NBER WORKING PAPER SERIES

# CAN PRODUCTIVE CAPACITY DIFFERENTIALS REALLY EXPLAIN THE EARNINGS DIFFERENTIALS ASSOCIATED WITH DEMOGRAPHIC CHARACTERISTICS? THE CASE OF EXPERIENCE

James L. Medoff

Katharine G. Abraham

Working Paper No. 363

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

June 1979

We are grateful to Gary Chamberlain and Zvi Griliches for their economic advice and to Martin Van Denburgh for his care and efficiency in doing the programming required for this study. The project has been supported by the NBER under its program for research on Labor Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper 363
June 1979

# Can Productive Capacity Differentials Really Explain the Earnings Differentials Associated with Demographic Characteristics? The Case of Experience

## ABSTRACT

This study uses computerized personnel microdata on the white male managerial and professional employees in a major U.S. corporation to address the following question: Can the additional earnings which are associated with more labor market experience at a point in time really be explained by higher productivity at the same point in time? Our answer to this question, based on both cross-sectional and longitudinal information, is that performance plays a substantially smaller role in explaining cross-sectional experience-earnings differentials or earnings growth than is claimed by those who have adopted the human capital explanation of the experience-earnings profile. Since our response to the question under analysis depends critically on our having assumed that the performance ratings which supervisors give (individually or collectively) to their white male managerial and professional subordinates adequately reflect the subordinates' relative productivity in the year of assessment, we present a great deal of evidence which very strongly supports this assumption.

Professor James L. Medoff
Katharine G. Abraham

National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, Massachusetts 02138

(617) 868-3914

The human capital explanation of the experience-earnings profile has dominated most economists' thinking about why workers' relative earnings increase with labor market experience. This theory has at its heart the assumption that the observed growth in relative earnings reflects growth in relative productivity, attributable to individuals' investment in on-the-job training. Since there are numerous other plausible explanations of why relative earnings grow with time on the job in which relative productivity growth plays virtually no role, the superiority of the human capital model over these alternatives must be established by demonstrating that when relative earnings are growing with labor market experience, relative performance is growing as well. Despite the widespread acceptance of the human capital model of the experience-earnings profile and the straightforward nature of the experiment required to empirically establish its superiority over alternative models in which factors other than productivity growth determine earnings growth, no one has ever provided evidence which demonstrates that experience-earnings differentials can in fact be explained by experience-productivity differentials. Thus, as is not atypical in economics, acceptance of the human capital theory of why relative earnings grow with experience is based solely on a priori logic.

The first experiments designed and conducted to ascertain the validity of the human capital interpretation of experience-earnings differentials are discussed in Medoff and Medoff and Abraham. These studies analyzed the earnings and rated performance of white male managerial and professional employees in three major U.S. corporations. They found that, among employees doing comparable tasks, while greater experience typically was

associated with substantially greater relative earnings it generally was associated with no greater or lesser relative rated performance. Hence the results of this experimentation lend no support to the human capital model of productivity-augmenting on-the-job training or to any other theory which attempts to explain experience-earnings differentials in terms of experience-productivity differentials.[1]

As would be hoped, the experimental design employed in our earlier studies has elicited a fair amount of criticism. The two most fundamental questions raised by our critics are: Do, as we assumed, the performance ratings given subordinates by their supervisors each year adequately reflect differences in the subordinates' current relative productivity? And, would our conclusions have been very different if our analysis were based on longitudinal data as opposed to the cross-sectional data employed?

The purpose of this study is to use a very rich body of personnel data from a fourth company (hereafter referred to as Company C) to address these two queries. In Section I, we discuss the nature of our experimental design in greater detail, summarize the results presented in the two previous pieces, and outline some potential criticisms of our methodology. We describe the data drawn from Company C's computerized data file and their transformation for use in this analysis in Section II. The Company C data are employed in Section III both to provide additional evidence which supports our use of performance ratings as an index of relative productivity and to demonstrate that our previous findings were not artifacts produced by cross-sectional data sets. The final section restates our major conclusion that, at least in the case of experience, a very substantial fraction of observed earnings differentials does not really reflect productivity differentials and discusses how we

might begin developing an empirically-based explanation of what experience-earnings differentials are really about.

## I. The Experiment

In this section we first briefly lay out the human capital explanation of the upward-sloping experience-earnings profile. We then discuss the design of the experiment used to test the explanation's validity and summarize the results of our earlier experimentation. Finally we consider criticisms of the experiment, some of which were anticipated and dealt with in Medoff and in Medoff and Abraham and some of which will be addressed for the first time in Section III.

### The Belief to Be Tested

The human capital model of the age-earnings profile posits that the upward sloping segments of an experience-earnings profile reflect on-the-job training, which causes the corresponding segments of an underlying experience-productivity profile to slope upwards.[2] There are, however, other potential explanations of the relationship between experience and earnings in which productivity growth plays a very minor, if any, role. For instance, Mincer (p. 80) recognizes the possibility that the positive association between experience and earnings might only "reflect the prevalence of institutional arrangements such as seniority provisions in employment practices." He then makes an observation that implicitly describes the nature of a seemingly good experiment for testing the human capital belief: "Such practices, however, do not contradict the productivity-augmenting hypothesis, unless it can be shown that growth of earnings under seniority provisions is largely independent of productivity growth."[3]

Thus the human capital interpretation of the experience-earnings profile is distinguished from other interpretations by the prediction:

earnings growth reflects productivity growth. Therefore it is possible to falsify the theory by demonstrating that, over much of an employee's work life, the corresponding pieces of experience-earnings and experience-productivity profiles do not have the same sign.

Experimental Design

The primary problem in designing an experiment which could provide evidence inconsistent with the human capital hypothesis is the measurement of an individual's productivity. In our two earlier studies, this problem was dealt with by assuming that the annual job performance ratings done by immediate supervisors are valid indicators of the relative productivity in the year covered by the appraisal of white male managerial and professional employees engaged in comparable work (i.e., that an employee rated ahead of X percent of his peers was, in the year of the rating, more productive than X percent). Under this assumption, it was possible to compare, for a given year, the relative earnings and relative productivity of employees in similar jobs. In particular, we could assess whether at a stated point in time the typical senior worker in a job received higher pay than a comparable junior worker in a similar job because the senior worker had higher productivity.

Our first experimental design employed standard semilog earnings functions to address the following three questions: 1. By how much does the pay of more experienced managers and professionals exceed that of otherwise comparable but less experienced managers and professionals? 2. Does a substantial part of the return to experience occur within grade levels, where the availability of performance ratings permits inference regarding the relative current productivity of more and less experienced managers and professionals? 3. Within grade levels, can the higher earnings of more experienced managers and professionals be explained by their current performance?

To answer the first question, we fit the following equation for each

company's white male managers and professionals:

1) $\ln(y) = \alpha + \underset{\sim\sim}{\beta X} + \varepsilon$,

with y equal to annual salary rate, the vector $\underset{\sim}{X}$ capturing educational attainment, pre-company experience and company service, $\alpha$ and $\underset{\sim}{\beta}$ representing parameters to be estimated and $\varepsilon$ representing the equation error. The estimated coefficients implied in each case that, at the appropriate mean of the relevant experience variable, an additional year of pre-company experience was associated with approximately one half of 1 percent higher earnings and an additional year of company service with approximately 1 percent higher earnings.

Next, regressions similar to equation 1 but with grade level dummies added on the right hand side were estimated for each of the three white male managerial and professional samples. The results from this second set of equations indicate that approximately 40 percent of the implied return at the mean to an additional year of either pre-company experience or company service occurred within grade level.

The finding that within-grade-level experience-related earnings differentials are a substantial fraction of total experience-related earnings differentials for white male managers and professionals at the three companies studied was very important because the available information on employee performance only permitted inference regarding the relative productivity of managers and professionals in the same grade. In each of the data sets employed, an overall performance rating awarded by a superior was recorded for all managers and professionals who had been on the job for at least four to six months. Superiors were expected to base their ratings of each subordinate on the extent to which during the year of appraisal that subordinate fulfilled the requirements of his or her

job. Since companies group their managerial and professional jobs into grade levels on the basis of the difficulty and importance of each job, it seemed reasonable to assume that within a grade level white male managers and professionals with high ratings were more productive in the year covered by the rating than were white male managers and professionals with lower ratings. However, no clear inference regarding relative productivities could be drawn from the performance ratings of persons in different grade levels.

