigh the random residual variability in earnings. Sample 3 provides stronger support for OJT effects within schooling attainment classes. Recalled full-time earnings data may have some error in measurement, but recalling may effectively reduce the transitory earnings variations that tend to mask OJT effects. The parametric approach of estimating standard deviations of earnings profile slopes is obviously more appealing for determining the potential empirical significance of OJT, but further application of the partial correlation test for the potential existence of net OJT effects may be worthwhile for fragmentary cohort data on earnings.

The negative partial correlations do not prove the existence of OJT effects (including the possible effect of systematic differences in occupational profiles, of course), but they provide evidence that the earnings profile slopes differ systematically within schooling attainment classes. To the extent that a sample member could anticipate these systematic differences at the time he entered the labor force, it would be expected that he would have taken them into account in considering the relative attractiveness of different careers.

ence between the OLS and the GLS dispersion matrices of the two earnings profile parameters is positive semidefinite if complete observation vectors are present, and this result holds for both large and small samples. An attempt to obtain estimates for other schooling levels analogous to the set C estimates was also unsuccessful, probably again because of missing observations.

TABLE 2

PARTIAL CORRELATIONS OF EARNINGS (OR INCOME) AT DIFFERENT AGES AND LEVELS OF SCHOOLING ATTAINMENT
(FIGURES IN BRACKETS ARE NUMBER OF OBSERVATIONS IN SAMPLE; THE SUBSCRIPTS OF r CORRESPOND TO $ij \cdot k$)

Educational Attainment					
Sample 1					
Folkskola	$r_{17,28.22}$ −.004 [22]	$r_{18,28.22}$ .031 [23]	$r_{21,28.24}$ −.036 [47]	$r_{22,28.24}$ .016 [22]	$r_{24,28.26}$ .059 [56]
Folkskola and vocational school	$r_{21,28.26}$ .004 [11]	$r_{24,28.26}$ .173 [14]	$r_{24,28.27}$ −.504 [14]		
Realskola	$r_{21,28.24}$ .176 [7]	$r_{24,28.26}$ .555 [10]			
Realskola and vocational school	$r_{21,28.25}$ .007 [13]	$r_{24,28.26}$ .397 [15]			
Sample 2					
Folkskola plus folkskola and vocational school	$r_{27,31.29}$ .385 [156]	$r_{27,34.30}$ .312 [114]	$r_{26,39.31}$ −.108 [92]	$r_{30,39.35}$ .190 [90]	
Realskola plus realskola and vocational school	$r_{28,34.31}$ .220 [28]	$r_{26,39.31}$ .118 [16]	$r_{26,39.35}$ −.599 [16]	$r_{30,39.35}$ .008 [15]	$r_{34,39.37}$ .752 [16]
Sample 3	$r_{65,55.60}$	$r_{60,50.55}$			
Nonhigh school graduate	−.082 [60]	.266			
High school graduate	−.346 [117]	−.023			
Some college	−.046 [51]	−.014			
College graduate	.049 [70]	.020			
College graduate, two or more degrees	−.374 [47]	−.185			

Source: Sample 1 is from the subset of LIU data used in Table 1. Sample 2 is from another subset of the LIU sample of men born 1924–1928, with income data collected by Erikson (1970). Sample 3 is from data collected by Daniel C. Rogers, mostly on Connecticut eighth graders in 1935. The partial correlations are from recalled full-time earnings in 1966 for 1950, 1955, 1960, and 1965. The subscripts on the partial correlations in sample 3 are for calendar year, not age. Assuming eighth graders are about 14 years old the corresponding partial correlation subscripts for age would be roughly $r_{44,34.39}$ and $r_{39,24.29}$.