One approach to ascertaining whether performance mediated the documented link between labor force experience and higher within-grade earnings among white male managers and professionals was to add performance rating dummies to the variables included on the right hand side of the second set of semilog earnings function regressions (the set with grade level dummies included). If more experienced people received higher within-grade salaries in a year even in part because they were more productive in that year, the introduction of performance rating dummies as controls for relative within-grade productivity should have moved the estimated within-grade return to experience towards zero. In no case did this occur.

Another experimental design for examining the same set of interrelationships entailed multinomial logit estimation of the effect of labor force experience (and other personal characteristics) on the probability of a white male manager or professional being at the top, in the middle or at the bottom of the salary distribution for his grade level and on the probability of a white male manager or professional being at the top, in the middle or at the bottom of the performance distribution for his grade level. The results of the multinomial logit estimation implied in all cases that while an additional year of either

pre-company experience or company service substantially and significantly reduced the probability of a sample member with mean characteristics being at the bottom of the relevant within-grade salary distribution, the same year of either pre-company experience or company service increased or did not affect the probability of a sample member with mean characteristics being at the bottom of the relevant within-grade performance distribution.

Thus, results from three companies support the conclusion that at least a substantial fraction of the higher intrafirm earnings associated at a given point in time with additional labor force experience among managers and professionals is unrelated to productivity at that same point in time. Comparable cross-sectional results presented in Section III below based on data from the fourth company we have called Company C reinforce this basic conclusion.

## Criticisms and Responses

One criticism of the basic cross-sectional result might start out by noting that although four to six performance ratings were available to supervisors at each of the three companies originally studied, in fact at each of these enterprises most people received one of only two ratings so that in effect only one cut in each within-grade performance distribution was made. Thus, this line of argument might continue, although more experienced employees were no more likely to be above than below the one cut in the relevant performance distribution, within each of the two parts of the performance distribution more experienced employees might be uniformly more productive than otherwise comparable less experienced employees. If such were the case, while introduction of crude performance rating dummies into the semilog earnings function which controlled for grade level would not produce any movement of the estimated within-grade return to

experience towards zero, introduction of finer performance measures would lead to a substantial reduction in the estimated return to experience. Better performance measures contained in the Company C data set permit the exploration of this possibility. As discussed further in Section II of this study, supervisors at Company C were required to follow a forced distribution in awarding one of the six (beginning in 1976) or nine (prior to 1976) possible ratings to each of their subordinates so that bunching of ratings was not a problem. In addition, each employee was ranked annually by a group of supervisors against a group of his colleagues, providing an even more precise measure of relative performance. Results reported in Section III below demonstrate that the use of these finer performance variables strengthens the basic conclusion that, within grade levels at a given point in time, higher earnings of more experienced white male managers and professionals are unrelated to productivity.

Another approach to criticizing our basic finding might be to assert that performance ratings are too flawed to offer any reliable information about the relative productivity of workers within a grade.[4] One commonly cited problem with performance ratings is that most supervisors tend to be overly lenient; hence, it is argued, ratings are inflated. This problem is irrelevant for present purposes provided that leniency in rating does not lead to a distortion of the relative position of employees. Further, the performance measures available in the data set explored in this study appear to be much less susceptible to charges of inflation than those available in the original three data sets; the Company C ratings follow a forced distribution and rankings are inherently immune to inflation.

Another potential problem with performance ratings is that different supervisors might employ divergent standards of evaluation. Even if divergences in individual supervisors' rating standards were typically

substantial (available evidence suggests they may typically be small) this source of variation in employee's rated performance could not explain our results unless employees with different observed characteristics were systematically paired with supervisors applying different sets of standards.[5,6] We know of no evidence which indicates that a pairing of this nature would be other than a very rare event.

Furthermore, variation in the standards applied in awarding the Company C ratings and rankings should be minimal. Since (as stated above) neither performance measure can be seriously inflated, variation in leniency on the part of supervisors should not be a significant source of variation in rating standards. The fact that both the ratings and the rankings reflect a consensus reached by a group of supervisors rather than just one person's opinion also argues against there being significant variation in rating standards.

It is also possible that employee's personal characteristics (race, sex, age, time with company) might influence supervisors' performance assessments. We see no reason to conclude a priori and have seen no evidence suggesting that supervisors discriminate against older or more experienced employees in giving out performance ratings. In fact, available evidence suggests that if there is any experience-related distortion in performance ratings, it works in the direction of overstating the more experienced person's true relative performance in the year of appraisal.[7]

Another possibility is that the rated performance of a group of more experienced employees would most likely be an understatement of their true relative productivity if their supervisor had greater expectations for them as opposed to their less experienced grade-level compatriots and systematically gave them more difficult assignments. It seems unlikely <u>a priori</u> that supervisors will expect more of and give more difficult tasks to individuals

in a given grade level with greater than average experience, since this group of "stayers" can be expected to be less able. Evidence supporting this claim was unearthed in a study (supported by the Harvard Business School) of 2,500 engineers and managers at six technology-based companies done by Dalton and Thompson. Their findings indicate a nontrivial (in absolute value) negative relationship between supervisors' assessment of the complexity of the tasks a subordinate engineer was typically asked to perform and the subordinate's age. Two reasons why older engineers might be asked to perform less complex tasks than younger engineers are, first, that the typical older engineer is most likely less able than the typical younger engineer since the more able members in any cohort of engineers tend to be promoted to supervisory positions as they become older and, second, that the skills possessed by the typical older engineer are likely to be somewhat obsolete. Similarly, one would expect that older managers in any grade level might be asked to perform less complex tasks than otherwise comparable younger managers in the same grade level because older managers in any grade level are probably less able than younger managers in the same grade level having the same amount of education and because the skills of older managers are probably somewhat obsolete.

Some readers of our earlier studies have asserted that we should seek a "hard" or "objective" measure of the present value of the marginal revenue product of each employee's current actions since it would clearly dominate our "soft" or "subjective" performance rating or ranking. For this statement to be correct, two thing would have to be true. First, the dimensions of an employee's current true value to his or her firm would have to be quantifiable. Second, either there would have to be only one relevant dimension in assessing the employee's true worth or the researcher would have to know the proper set of weights or shadow prices to attach to each relevant dimension.

Unfortunately, for the vast majority of occupations in an advanced industrial society these two prerequisites will not hold. To see the nature of the problem, consider the job of a secretary. No one quantifiable measure (e.g., words typed per hour, phone calls answered per hour, etc.) would adequately capture a secretary's true value to his or her employer. Moreover, it is unlikely that all of the relevant dimensions of a secretary's value marginal product are quantifiable. Furthermore, even if all of the relevant factors could be quantified, we see little reason to believe that any set of aggregation weights chosen by a researcher for creating a single productivity measure based on the various relevant factors would be better than those implicitly chosen by the secretary's supervisor in awarding a performance rating or ranking. Even for blue-collar jobs in which workers are paid by the piece, deriving a satisfactory "hard" or "objective" measure of current productivity is likely to be difficult. For instance, one tire assembler might assemble more tires than another tire assembler without having any greater true current productivity by turning out a relatively large number of flawed tires (which sneak by inspectors), being relatively remiss in maintaining the machinery he or she uses, being relatively wasteful of materials, or being less willing to assist his or her compatriots than the typical assembler acting in a way consistent with long run profit maximization.[8]

In the case of managerial and professional jobs, the problems of deriving a "hard" or "objective" measure of relative value to the firm are particularly complex. While the bottom line might at first blush appear to be the current output to input ratio in the unit under an employee's supervision, it is not at all obvious that this measure would dominate the performance ratings and rankings employed in our analysis. This is because the current output to input ratio can most likely be increased by utilizing capital and subordinate labor in ways clearly inconsistent with long-run profit

maximization. For this reason, most companies tell supervisors to consider factors other than short-run unit success in doing their assessments of subordinates' worth. For instance, at the airline analyzed in Medoff an "Overall Evaluation" was to be based on "Subordinate Coaching" and "Teamwork" in addition to "Unit Achievement," and currently at Company C the "Overall Appraisal" is to be based on 20 factors in addition to "Quantity of Work", those factors including "Quality of Work", "Resource Utilization", "Planning Effectiveness", "Developing People", "Security Sensitivity", and "Safety Consciousness". It seems most unlikely that a researcher could attach a better set of weights to the various factors determining a manager or professional's value to his company (e.g., a weight of one for "Unit Achievement" or "Quantity of Work" and a weight of zero for all other factors) than those implicitly assigned by the supervisors of jobs in rating or ranking the performance of their subordinates. In sum, it is not at all obvious that researchers can derive "hard" or "objective" indexes of relative true productive value which will in fact dominate the index based on performance ratings and rankings we used.