4.2 *Compensatory Effects of OJT on Lifetime Relative Earnings*

As I discussed in the Introduction, the second major implication of the OJT model concerns the possible compensatory role played by OJT in reducing lifetime relative earnings variability compared with earnings variability in a single year. Crude estimates of that effect can be obtained using the Rogers sample. In the model developed in section 2, there are two factors that tend to make the standard deviation of the logarithm of discounted lifetime earnings smaller than the corresponding deviation for a single year. One is competitive behavior, which tends to equalize the economic returns from career choices with earnings profiles of different shapes, as exemplified by the OJT model. The other is random variation in the profile due to transitory fluctuations. The latter becomes relatively less important since it tends to be averaged out in the computation of discounted lifetime earnings, which is a weighted average of single-year earnings. The relative importance of the two cannot be determined unless estimates of the variance of the transitory component are available. However, the transitory random term is expected to be small in the Rogers sample because it contains data on full-time, not actual, earnings. The main defect of the sample is that the earnings data approximately span the ages from 29 through 44 years at five-year intervals. Although most members in the sample had their civilian careers delayed because of military service during World War II, a much better test of the potential OJT effects could be made if earnings data were available beginning at age 25 for the college graduates, and even younger for those with less schooling. Specifically, I expect this data limitation reduces our ability to infer compensatory OJT effects.

Table 3 contains some estimates of the standard deviation of the natural logarithm of earnings at ages 29 and 44, and the standard devia-

TABLE 3

RELATIVE VARIABILITY OF ANNUAL AND DISCOUNTED EARNINGS

Schooling Attainment	No. of Observations (1)	Standard Deviation of Natural Logarithm of Earnings (Age ~ 29) (2)	(Age ~ 44) (3)	Discount 4% (4)	Discount 8% (5)	% Diff.: (Col. 4 Less Col. 3)/ Col. 3 (6)
Non-high-school graduate	55	.257	.254	.211	.214	17
High school graduate	115	.335	.357	.281	.279	21
Some college	50	.437	.516	.400	.396	22
College graduate	63	.351	.452	.335	.323	26
College graduate (2 or more degrees)	46	.627	.518	.477	.473	8

Source: See Source note to Table 2 for sample 3.

tion of the natural logarithm of the discounted sum of the available data on earnings at a 4 percent and an 8 percent annual rate, i.e.,[13]

$$\log \sum_{t=0}^{X} \frac{Y_{30+t}}{(1+r)^{t}}$$

The procedure is crude, but the results are very suggestive. In every case, the log of the discounted sum has a lower standard deviation than 1950 or 1965 earnings (for example, cf. column 6 of the table).

It is well known that significance tests for differences in variances based on normal variates are very sensitive to departures from normality. I therefore adopt a more robust jackknife procedure, proposed by Miller (1968). He suggests basing the test on estimates of the logarithm of the variance rather than the variance itself. Table 4 contains the jackknife estimates of this statistic for 1965 log earnings and for the log of earnings discounted at 4 percent for each of the schooling classes. For comparison, I also present the corresponding full sample estimates of the log of the variances, which are slightly smaller than the jackknife estimates. As one would expect from Table 4, the discounted earnings statistics are all algebraically less than the 1965 statistics.

The samples are of modest size. In conventional one at a time, one tailed t tests with one degree of freedom, the null hypothesis is rejected at the 5 percent level only for both high school and college graduates. Fisher (1958, pp. 99–101) has proposed a pooled test. The natural logarithm of the significance level for each schooling level is taken. The sum of these over schooling levels multiplied by (−2) yields a χ^2 value of 27.4 with 10 degrees of freedom (twice the number of schooling levels), and it is highly significant at the 5 percent level. I conclude that the variance of the logarithm of discounted earnings is significantly smaller than the corresponding variance for the log of earnings in 1965.

[13] The Rogers data were purged of extreme earnings observations to make formal testing of the annual and discounted earnings feasible. Miller (1974, p. 13) provides an explicit warning that the jackknife is not a technique for correcting outliers. Rogers himself had estimated discounted lifetime earnings by imputing earnings for younger ages and 29 years and making an involved extrapolation for ages beyond 44. The imputation for the younger years within schooling attainment class consisted of an approximately constant term added to the sum of discounted earnings for which individual data were available. These estimates are inappropriate for our use, since the procedure by which they are made biases the results by making the variance of the log of discounted earnings smaller than the variance of the log of that portion of the earnings profile for which data are available. The reason is that if x_i is the known portion of the discounted earnings profile, then the variance and standard deviation of $\log x_i$ are larger than the variance and standard deviation of $\log(x_i + c)$, where c is the approximately constant positive imputed term. Heuristically, this occurs because the log is an increasing but convex function, and adding c to each x_i "squeezes" the transformed values, $\log(x_i + c)$, closer together.