Substantiation of the validity of managerial and professional performance ratings can be sought by assuming that, holding other personal characteristics constant, firms should be more willing to promote and give larger raises to more productive managers and professionals, and subsequently interpreting evidence that managers and professionals with higher performance ratings are significantly more likely to be promoted and receive larger raises as supportive of the proposition that there must be a significant positive correlation between measured performance and true relative productivity. Results reported in the original Medoff and Medoff and Abraham studies demonstrate a strong positive partial correlation (after controlling for education, experience, grade level, and, in the salary increase equations, ln(annual salary)) both between higher performance ratings and promotion in subsequent years

and between higher performance ratings and percentage salary increases in subsequent years. Replication of these promotion and percentage salary increase experiments using the Company C data yields the same conclusions; the results are reported in Section III.

While it is reasonable that performance ratings would be one piece of information utilized by firms in making promotion decisions, there would be cause for concern if performance ratings did not reflect solely a manager or professional's productivity in his current job but also his supervisor's assessment of his potential for advancement within the company. Since, all else equal, more experienced (older) managers and professionals are typically less likely to be promoted out of any given grade level than less experienced (younger) managers and professionals and hence apt to be perceived as having limited potential for future advancement, to the extent that performance ratings are affected by supervisors' assessments of individuals' future potential, more experienced workers might typically receive lower performance ratings than warranted simply on the basis of their current productivity.[9]

The institutional basis for this argument is weak insofar as at all four companies whose data have been analyzed the instructions to supervisors preparing ratings clearly ask for an assessment of how well each employee is fulfilling the requirements of his or her current job, not for a forecast of the employee's productivity profile over his or her work life. At the airline studied in Medoff, the "Reviewer's Workbook" gave supervisors the following instruction concerning the time period covered by their appraisals: "Considering the specific performance evaluations on the preceding pages, state below your overall evaluation of the manager's performance <u>over the past 12 months</u>" [our italics]. At Company A (discussed in Medoff and Abraham) supervisors are asked to make entries on

a performance review worksheet "during the year for the purpose of providing supporting information for the annual employee assessment" [our italics]; the intent is clearly that each employee's annual performance rating should reflect his or her performance during the appraisal year. At Company B (also discussed in Medoff and Abraham) the "Supervisor's Guide for Performance Review and Development Planning" states that "[r]eviews should be based on the employee's performance in his present position and only for the period [typically one year] since his last review" [our italics]. At Company C, as will be discussed below, the instructions to raters tell them that "[e]ach employee should be rated on his current performance and contributions based on requirements of his present assignment" and that "career potential and promotability should not enter into ratings of an individual's performance" [our italics].

It should also be pointed out that the Dalton and Thompson study of age-earnings and age-performance differentials among engineers at six companies, which provided evidence consistent with our findings, derived performance rankings by asking supervisors to assess the "contribution made to the company during the past year" [our italics] by their subordinates. The Dalton and Thompson analysis was not based on ratings and rankings done under companies' actual appraisal programs which would provide inputs into salary growth or promotion decisions. Finally, it should be noted that at many companies, including two of the four whose data we have analyzed, each employee's promotability is formally assessed separately from his performance; at Company C, for example, each employee's supervisor annually records the highest grade level within the company which he or she believes the employee is likely to attain. The results of statistical analysis presented in Section III imply that the Company C performance ratings and the Company C potential measures really do capture different things; while performance ratings do very much better than potential ratings in explaining current salary, potential ratings

strongly dominate in predicting promotions.

Some critics have accepted our assumption that performance ratings adequately capture current relative productivity but have argued that our basic finding is an artifact of the cross-sectional data employed and would not be observed with longitudinal data. One possible story about why this might be true assumes that at each of the companies studied there had been a secular decrease in the share of the costs of firm specific on-the-job training borne by employees. It would then be possible (as is shown more formally in Section III) for longer service employees to earn more and perform less well than shorter service employees without contradicting the tenets of human capital theory. Adequate response to this line of argument requires the presentation of longitudinal evidence showing that at least some individuals have increasing relative earnings but decreasing relative productivity over time. Perhaps the most important result reported in this study is the finding that at Company C the relative within-grade earnings position of white male managers and professionals remaining in the same grade level over a period of years improves significantly and substantially with the passage of time while the relative within-grade performance position of those same managers and professionals deteriorates over the same time period. Section III discusses this finding in detail.

## II. The Company Personnel File

All of the results reported in the next section of this paper are based on information extracted from a computerized personnel file made available to us by a large U.S. manufacturing corporation which we have called Company C. The raw Company C data file consisted of five annual segments, one for each year from 1973 through 1977, each segment containing a record for every exempt employee who was "active" with the company at any time during the given year. Each annual employee record included information on the employee's education, length of company service, date of birth, physical work location, current job grade, date of entry into current job grade, current salary and recent salary increases; in addition, it included two separate assessments of the employee's performance and an assessment of the employee's potential for advancement within the company prepared during the summer of the previous year (so that, for example, the 1973 records contained performance assessments and potential assessments prepared during the summer of 1972). Under the assumptions that all changes in grade level during 1973 were promotions from one grade level to the next higher grade level, that no persons changed grade level more than once during 1973, that the salary increase information on the 1973 records captured all salary changes made during 1973, and that no employee changed educational status or physical work location during 1973, a complete fix on each employee's year-end status (including demographic information, grade level, salary and evaluations of performance and potential received during the year) was constructed for each year from 1972 through 1976; similar data except for the performance and potential evaluations were available for 1977.[10]

One further adjustment was needed to prepare the year-end data for

analysis. Because the performance and potential evaluations done each year were completed during the summer while each employee's grade level was recorded as of year end, the grade levels originally present on the annual employee records corresponded imperfectly to the grade levels the employees were in at the time of their annual performance and potential reviews. To correct for this shortcoming in the data, it was assumed that employees with a grade change date later than June 30 of the relevant year had been evaluated prior to changing grade level. For these employees, the grade level entered on the relevant annual record was reduced by one to the next lower grade level. Whenever subsequent reference is made to an employee's grade level, it is this adjusted grade level which is indicated.[11]

The populations for each year 1972 through 1976 on which all of the subsequently discussed data transformations were based consisted of all domestically based white males "active" with full time exempt status as of the relevant year end and with (adjusted) grade level in the Company C managerial and professional range for whom complete performance and year end salary information for the relevant year was available.[12] Those included in the samples used for the experiments reported in Section III had in addition to meet more restrictive information availability requirements as dictated by each particular experiment.

The schooling information on the Company C file was used to categorize population members in each of the years 1972 through 1976 by highest level of educational attainment: less than high school, high school diploma, bachelors degree, masters degree or doctorate. Pre-company experience variables were computed to equal age as of each year end minus schooling minus company service minus five. For this purpose, it was assumed that

non-high school graduates had spent 10 years in school, high school graduates 13 years, college graduates 16 years, masters degree holders 18 years and Ph.D.'s 21 years. Dummy variables were created which placed each Company C physical work location in one of four regions: Northeast, North Central, South or West.[13]

Performance Rating and Ranking

At Company C, supervisors formally assess the performance of each managerial or professional employee once a year. Two separate performance measures are recorded: an overall performance rating and a ranking of each employee relative to others in an appropriate comparison group.

Performance ratings are prepared initially by each employee's immediate supervisor. The rating form in use during the period under analysis gave supervisors the following instructions:

> Each employee should be rated on current performance and contributions based on requirements of his present assignment. An employee should be measured both as to his contributions in terms of the standards of his job and against others performing similar work at similar levels. Career potential and promotability should not enter into ratings of an individual's performance.

Prior to 1976, the overall rating scale comprised nine categories; beginning in 1976, the number of appraisal categories was cut to six by merging the top two categories and merging the bottom three categories from the old rating scale.