TABLE 4

JACKKNIFE ESTIMATES AND SIGNIFICANCE TESTS FOR DIFFERENCES OF LOG VARIANCES OF SINGLE-YEAR AND DISCOUNTED EARNINGS (FIGURES IN PARENTHESES ARE ESTIMATED STANDARD DEVIATIONS)

Schooling Attainment	Jackknife Estimates of Log σ^2 1965	Jackknife Estimates of Log σ^2 4%	t Test	Deg. of Freedom	Full-Sample Estimate of Log σ^2 1965	Full-Sample Estimate of Log σ^2 4%
Non-high-school graduate	−2.72 (0.18)	−3.10 (0.18)	1.49	108	−2.74	−3.12
High school graduate	−2.05 (0.18)	−2.52 (0.18)	1.87	228	−2.06	−2.54
Some college	−1.28 (0.30)	−1.77 (0.36)	1.04	98	−1.33	−1.83
College graduate	−1.57 (0.17)	−2.17 (0.18)	2.35	124	−1.59	−2.19
College graduate (2 or more degrees)	−1.30 (0.16)	−1.46 (0.23)	0.57	90	−1.31	−1.48

Source: See Source note to Table 2 for sample 3.

5. POSSIBLE EXTENSIONS

The empirical results reported in the preceding section provide evidence consistent with two important implications of the OJT hypothesis. First, reasonable upper-bound estimates of OJT effects were obtained that suggest substantial systematic dispersion in earnings profile slopes for workers in their twenties, even after controlling for educational attainment. Second, one sample provides evidence that discounted lifetime earnings vary less than single-year earnings, a finding compatible with the hypothesis that OJT has a compensatory effect on lifetime earnings. Both features reflect significant structure in earnings profiles that cannot be directly observed in cross-sectional data, and illustrate the importance that longitudinal data have for uncovering the covariance structure of earnings. However, much remains to be done in verifying and extending these results, provided appropriate longitudinal data sets that cover the life cycle of earnings more completely and that provide more information on personal characteristics that affect labor market productivity.

In the model developed in section 2, the parameter m measures earnings at the crossover point in the absence of random variation, and it was assumed that m is correlated with economic ability (earnings potential).[14]

[14]It would also be useful to test directly for positive correlation of the estimated m's with a specific measure of economic ability. A significantly positive result would help explain my earlier finding (Hause 1972) that the effect of ability on earnings becomes stronger with increasing job experience.

It is not easy to test this assumption directly even if an adequate independent measure of ability is available. The age at which crossover occurs, t_c, is not observable, and in the present study I inferred this age by requiring a modest, positive covariance of m and w. I was led to this by the assumption, based on an earlier study of mine (Hause 1972), that there is a modest positive correlation between measured ability and the slope of the earnings profile. Since the value computed for m depends on the age at which crossover is assumed to take place, a test for the correlation of the calculated m and independently measured ability does not provide a very convincing test of my interpretation of m because the requirement that σ_{mw} be positive had already been imposed. The exact status of the crossover concept for future work is not very clear. To the extent that there is a central tendency for earnings profiles to cross each other over a relatively narrow interval of job experience, the concept appears to be a convenient construct. The results of the test for a negative partial correlation of earnings as evidence of OJT effects, developed in section 3.1 indicated that the crossover characteristic is useful for determining points along the earnings profile where a negative partial correlation is most likely to occur. It would be worthwhile to consider additional statistical procedures to characterize the extent to which there is a central interval where the profiles cross.

Another topic for future research is the relation of personal characteristics to different earnings profiles. In most empirical work, earnings or the logarithm of earnings for a single year are regressed on a set of explanatory variables including some function of years of experience unless the variance of the experience variable is very low in the sample. These regression functions provide some insight into personal characteristics associated with earnings, but they do not take into account that these characteristics define the potential *earnings profiles* available to individuals, not just the earnings for a single year. For example, the availability of financial resources to be invested in human capital or of personal abilities are both characteristics that affect the feasible investment strategies by affecting both costs and returns. A simple analogy with financial investments illustrates the point: Even if two companies have the same earnings per share, their price-earnings ratios can still differ greatly because investors' expectations about the prospects for future earnings are different for each. Similarly, two persons whose earnings are the same at a particular time may have very different earnings prospects. However, the serial correlation of personal earnings is presumably substantially higher than

The tendency for the ability coefficient to become stronger with work experience could reflect a steeper slope of the earnings profile for those with greater ability, as is suggested in the text. It is also possible that the profile of more able people continues to rise over a longer interval of time as they successfully demonstrate their superior capacity through job performances and become eligible for further advancement.

the serial correlation of company earnings; so one year of earnings data for an individual may have greater predictive power than one year of company earnings.

A final problem that warrants more study relates to the strong assumption that the covariance structure of *residual* earnings (after excluding OJT ability) is the same for all members of a given schooling cohort. It is presumably better to make this assumption than to carry out the estimation of the earnings profile slopes as if nothing were known about the covariance structure, which is the tacit assumption in OLS estimation. Still, I expect there is some difference in the residual covariance structure, depending on occupation, personal characteristics, or both. If sufficiently large samples are available, it would be desirable to evaluate the differences that may exist in the residual covariance matrix. This matrix is of great interest in its own right, since it is our estimate of the risk and uncertainty in the earnings stream faced by individuals once we have accounted as thoroughly as we can for systematic structural evolution of earnings the individuals may be able to predict.

University of Minnesota and
National Bureau of Economic Research

Submitted December 1973

Appendix A: Technical Notes on Testing for OJT Effects from Partial Correlations of Earnings

Negative partial correlations of earnings ($r_{ik \cdot j}$, where $i < j < k$) provide evidence of systematic slope differences in individual earnings profiles consistent with the OJT model. In this appendix I outline the formal arguments that determine the conditions under which $r_{ik \cdot j} < 0$ is most likely to be observed. I commented in section 3.1 that the sign of $r_{ik \cdot j}$ is determined by the sign of the determinant $D = (\sigma_{x_i x_k} \sigma_{x_j x_j} - \sigma_{x_j x_j} \sigma_{x_j x_k})$ where the σ's denote variances and covariances of earnings in periods i, j, and k.

Given equation 1 in the main text for individual earnings, the covariance of earnings for periods g and h (which may be the same), is

(A-1) $$\sigma_{x_g x_h} = \sigma_{mm} + gh\sigma_{ww} + (g + h)\sigma_{mw} + \sigma_{u_g v_h},$$

where σ_{mm}, σ_{mw}, and $\sigma_{u_j v_j}$ are the variances of m, w, and u_j; σ_{mw} and $\sigma_{u_g u_h}$ are the corresponding covariances; and the time origin is at the crossover point, $t_c = 0$. Substituting the terms from A-1 into the determinant D, we obtain

(A-2) $D =$

(a) $-(\sigma_{mm}\sigma_{ww} - \sigma_{mw}^2)(j - i)(k - j)$

$$\text{(b)} \quad +\sigma_{mm}(\sigma_{u_ju_j} + \sigma_{u_iu_k} - \sigma_{u_iu_j} - \sigma_{u_ju_k})$$

$$\text{(c)} \quad +\sigma_{ww}\{ik\sigma_{u_ju_j} + j(j\sigma_{u_iu_k} - k\sigma_{u_iu_j} - i\sigma_{u_ju_k})\}$$

$$\text{(d)} \quad +\sigma_{mw}\{2j\sigma_{u_iu_k} + (i + k)\sigma_{u_ju_j} \quad (j + k)\sigma_{u_iu_j} - (i + j)\sigma_{u_ju_k}\}$$

$$\text{(e)} \quad +(\sigma_{u_ju_j}\sigma_{u_iu_k} - \sigma_{u_iu_j}\sigma_{u_ju_k})$$