In addition to rating the performance of each subordinate, supervisors at Company C are required to rank each employee relative to a group of his peers. Management provides each supervisor in a department or other

appropriate organizational unit with a list of employees doing reasonably comparable levels of work.[14] The supervisor is first told to strike off the list employees "whose work you do not know well" or "whose work in your opinion is so different from most of the others that you do not think he (or she) can be compared with them" and then instructed to rank the remaining people. The ranking is done by first picking the best employee on the list, then picking the worst employee on the list, next designating the best employee of those remaining on the list, and then designating the worst employee of those remaining on the list, and so on until all of the employees have been ranked. The same criteria used in assigning the performance ratings underlie the rankings. That is, employees are ranked on the basis of how well each is meeting the requirements of his or her own particular assignment compared to how well others are meeting the requirements of their assignments. Rank group lists are designed to be as large as practicable, with as many supervisors as possible serving as rankers for each group consistent with their having adequate knowledge of the employees being ranked.

After all of the supervisors at a given level of the corporate hierarchy have assigned performance ratings to their subordinates and ranked those employees whose work they are qualified to assess, the rating and ranking forms are reviewed by managers at the next level of the corporate hierarchy. Any major disparities among rankers are reconciled and a consensus is reached regarding each employee's position in his or her ranking group.

Finally, the performance ratings made by the immediate supervisors may be modified so that they are consistent with the consensus rankings and so that the overall distribution of ratings within the ranking group is reasonably consistent with the company's expected performance

distribution. In 1976, for example, the enunciated goal was that no more than 10 percent of all employees rated should receive the highest possible rating, approximately 25 percent should receive each of the next three ratings and approximately 15 percent should receive one or the other of the bottom ratings. In all years, the actual distribution of ratings ultimately given closely matched the desired distribution. In 1976, to continue our example, the actual proportions of those included in our population versus the desired proportions of employees receiving each rating were (from top to bottom) as follows: 9 percent versus 10 percent; 25 percent versus 25 percent; 26 percent versus 25 percent; 25 percent versus 25 percent; and (for the bottom two rating categories combined) 15 percent versus 15 percent.

As noted above, both performance rating and ranking are designed to measure an employee's performance relative to "the total requirements and standards of the job." According to a manual describing their salary system which we received from Company C:

> Every management and professional position at [the Company] is assigned to a classification level which is determined by an evaluation of the job and a measurement of the value of that job relative to other jobs in the Company. Thus positions of similar value are placed in the same classification level.

Given the way jobs are assigned to grade levels at Company C, we felt comfortable assuming that within a given grade level those with high performance ratings or high rankings were more productive than those with low performance ratings or low rankings.

For some experiments, sets of dummy variables were constructed from the recorded performance ratings. For other experiments, the recorded performance ratings were transformed into variables reflecting percentile position in the relevant within-grade performance distribution.

One rating percentile variable (P) was calculated based on the 1976 population using the formula $P = F/2$ for persons with any of the bottom three performance ratings in 1976 and the formula $P = F + (1 - F)/2$ for persons with any of the top three performance ratings in 1976, where F represents the proportion of 1976 population members in the relevant grade level who were given any one of the bottom three performance ratings in 1976. Five more refined rating percentile variables (P') were calculated, one set for each of the years 1972 through 1976, using the formula $P' = F(i-1) + [F(i) - F(i-1)]/2$, where F(i-1) represents the fraction of the relevant year's population members in the appropriate grade level with a rating in the given year of i-1 or below; F(i) represents the fraction of the same group with a rating in the given year of i or below; and i represents the rating received in the given year by the individual for whom P' is being calculated, with i having possible values of one (worst) through nine (best) in 1972 through 1975 and possible values of one (worst) through six (best) in 1976.

The raw ranking information for each of the years 1972 through 1976 was transformed into continuous ranking percentile variables (R) using the formula $R=(N-Q+1/2)/N$, where Q represents the individual's rank within the relevant comparison group in the given year and N represents the number of people against whom the individual was ranked in the given year. Ranking percentile variables with a limited number of possible values (R') were created for each of the years 1972 through 1976 for use in those experiments where both a ranking variable and an analagous salary variable were required. Where R was between zero and .10, R' was set equal to .05; where R was between .10 and .20, R' was set equal to .15; and so on.

Salary Variables

A salary variable (SP) analagous to the 1976 two category (low or high) rating percentile variable (and also based on the 1976 population) was constructed by cutting each year-end within-grade salary distribution at the cumulative frequency corresponding to the proportion of 1976 population members in the relevant grade level with any one of the bottom three ratings, then setting SP=F/2 for those individuals with salaries at the end of 1976 below the cut and setting SP=F+(1-F)/2 for those individuals with salaries at the end of 1976 above the cut, where F again represents the proportion of 1976 population members in the relevant grade level with any one of the bottom three performance ratings in 1976.

Five salary variables (SP') analagous to each of the more refined rating percentile variables (and also based on the 1972 through 1976 populations) were constructed by cutting each year-end within-grade salary distribution at the cumulative frequences corresponding to the proportions of the relevant year's population members in the appropriate grade level with the given rating or a worse rating in the given year, then, depending on the individual's year-end salary and grade level in the relevant year, setting SP'=F(i-1)+[F(i)-F(i-1)]/2, where F(i) represents the cumulative proportion of the relevant year's population members in the appropriate grade level with year-end salaries at or below the cut closest above the individual's salary and F(i-1) represents the cumulative proportion of the same group with year-end salaries at or below the cut closest below the individual's salary.

For comparability with the R' ranking percentile variables, a set of salary variables (SR') (also based on the 1972 through 1976 populations) were calculated. To do this, each relevant year-end within-grade salary

distribution was divided into deciles; individuals with salaries at the end of the relevant year in the first decile were given a percentile score SR' equal to .05, individuals with salaries at the end of the relevant year in the second decile were given a percentile score SR' equal to .15, etcetera.

To summarize, for each individual for whom the requisite raw rating, raw ranking and raw year-end annual dollar salary figures were available, the following transformed variables were created: 1. Sets of performance rating dummies; 2. A two category (low or high) 1976 rating percentile variable and an analagous two category 1976 salary percentile variable; 3. A more refined rating percentile variable and an analagous salary percentile variable for each year 1972 through 1976; 4. A continuous ranking percentile variable for each year 1972 through 1976; and 5. A ranking decile variable and an analagous salary decile variable for each year 1972 through 1976.

Rated Potential

In addition to evaluating the current performance of each employee and awarding performance ratings and rankings, supervisors at Company C were required to perform an annual assessment of each employee's potential for advancement within the Company. Based on consideration of such factors as quality of performance, capacity to learn, judgement and motivation, supervisors were first asked to give one or more examples of the highest level of job which each employee had the potential ability to achieve. Next, supervisors were requested to consider whether such factors as an employee's preference for a particular type of work or an employee's health would hold the employee back, have no effect, or help the employee achieve full potential. Finally, having considered both of the above sets

of factors, supervisors were asked to record the grade of the highest level of job which each employee could realistically be expected to achieve during the course of his career.

## III. Basic Results and Responses to Criticisms

In analyzing the Company C personnel data, our first question was, controlling for level of educational attainment and region, how large are the earnings differentials between managers and professionals with different amounts of labor force experience? The first regression in Table 1 indicates that at the mean of the relevant variables, an additional year of pre-company experience is associated with a small but significant positive increment to earnings and an additional year of company service is associated with somewhat more than a 1 percent increment to earnings. The regression coefficients imply that, all else equal, a person with one standard deviation more than the mean number of years of pre-company experience would earn 3 percent more than a person with no pre-company experience and that a person with one standard deviation more than the mean number of years of company service would earn 29 percent more than a person with one standard deviation less.