In the absence of random disturbances, u_t, only term a in equation A-2 is nonzero, and the main text shows that this term is necessarily negative unless there is perfect correlation between m and w (or unless σ_{ww} or σ_{mm} is zero). It is necessary to impose some restrictions on the autocovariance structure of the u_t's in order to deduce further conclusions from equation A-2. In this appendix I assume that u_t has the relatively simple structure $u_t = y_t + z_t$, where y_g and z_h are uncorrelated for all g and h, y_t is a random walk, and z_t is a white noise, i.e., has zero autocorrelation. With this specification, $\sigma_{u_gu_h} = \sigma_{y_gy_g}$ for $g < h$; $\sigma_{u_ju_j} = \sigma_{y_jy_j} + \sigma_{z_jz_j}$. For the pure random walk, $\sigma_{y_gy_g} < \sigma_{y_hy_h}$ for $g < h$. Under these assumptions we obtain from equation A-2:

(A-2′) $\quad D =$

$$\text{(a}'\text{)} \quad -(\sigma_{mm}\sigma_{ww} - \sigma^2_{mw})(j - i)(k - j)$$

$$\text{(b}'\text{)} \quad +[\sigma_{mm} + ik\sigma_{ww} + (i + k)\sigma_{mw}]\sigma_{z_jz_j}$$

$$\text{(c}'\text{)} \quad +\sigma_{ww}[i(k - j)\sigma_{y_jy_j} + j(j - k)\sigma_{y_iy_i}]$$

$$+\sigma_{mw}[(j - k)\sigma_{y_iy_i} + (k - j)\sigma_{y_jy_j}]$$

$$\text{(d}'\text{)} \quad +\sigma_{y_iy_i}\sigma_{z_jz_j}$$

It was established that term a' is negative. If only the white noise disturbance z_t is present, then it is necessary to determine the effect of component b' on the sign of D, and thus of $r_{ik\cdot j}$. The first term of the coefficient of $\sigma_{z_jz_j}$ is σ_{mm}, which is positive. If $i < j = 0 < k$, the second term of this coefficient is negative, and achieves its maximum value for a given time interval $k - i$ when i and k are equidistant from t_c, i.e., when $|i| = k$. For these values of i and k, $(i + k)\sigma_{mw} = 0$. Thus the opportunity of observing a negative $r_{ik\cdot j}$ is enhanced if we have $j = 0$ and $|i| = k$. These circumstances should maximize the probability of detecting the influence of OJT with a white noise.

On the other hand, if only the random walk disturbance y_t is present, then it is necessary to determine the effect of component c' on the sign of $r_{ik\cdot j}$. In c', the coefficient of σ_{mw} must be positive, since for a random walk $\sigma_{y_jy_j} > \sigma_{y_iy_i}$ for $j > i$. We have already indicated why σ_{mw} is expected to be positive. Thus, the second term of c' is expected to be positive. The first term in c' is obviously negative if $j = 0$ and $i < j < k$. A stronger assertion can be made about the first term of c' if the uncorrelated random shocks ϵ_t that generate the random walk have constant variance $\sigma_{\epsilon\epsilon}$. With this additional assumption the coefficient of σ_{ww} in c' is $t_0(j - i)(k - j)\sigma_{\epsilon\epsilon}$, where t_0 is the period of entry into the labor force. This coefficient is clearly negative for $t_0 < i < j < k$ (remembering that t_0 is negative), and the magnitude of the coefficient is maximized for a given time interval $k - i$ if $k - j = j - i$. Finally, the term $\sigma_{y_iy_i}\sigma_{z_jz_j}$, which is necessarily positive, appears if both random walk and white noise disturbances are present.