Next we asked whether a substantial fraction of the higher earnings associated with additional labor force experience occurred within grade levels rather than as a result of more experienced managers and professionals holding jobs in grade levels with higher than average salaries. To answer this question, we added grade level dummies to the controls present in the first model to get model 2 and calculated the fraction of the total return to experience they captured. The results of this procedure imply that all of the return at the mean to an additional year of pre-company experience occurs within grade level and that 42 percent of the return at the mean to an additional year of company service occurs within grade; a person with one standard deviation more than the mean number of years of pre-company experience would earn 4 percent more than an otherwise comparable

Table 1: Influence of Education, Experience and Performance on Managerial and Professional Earnings[a]

Dependent Variable: ln (Annual Salary as of December 1976)[b]

(N = 8,238)

| | Mean [S.D.] | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|---|
| Less than High School Diploma (yes = 1) | .021 [.143] | -.281 (.017) | -.056 (.006) | -.066 (.006) | -.069 (.006) | -.070 (.006) |
| High School Diploma (yes = 1) | .343 [.475] | -.204 (.006) | -.015 (.002) | -.025 (.002) | -.030 (.002) | -.030 (.002) |
| Masters Degree (yes = 1) | .145 [.352] | .104 (.007) | .025 (.003) | .025 (.003) | .026 (.002) | .025 (.002) |
| Doctorate (yes = 1) | .043 [.203] | .238 (.012) | .052 (.004) | .055 (.004) | .057 (.004) | .056 (.004) |
| Years of Pre-Company Experience/10 | .521 [.532) | -.019 (.010) | .035 (.004) | .047 (.004) | .054 (.004) | .052 (.004) |
| (Years of Pre-Company Experience)$^2$/100 | .554 [1.247] | .030 (.004) | .003 (.002) | .002 (.001) | .002 (.001) | .002 (.001) |
| Years of Company Service/10 | 1.946 [1.125) | .323 (.008) | .110 (.003) | .115 (.003) | .119 (.003) | .117 (.003) |
| (Years of Company Service)$^2$/100 | 5.053 [4.504] | -.050 (.002) | -.014 (.001) | -.014 (.001) | -.013 (.001) | -.013 (.001) |
| Performance Rating 1 (Worst; yes = 1) | .011 [.102] | -- | -- | -- | -.116 (.008) | -- |
| Performance Rating 2 (yes = 1) | .137 [.344] | -- | -- | -- | -.083 (.003) | -- |
| Performance Rating 3 (yes = 1) | .253 [.435] | -- | -- | -- | -.045 (.002) | -- |
| Performance Rating 4 (yes = 1) | .258 [.437] | -- | -- | -- | -.027 (.002) | -- |
| Performance Rating 6 (Best; yes = 1) | .092 [.289] | -- | -- | -- | .032 (.003) | -- |
| Performance Rating 1-3 (yes = 1) | .400 [.490] | -- | -- | -.050 (.002) | -- | -- |
| (No. Ranked - Rank + 1/2)/No. Ranked | .504 [.289] | -- | -- | -- | -- | .109 (.003) |
| Constant | -- | yes | yes | yes | yes | yes |
| Region Dummies (3) | -- | yes | yes | yes | yes | yes |
| Grade Dummies (11) | -- | no | yes | yes | yes | yes |
| $R^2$ | -- | .376 | .913 | .921 | .926 | .925 |
| SEE | -- | .204 | .076 | .073 | .070 | .071 |

Notes: [a] Standard errors are enclosed in parentheses below coefficient estimates.

[b] Mean [S.D.] = 10.156 [.258].

individual in the same grade level with no pre-company experience and a person with one standard deviation more than the mean number of years of company service would earn 12 percent more than an otherwise comparable individual in the same grade level with one standard deviation less.[15]

Having established the existence of substantial within-grade returns to pre-company experience and to company service among managers and professionals at Company C, our third question was whether the higher within-grade earnings of more experienced employees were justified in terms of their higher relative within-grade productivity. If more experienced employees in a grade level in fact tended to receive higher salaries because they were typically more productive, one would expect the introduction of variables capturing relative within-grade performance into the ln(annual salary) regression to move the estimated within-grade returns to experience towards zero.

Formally, suppose that experience were one of the things having a positive impact on relative within-grade productivity:

$$2) \quad P = \beta_o + \beta_1 X_1 + \underset{\sim}{\beta_2} \underset{\sim}{X_2} + \underset{\sim}{\beta_3} \underset{\sim}{X_3} + \varepsilon,$$

where P represents relative within-grade productivity, $X_1$ represents experience, $\underset{\sim}{X_2}$ represents a vector containing educational attainment dummies, region dummies and grade level dummies, $\underset{\sim}{X_3}$ represents additional factors which affect productivity, the $\beta$'s represent parameters, and $\varepsilon$ represents the equation error. In a standard earnings function:

$$3) \quad \ln(Y) = \theta_0 + \theta_1 X_1 + \underset{\sim}{\theta_2} \underset{\sim}{X_2} + \mu,$$

where Y represents annual salary, the $\theta$'s represent parameters and $\mu$ represents the equation error, the experience variable typically has a significant positive coefficient. If, however, more experienced persons within a grade earn more solely because they are more productive, when

a measure of relative within-grade productivity is introduced into the earnings function:

$$4) \quad \ln(Y) = \alpha_0 + \alpha_1 X_1 + \underset{\sim}{\alpha_2 X_2} + \alpha_3 P + \gamma,$$

where the $\alpha$'s represent parameters and $\gamma$ is the error term, the estimated coefficient on the experience variable should fall to zero. According to this model, the expected value of the estimated coefficient on experience in equation (3) should be:

$$5) \quad \hat{\theta}_1 = b_{PX_1 \cdot \underset{\sim}{X_2}} \alpha_3,$$

where $b_{PX_1 \cdot \underset{\sim}{X_2}}$ represents the coefficient on experience obtained from the auxiliary repression of relative within-grade productivity on experience and the $\underset{\sim}{X_2}$ variables. Once productivity is controlled for, experience should have no independent effect on ln(annual salary), hence $\alpha_1$ in equation (4) should equal zero.[16]

Regression 3 contains a performance dummy which is equal to 1 if the employee received any one of the top three performance ratings in 1976 and equal to 0 otherwise; introducing this dummy should capture relative performance with a degree of precision close to that obtained in the comparable experiments reported in Medoff and in Medoff and Abraham, since at each of the three companies whose data were analyzed in those studies between 85 and 95 percent of the sample members received one of only two ratings. Contrary to what would be predicted by the human capital model of experience-earnings differentials, controlling for whether a manager or professional did or did not receive one of the top three performance ratings actually increases the estimated within-grade return to both pre-company experience and company service (although not by a statistically significant amount).

This finding is susceptible to the criticism that the higher

within-grade earnings of more experienced employees could still be explicable on the basis of their higher relative within-grade productivity since it could be the case that within each of the two large performance categories captured by the crude performance dummies the more experienced employees were the best performers. However, the introduction of finer controls for relative within-grade productivity, five performance rating dummies in regression 4 and percentile rank in regression 5, produces results which strongly suggest that this criticism has little validity; rather than moving the estimated return to experience towards zero, replacement of the crude performance dummy with either of the more precise measures of relative within-grade performance actually moves the estimated return to both pre-company experience and company service even further away from zero (although not by a statistically significant amount).[17,18]

An Alternative Experimental Design

Table 2 offers an alternative approach to untangling the inter-relationship of experience, relative within-grade performance and relative within-grade earnings among managers and professionals at Company C. Three measures of each employee's 1976 within-grade performance (columns 1, 3 and 5) and three comparable measures of each employee's 1976 within-grade salary percentile position (columns 2, 4 and 6) are regressed on a right hand side vector which contains educational attainment dummies, pre-company experience and its square, company service and its square, region dummies, and grade level dummies. Provided that all employees in each grade level had borne the same proportion of the costs of their on-the-job training, standard human capital theory implies that the within-grade ordering of employees by performance as of any point in time should be the same as the within-grade rank ordering of employees by salary as of the same

Table 2: Influence of Education and Experience on the Within-Grade Performance Percentile and Salary Percentile Position of Managers and Professionals[a]

(N = 8,238)

| | Dependent Variable = 1976 Two-Category Rating Percentile Variable[b] (1) | Dependent Variable = 1976 Salary Variable Corresponding to Two-Category Rating Percentile Variable[b] (2) | Dependent Variable = 1976 Six-Category Rating Percentile Variable[b] (3) | Dependent Variable = 1976 Salary Variable Corresponding to Six-Category Rating Percentile Variable[b] (4) | Dependent Variable = 1976 Ranking Decile Variable[b] (5) | Dependent Variable = 1976 Salary Variable Corresponding to Ranking Decile Variable[b] (6) |
|---|---|---|---|---|---|---|
| Mean [Standard Deviation] of Dependent Variable | .501 [.243] | .548 [.235] | .501 [.281] | .502 [.280] | .502 [.287] | .502 [.287] |
| Less than High School Diploma (yes = 1) | .100 (.019) | -.108 (.016) | .120 (.021) | -.179 (.019) | .136 (.022) | -.189 (.020) |
| High School Diploma (yes = 1) | .104 (.007) | -.033 (.006) | .128 (.008) | -.058 (.007) | .135 (.008) | -.059 (.007) |
| Masters Degree (yes = 1) | -.003 (.008) | .050 (.007) | -.006 (.009) | .057 (.008) | .001 (.009) | .056 (.008) |
| Doctorate (yes = 1) | -.025 (.013) | .116 (.011) | -.036 (.015) | .147 (.014) | -.031 (.015) | .148 (.014) |
| Years of Pre-Company Experience/10 | -.123 (.011) | .072 (.010) | -.160 (.013) | .103 (.012) | -.158 (.013) | .107 (.013) |
| (Years of Pre-Company Experience)$^2$/100 | .011 (.005) | .006 (.004) | .015 (.005) | .009 (.005) | .016 (.005) | .009 (.005) |
| Years of Company Service/10 | -.051 (.010) | .229 (.008) | -.076 (.011) | .258 (.010) | -.063 (.011) | .268 (.010) |
| (Years of Company Service)$^2$/100 | -.009 (.002) | -.031 (.002) | -.009 (.003) | -.027 (.002) | -.012 (.003) | -.028 (.002) |
| Constant | yes | yes | yes | yes | yes | yes |
| Region Dummies (3) | yes | yes | yes | yes | yes | yes |
| Grade Dummies (11) | yes | yes | yes | yes | yes | yes |
| $R^2$ | .123 | .333 | .158 | .302 | .160 | .314 |
| SEE | .228 | .192 | .258 | .235 | .264 | .238 |

Notes: [a] Standard errors are enclosed in parentheses below coefficient estimates.

[b] The procedure used to construct the dependent variables is detailed in Section II of the text.

date. Hence, the estimated pre-company experience and company service coefficients in each Table 2 performance equation should equal the estimated pre-company experience and company service coefficients in the corresponding Table 2 salary equation.

This argument holds even if important factors are omitted from the list of explanatory variables included in the Table 2 regressions. Suppose, for example, that innate ability exerts an important influence on a manager or professional's relative within-grade productivity:

$$6) \quad P = \beta_0 + \beta_1 X_1 + \underset{\sim}{\beta_2} \underset{\sim}{X_2} + \beta_3 A + \varepsilon,$$

where P represents within-grade performance percentile, $X_1$ represents experience, $\underset{\sim}{X_2}$ represents a vector containing educational attainment dummies, region dummies and grade level dummies, A represents uncaptured individual-specific differences, say in innate ability, which do not change over time, the $\beta$'s represent coefficients to be estimated and $\varepsilon$ represents the equation error. Let relative earnings be a function of the same variables:

$$7) \quad S = \theta_0 + \theta_1 X_1 + \underset{\sim}{\theta_2} \underset{\sim}{X_2} + \theta_3 A + \mu',$$

where S represents within-grade percentile earnings position, the $\theta$'s represent coefficients to be estimated, $\mu'$ represents the equation error, and $X_1$, $\underset{\sim}{X_2}$ and A are as in equation 6. If no measure of ability is available so that A is omitted both from the performance equation and from the salary equation, the estimated experience coefficients will have expected values:

$$8) \quad \hat{\beta}_1 = \beta_1 + b_{AX_1 \cdot \underset{\sim}{X_2}} \beta_3$$

and

$$9) \quad \hat{\theta}_1 = \theta_1 + b_{AX_1 \cdot \underset{\sim}{X_2}} \theta_3,$$

where $b_{AX_1 \cdot \underset{\sim}{X_2}}$ represents the coefficient on experience from an auxiliary regression of ability on experience and the $\underset{\sim}{X_2}$ variables. If the rank

ordering of employees by performance within each grade level were in fact equal to the rank ordering of employees by salary within each grade level, $\beta_1$ would equal $\theta_1$ and $\beta_3$ would equal $\theta_3$, so that $\hat{\beta}_1$ would equal $\hat{\theta}_1$. Furthermore, if unobserved ability were to affect performance position and salary position in the same way ($\beta_3$ equal to $\theta_3$) but experience were to have a different effect on relative performance than on relative salary ($\beta_1$ not equal to $\theta_1$), even though $\hat{\beta}_1$ and $\hat{\theta}_1$ would be biased estimators of $\beta_1$ and $\theta_1$, $\hat{\beta}_1 - \hat{\theta}_1$ would be an unbiased estimator of $\beta_1 - \theta_1$.[19,20]

Examination of the Table 2 regressions quickly reveals that the experience coefficients in the performance regressions differ greatly from the experience coefficients in the corresponding earnings regressions. Regressions 1 and 2 imply that, at the mean of the relevant variables, an additional year of pre-company experience is associated with a 1 point reduction in within-grade performance percentile but with a 1 point increase in within-grade salary percentile (a divergence of 2 points) and that an additional year of company service is also associated with a 1 point reduction in within-grade performance percentile and a 1 point increase in within-grade salary percentile (a divergence of 2 points).

One might argue that the use of dependent variables constructed by making one cut in each within-grade performance distribution and in each within-grade salary distribution as is done in regressions 1 and 2 could produce misleading results; if more experienced employees in each grade level tended to be towards the top within either or both of the two performance categories and/or towards the bottom within either or both of the two salary categories on which the regression 1 and regression 2 percentile variables were based, the divergence between performance and salary estimated to be associated with additional years of both pre-

company experience and company service could be seriously overstated. However, the results of comparable regressions with dependent variables based on six cuts (regressions 3 and 4) and ten cuts (regressions 5 and 6) in each within-grade performance distribution and in each within-grade salary distribution indicate that using more than two cuts in each relevant distribution actually enhances both the estimated negative effect of experience on within-grade performance percentile and the estimated positive effect of experience on within-grade salary percentile. Regressions 3 and 4 imply that for an employee with the mean amount of both pre-company experience and company service, an additional year of either pre-company experience or company service is associated with a 3 point divergence between performance percentile and salary percentile; regressions 5 and 6 imply a 3 point divergence associated with an additional year of pre-company experience at the mean and a 2 point divergence associated with an additional year of company service at the mean. It appears unarguable that the cross sectional relationship between labor force experience and relative within-grade performance differs substantially from the cross sectional relationship between labor force experience and relative within-grade earnings.[21,22]

On the Nature of Performance Ratings

The finding that experience bears a different partial relationsip to performance percentile than to salary percentile could conceivably be the result of the performance measures not accurately capturing true relative productivity in the year of appraisal. Possible problems with performance measures which might lead to distorted representation of the relative productivity of employees in a grade level include leniency in evaluations, divergent standards of evaluation on the part of different supervisors and age discrimination in awarding evaluations. These potentially worrisome flaws were discussed in Section I and dismissed as unlikely to be

the cause of any significant experience-related bias in the performance ratings and rankings used in this paper.

It seems probable that, all else equal, those managers and professionals who are relatively more productive in their current jobs should typically be viewed as more likely to succeed in a higher level job than those who are relatively less productive in their current jobs; additionally, among the group of employees in any grade level who are perceived by their supervisors as likely ultimately to reach any given higher grade level those with the highest productivity at a given point in time would typically be the most likely to be promoted at that point in time or in the near future. Thus, high productivity employees in any grade level should have a higher probability of receiving a promotion than otherwise comparable low productivity employees in the same grade. Similarly, it seems probable that, controlling for a list of other factors including current rate of compensation, those employees in any grade level who are most productive would receive larger salary increases than employees in the same grade level who are less productive.

One method of assessing whether the Company C performance ratings and rankings really reflect employees' relative within-grade productivity is first to assume that, controlling for other factors, in particular seniority, the company is indeed more likely to promote and give larger raises to more productive employees in a given grade level, and subsequently to interpret strong positive partial relationships between measured performance and probability of promotion and between measured performance and size of raise as good evidence that the Company C performance measures are significantly correlated with true relative within-grade current productivity. Table 3 presents results which indicate that Company C employees with high performance ratings are in fact more likely to be promoted and do tend to receive larger salary

increases than those receiving lower performance ratings.[23] Equation 1 indicates that, holding education, pre-company experience, company service, region, and grade level constant, employees who received either of the top two performance ratings in 1973 were significantly and substantially more likely to have been promoted between year-end 1973 and year-end 1974 than those who received any of the lower performance ratings while those who received any of the bottom five performance ratings in 1973 were significantly and substantially less likely to have been promoted between year-end 1973 and year-end 1974 than those who received any of the higher performance ratings. A similar statement based on equation 2 can be made regarding the partial relationship between 1973 performance rating and likelihood of having been promoted between year-end 1974 and year-end 1977.