This analysis leads to the conclusion that the likelihood of observing $r_{ik\cdot j} < 0$ is greatest for a given time interval $k - i$ if $j = t_c(= 0)$, $|i| = k$. The obstacle to observing the negative partial correlation stems from the positive terms

$$\sigma_{mm}\sigma_{z_jz_j}, \quad \sigma_{y_iy_i}\sigma_{z_jz_j}, \quad \text{and} \quad \sigma_{mw}(k - j)(\sigma_{y_jy_j} - \sigma_{y_iy_i}).$$

The last of these terms may not be very important if σ_{mw} is small. It can be observed directly from equation A-2′ that if there is no systematic dispersion in the slopes of the earnings profiles, σ_{ww} and σ_{mw} are both zero, and $r_{ik\cdot j} > 0$, contrary to the implication of the OJT part of the model.

What if the autocorrelation structure of u_t assumes some other form? It can be shown that as long as the log of the autocorrelation function is convex, $r_{ik\cdot j} > 0$ if $\sigma_{ww} = 0$, i.e., if the profile slopes do not differ systematically.

Appendix B: An Application of a Theorem for Best Linear Unbiased Estimation of Regression Coefficients in the Correlated Case

By Urban Norlén

An application of a theorem by Rao (1967) for generalized least squares (GLS) estimation is shown to facilitate the procedure for estimating a set of regression coefficients. GLS estimation in its original form involves the disturbance covariance matrix, which is unknown here. With Rao's theorem as a point of departure we select another covariance matrix and use it in place of the disturbance covariance matrix. The replacement matrix is selected from a class of covariance matrices as given by the theorem and with the property that they all leave the GLS estimates unaffected.

1. Theory

We consider the generalized linear regression model

(1) $$\underset{T\times 1}{y} = \underset{T\times p}{Z}\,\underset{p\times 1}{\beta} + \underset{T\times 1}{\epsilon}$$

where y is the observation vector on the dependent variable. Z is a given observation matrix on the independent variables of rank $p(< T)$. β is the vector of parameters to be estimated. and ϵ a disturbance vector: y has the mean vector and positive definite covariance matrix

$$E(y) = Z\beta \tag{2}$$

$$E(y - Z\beta)(y - Z\beta)' = \Omega \tag{3}$$

respectively. We shall be concerned with the GLS estimator

$$b = (Z'\Omega^{-1}Z)^{-1}Z'\Omega^{-1}y \tag{4}$$

of β, which is the best linear unbiased estimator (BLUE) with covariance matrix

$$E(b - \beta)(b - \beta)' = (Z'\Omega^{-1}Z)^{-1} \tag{5}$$

One difficulty in employing this estimator is that the matrix Ω must be known, at least up to a factor of proportionality. This knowledge is often hard to get. In practice, therefore, this problem has often led to use of the usually less efficient least squares (LS) estimator

$$b^0 = (Z'Z)^{-1}Z'y \tag{6}$$

with covariance matrix

$$E(b^0 - \beta)(b^0 - \beta)' = (Z'Z)^{-1}Z'\Omega Z(Z'Z)^{-1} \tag{7}$$

which in general is larger than (5) in the sense that the difference between (7) and (5) is a positive definite matrix. In some of these problem situations, however, the following theorem by Rao (1967) may be used in obtaining GLS estimates of β although the disturbance matrix still is unknown. When Z and Ω are of full rank the theorem may be stated as follows.

THEOREM (Rao 1967): Let X be a $T \times (T - p)$ matrix of rank $T - p$ such that $X'Z = 0$, and let Ω^* be a matrix of the form

$$\Omega^* = \Omega + Z\phi Z' + \Omega X\psi X'\Omega, \tag{8}$$

where ϕ and ψ are arbitrary. Then the GLS estimator with Ω^* is the same as that for Ω.

If the matrix

$$A = I + Z'\Omega^{-1}Z\Omega \tag{9}$$

is nonsingular, a simple proof of sufficiency may be given. Consider the identity

$$\Omega^{-1} - \Omega^{*-1} = (\Omega^{-1}\Omega^* - I)\Omega^{*-1} = (\Omega^{-1}Z\Phi Z' + X\psi X'\Omega)\Omega^{*-1} \tag{10}$$

Premultiplication by Z' gives after simplification

(11) $$Z'\Omega^{-1} = (I + Z'\Omega^{-1}Z\Phi)Z'\Omega^{*-1} = AZ'\Omega^{*-1}$$

(12) $$Z'\Omega^{*-1} = A^{-1}Z'\Omega^{-1}$$

Postmultiplication of (12) by Z and inverting

(13) $$(Z'\Omega^{*-1}Z)^{-1} = (Z'\Omega^{-1}Z)^{-1}A,$$

which combined with (12) establishes the relation

(14) $$(Z'\Omega^{*-1}Z)^{-1}Z'\Omega^{*-1} = (Z'\Omega^{-1}Z)^{-1}Z'\Omega^{-1},$$

i.e., the GLS estimator with Ω^* and that for Ω are the same.