Equation 3 demonstrates the existence of an almost perfectly monotonic partial relationship between 1973 performance rating and size of percentage increase in salary received during calendar year 1974. The point estimates of the coefficients imply that (with the exception of the difference between the percentage salary increase associated with receiving the second lowest performance rating and the percentage salary increase associated with receiving the bottom performance rating), holding constant education, pre-company experience, company service, region, grade level and ln(annual salary) as of the end of 1973, receiving any given rating rather than the next lower rating was associated with roughly a two point increment to an employee's percentage increase in salary.[24] According to the equation 4 results, persons who received either of the top two performance ratings in 1973 tended to receive larger cumulative percentage increases in salary between year-end 1974 and year-end 1977 than those who received any of the lower performance ratings, while those who received any of the bottom five performance ratings in 1973 tended to receive smaller cumulative

Table 3: Estimates of the Effect of Schooling, Experience and Performance Rating on the Probability of Being Promoted and on Percentage Salary Increases[a]

| | Dependent Variable = 1 if Promoted Between Year-end 1973 and Year-end 1974, 0 Otherwise (N = 5,575)[b] (1) | Dependent Variable = 1 if Promoted Between Year-end 1974 and Year-end 1977, 0 Otherwise (N = 5,606)[b] (2) | (Percent Salary Increase)/100; Year-end 1973 to Year-end 1974 (N = 5,606)[b] (3) | (Percent Salary Increase)/100; Year-end 1974 to Year-end 1977 (N = 5,606)[b] (4) |
|---|---|---|---|---|
| Mean [Standard Deviation] of Dependent Variable | .182 | .364 | .194 [.057] | .243 [.091] |
| Less than High School Diploma (yes = 1) | -1.389 (.484) | .025 (.244) | -.006 (.005) | .017 (.008) |
| High School Diploma (yes = 1) | -.543 (.108) | -.225 (.084) | -.007 (.002) | .003 (.003) |
| Masters Degree (yes = 1) | .264 (.113) | .180 (.099) | .007 (.002) | -.003 (.004) |
| Doctorate (yes = 1) | .459 (.192) | .360 (.172) | .015 (.004) | -.011 (.006) |
| Years of Pre-Company Experience/10 | -.430 (.183) | -.440 (.150) | .014 (.003) | -.015 (.005) |
| (Years of Pre-Company Experience)$^2$/100 | .145 (.087) | .028 (.072) | -.001 (.001) | -.002 (.003) |
| Years of Company Service/10 | -.510 (.168) | -.926 (.139) | .006 (.003) | -.061 (.005) |
| (Years of Company Service)$^2$/100 | .065 (.047) | .114 (.037) | .002 (.001) | .010 (.001) |
| ln(Annual Salary) | -- | -- | -.356 (.008) | -.169 (.013) |
| Performance Rating 1 (Worst; yes = 1)[c] | -- | -- | -.084 (.045) | -.038 (.079) |
| Performance Rating 2 (yes = 1) | -- | -- | -.092 (.006) | -.041 (.011) |
| Performance Rating 1 or 2 (yes = 1)[d] | -.866 (.424) | -.594 (.320) | -- | -- |
| Performance Rating 3 (yes = 1) | -1.449 (.174) | -.850 (.119) | -.069 (.002) | -.044 (.004) |
| Performance Rating 4 (yes = 1) | -1.078 (.122) | -.462 (.092) | -.036 (.002) | -.042 (.003) |
| Performance Rating 5 (yes = 1) | -.587 (.095) | -.345 (.080) | -.026 (.002) | -.018 (.003) |
| Performance Rating 7 (yes = 1) | .516 (.118) | .159 (.112) | .025 (.002) | .007 (.004) |
| Performance Rating 8 (Best; yes = 1) | 1.018 (.398) | .746 (.406) | .042 (.008) | .041 (.014) |
| Region Dummies (3) | yes | yes | yes | yes |
| Grade Dummies (11)[e] | yes | yes | yes | yes |
| $\chi^2$ | 709.60 | 878.77 | -- | -- |
| d.f. | 27 | 28 | -- | -- |
| $R^2$ | -- | -- | .375 | .239 |
| SEE | -- | -- | .045 | .079 |

Table 3: Notes

/a Standard errors are enclosed in parentheses below coefficient estimates.

/b Equations 1 and 2 were estimated using a maximum likelihood logit procedure, equations 3 and 4 using OLS. All the values of the independent variables in all four equations are as of December 1973.

/c No one who received the lowest of the nine possible ratings in 1973 was still with the company at the end of 1977; thus performance rating 1 is actually the second lowest possible rating.

/d Since no one who received performance rating 1 in 1973 was promoted between year-end 1973 and year-end 1977, the performance rating 1 and performance rating 2 groups were merged in estimating the logit equations.

/e Individuals in one grade where no one received a promotion between year-end 1973 and year-end 1974 were excluded from the equation 1 sample.

percentage increases in salary between year-end 1974 and year-end 1977 than those receiving any of the higher ratings.

Current Productivity or Potential?

Doubters might still argue that, rather than measuring just a manager or professional's current productivity, performance ratings could also be partially dependent upon a person's potential for advancement with the company. As the results discussed previously imply, holding schooling, region and grade level constant, employees with more seniority and pre-company experience generally have a lower probability of advancing in the corporate hierarchy than employees with less; this reflects within-grade heterogeneity in innate ability to handle a higher level job, "state dependence", and other factors which cause the probability of promotion in period t to be positively related to the frequency with which that event has occurred in previous periods.[25] For whatever reason, the fact that more experienced (older) employees in any given grade level typically are less likely to be promoted than less experienced (younger) employees in the same grade level implies that if performance ratings were a function of both likelihood of advancement and current productivity, the performance ratings of more experienced (older) employees might typically provide a downward biased measure of their true relative productivity in the year of appraisal.

Given that at Company C supervisors are asked to provide separate assessments of each employee's current performance and potential for advancement, and given that, as stated above, they are told in doing their performance evaluations that "career potential and promotability should not enter into ratings of an individual's performance," it seems implausible that supervisors' performance assessments would be reflecting promotability. If a supervisor wished to communicate to

higher management that a particular employee would do well in a higher level job, the message would not have to be conveyed indirectly by awarding the employee a higher-than-deserved performance rating but could be conveyed directly by giving him a high potential score, which is most certainly the way Company C seems to want this information transmitted. The equation estimates presented in Table 4 lend strong support to the belief that the performance ratings and the potential measure mirror very different things, with the former reflecting the level of an employee's relative productivity in his current assignment and the latter the extent to which the employee is likely to advance in the Company C hierarchy.

Equations 1 through 3 allow us to examine the partial effects (holding education, pre-company experience, company service and grade level constant) of 1976 performance (equation 1), 1976 potential (equation 2) and 1976 performance plus 1976 potential (equation 3) on the probability of an employee being promoted between year-end 1976 and year-end 1977.[26] The results obtained indicate that performance makes a significant contribution to explaining probability of promotion when added to the promotion equation with potential already present and that potential makes a significant contribution to explaining probability of promotion when added to the promotion equation with performance already present. By using the logit analogue to the multiple-correlation coefficient in the linear statistical model ($\rho^2$) presented and discussed in McFadden, we can ascertain the relative explanatory contributions of the potential and performance measures.[27] It turns out the change in $\rho^2$ associated with the potential variable is more than twice as large as the change in $\rho^2$ associated with the performance dummies. Thus, as would most likely be expected if the potential variable were directly capturing likely future success while the performance ratings were capturing current level of