In passing we notice that an important class of covariance matrices is defined by (8) with $\Omega = \sigma^2 I$. For these covariance matrices for GLS and OLS estimators are the same.

2. *Application*

In the application of Rao's theorem to the present study, the observed income matrix

(15) $$S = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})(y_i - \bar{y})'$$

is used in place of Ω for making approximate GLS estimates. The expected value of this matrix

(16) $$E(S) = Z\left(\frac{1}{N-1}\sum_{i=1}^{N}\beta_i\beta_i'\right)Z' + \Omega$$

belongs to the equivalence class (8) that was shown to leave the GLS estimates unaffected.

The expected difference in the variability of the OLS and approximate GLS parameter estimates can be shown analytically. One measure of the difference is

(17) $$D = \frac{1}{N-1}\left(\sum_{i=1}^{N} b_i^0 b_i^{0\prime} - \sum_{i=1}^{N} b_i b_i'\right)$$

$$= (Z'Z)^{-1}Z'SZ(Z'Z)^{-1} - ((Z'S^{-1}Z)^{-1}.$$

where b_i^0 is the OLS estimate, given in equation (6), and b_i is the approximate GLS estimate, of the regression coefficients for individual i. This matrix is in general a positive definite matrix.

References

Eriksson, I. 1971. *Ålder och Inkomst* [Age and income]. Monograph för Låginkomstutredningen. Stockholm: Allmänna Förlaget.

Fase, M. M. G. 1970. *An Econometric Model of Age-Income Profiles*. Rotterdam: Rotterdam University Press.

Fisher, R. A. 1958. *Statistical Methods for Research Workers*. 15th ed., rev. Edinburgh and London: Oliver and Boyd.
Friedman, M., and S. Kuznets. 1954. *Income from Independent Professional Practice*. New York: National Bureau of Economic Research.
Hanna, F. 1948. "The Accounting Period and the Distribution of Income." In F. Hanna, J. Peckman, and S. Lerner, *Analysis of Wisconsin Income*. Conference on Research in National Income and Wealth, Vol. 9. New York: NBER.
Hause, J. C. 1972. "Earnings Profile: Ability and Schooling." *Journal of Political Economy*, Supplement, S108–S138.
———. 1974. "The Risk Element in Occupational and Educational Choices: Comment." *Journal of Political Economy* 82.
Mendershausen, H. 1946. *Changes in Income Distribution During the Great Depression*. Studies in Income and Wealth, Vol. 7. New York: NBER.
Miller, R. C. 1968. "Jackknifing Variances." *Annals of Mathematical Statistics* 39: 567–582.
———. 1974. "The Jackknife—A Review." *Biometrika* 61–1: 1–15.
Mincer, J. 1970. "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach." *Journal of Economic Literature* 8: 1–26.
———. 1974. *Schooling, Experience, and Earnings*. New York: NBER.
Norlen, U. 1975. "The Covariance Matrices for which Least Squares Is Best Linear Unbiased." *Scandinavian Journal of Statistics*.
Rao, C. R. 1967. "Least Squares Theory Using an Estimated Dispersion Matrix and Its Application to Measurement of Signals." In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 2, pp. 355–372. Berkeley: University of California Press.
Rogers, D. C. 1969. "Private Rates of Return to Education in the United States: A Case Study." *Yale Economic Essays* 9 (Spring): 88–134.
Thatcher, A. R. 1971. "Year-to-Year Variations in the Earnings of Individuals." *Journal of the Royal Statistical Association*, Series A, 134: 374–382.