Table 4: Current Performance versus Potential[a]

| | Dependent Variable = 1 if Promoted Between Year-end 1976 and Year-end 1977, 0 Otherwise (N = 7,567)[b] | | | Dependent Variable = ln Annual Salary as of December 1976) (N = 8,235)[b] | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Less than High School Diploma (yes = 1) | -.475<br>(.368) | -.135<br>(.365) | -.315<br>(.368) | -.069<br>(.006) | -.054<br>(.006) | -.069<br>(.006) |
| High School Diploma (yes = 1) | -.325<br>(.103) | .115<br>(.104) | -.097<br>(.107) | -.030<br>(.002) | -.011<br>(.002) | -.029<br>(.002) |
| Masters Degree (yes = 1) | .063<br>(.097) | -.097<br>(.101) | -.055<br>(.101) | .026<br>(.002) | .021<br>(.003) | .026<br>(.002) |
| Doctorate (yes = 1) | .222<br>(.158) | .219<br>(.159) | .248<br>(.161) | .057<br>(.004) | .052<br>(.004) | .057<br>(.004) |
| Years of Pre-Company Experience/10 | -.177<br>(.164) | -.006<br>(.002) | .078<br>(.168) | .054<br>(.004) | .045<br>(.004) | .054<br>(.004) |
| (Years of Pre-Company Experience)$^2$/100 | -.030<br>(.072) | -.001<br>(.074) | -.019<br>(.073) | .002<br>(.001) | .003<br>(.002) | .002<br>(.001) |
| Years of Company Service/10 | -.673<br>(.141) | .427<br>(.162) | .222<br>(.164) | .119<br>(.003) | .137<br>(.004) | .122<br>(.003) |
| (Years of Company Service)$^2$/100 | .085<br>(.037) | -.120<br>(.040) | -.062<br>(.041) | -.013<br>(.000) | -.018<br>(.001) | -.014<br>(.001) |
| Performance Rating 1 (Worst; yes = 1) | -1.664<br>(.729) | -- | -1.342<br>(.732) | -.116<br>(.008) | -- | -.114<br>(.008) |
| Performance Rating 2 (yes = 1) | -1.470<br>(.171) | -- | -1.044<br>(.176) | -.083<br>(.003) | -- | -.082<br>(.003) |
| Performance Rating 3 (yes = 1) | -1.078<br>(.107) | -- | -.707<br>(.112) | -.045<br>(.002) | -- | -.044<br>(.002) |
| Performance Rating 4 (yes = 1) | -.591<br>(.090) | -- | -.361<br>(.093) | -.027<br>(.002) | -- | -.026<br>(.002) |
| Performance Rating 6 (Best; yes = 1) | .656<br>(.100) | -- | .354<br>(.107) | .032<br>(.003) | -- | .031<br>(.003) |
| ln(Mean Annual Salary as of December 1976 in Potential Grade)[c] | -- | 3.935<br>(.219) | 2.938<br>(.242) | -- | .098<br>(.006) | .011<br>(.006) |
| Constant | yes | yes | yes | yes | yes | yes |
| Region Dummies (3) | yes | yes | yes | yes | yes | yes |
| Grade Dummies (11) | yes | yes | yes | yes | yes | yes |
| $R^2$ | -- | -- | -- | .926 | .916 | .926 |
| SEE | -- | -- | -- | .070 | .075 | .070 |
| $\chi^2$ | 877.16 | 944.15 | 1,033.68 | -- | -- | -- |
| d.y. | 27 | 23 | 28 | -- | -- | -- |

Notes: [a] Standard errors are enclosed in parentheses below coefficient estimates.

[b] Equations 1, 2 and 3 were estimated using a maximum likelihood logit procedure; equations 4, 5 and 6 using OLS. The mean of the dependent variable in the logit equations is .151; the mean [standard deviation] of the dependent variable in the OLS equations is 10.156 [.258]. All values of the independent variables in all six equations are as of December 1976.

[c] Mean [S.D.] = 10.320 [.318].

relative productivity, the potential variable plays a much more important role in explaining promotions than do the performance rating dummies.

The coefficient estimates obtained in equations 4 through 6 reveal the partial effects (holding education, pre-company experience, company service and grade level and region constant) of 1976 performance (equation 4), 1976 potential (equation 5 ) and 1976 performance plus 1976 potential (equation 6) on ln(annual salary as of December 1976). While the performance rating dummies make a highly significant incremental contribution to $R^2$ when added to the equation with the potential variable already present, the potential variable does not make a significant incremental contribution to $R^2$ when added to the equation with the performance rating dummies already present. This is exactly what would be expected if the performance ratings were reflecting current productivity and the potential variable were reflecting likely future value to the company and if, holding demographic characteristics, in particular company service, constant, salary decisions were influenced by employees' current productivity much more than by their likely future productivity.[28]

Longitudinal Evidence

It must be conceded that the cross-sectional results presented above and in Medoff and Medoff and Abraham could be arguably consistent with the tenets of human capital theory. While many economists might initially be surprised by the finding that more experienced managers and professionals tend to earn relatively more and perform relatively worse than less experienced managers and professionals in the same grade level, the alert human capital theorist would undoubtedly be able to respond with a quick defensive parry. In particular, if the share of on-the-job training costs borne by longer tenure employees in any grade level were significantly above the share of on-the-job training costs borne by shorter tenure employees in the same grade level, the observed cross-sectional pattern

might be produced even though each individual employee were being compensated in accordance with what human capital theory would predict.

More formally, suppose that all employees were paid an amount which depended on the value of their marginal productivity and on the share of their on-the-job-training which they had financed:

$$10) \quad Y = [1 - f(X_1)]P,$$

where Y represents an employee's wage, P represents the value of the employee's marginal product, $X_1$ represents experience, and f is a function of $X_1$ such that $0 \leq f(X_1) \leq 1$ and $f'(X_1) < 0$. The stated condition on the sign of the derivative of f is hypothesized to result from older employees having financed a larger proportion of their on-the-job training than younger employees. Taking the logarithm of both sides of equation 10 and assuming that $f(X_1)$ is small enough such that $\ln[1 + f(X_1)] \approx f(X_1)$ we get:

$$11) \quad \ln(Y) \approx -f(X_1) + \ln(P),$$

which implies that, holding productivity constant, older employees should be paid more than younger employees.

While the preceding story provides a possible (although somewhat implausible) explanation of the cross sectional results presented in Medoff, in Medoff and Abraham, and thus far in this paper,[29] it could not explain a finding that over time the relative earnings of individual managers and professionals tended to rise while the relative performance of those same managers and professionals failed to keep pace or even tended to fall. We have been able to devise no reasonable human capital explanation for the findings presented in Table 5.

The Table 5 results were generated under the assumption that both within-grade performance percentile in any year and within-grade salary

Table 5: Movement in Performance Ratings and Rankings Over Time for Individuals Staying in the Same Grade [a]

| Period | Fraction of Those Active 1972 to Relevant End Year Not Changing Grade During Period [b] | Change in Within-Grade Performance Rating Percentile Variable for Stayers [c] | Change in Salary Variable Corresponding to Performance Rating Percentile Variable for Stayers [c] | Change in Ranking Decile Variable for Stayers [c] | Change in Salary Variable Corresponding to Ranking Decile Variable for Stayers [c] |
|---|---|---|---|---|---|
| 1972-1973 | .789 | .008<br>(.005) | .059<br>(.005) | -.005<br>(.006) | .061<br>(.005) |
| 1972-1974 | .608 | .005<br>(.009) | .108<br>(.009) | -.012<br>(.010) | .108<br>(.008) |
| 1972-1975 | .513 | -.004<br>(.012) | .130<br>(.012) | -.024<br>(.013) | .137<br>(.012) |
| 1972-1976 | .393 | -.019<br>(.016) | .156<br>(.016) | -.041<br>(.017) | .157<br>(.015) |

| Period | Fraction of Those Active 1972 to 1976 Not Changing Grade 1972 to 1976 [b] | Change in Within-Grade Performance Rating Percentile Variable for Stayers [c] | Change in Salary Variable Corresponding to Performance Rating Percentile Variable for Stayers [c] | Change in Ranking Decile Variable for Stayers [c] | Change in Salary Variable Corresponding to Ranking Decile Variable for Stayers [c] |
|---|---|---|---|---|---|
| 1972-1973 | .393 | -.005<br>(.010) | .044<br>(.011) | -.021<br>(.012) | .045<br>(.009) |
| 1972-1974 | .393 | -.012<br>(.013) | .088<br>(.012) | -.030<br>(.015) | .087<br>(.012) |
| 1972-1975 | .393 | -.018<br>(.015) | .112<br>(.014) | -.032<br>(.016) | .115<br>(.014) |
| 1972-1976 | .393 | -.019<br>(.016) | .156<br>(.016) | -.041<br>(.017) | .157<br>(.015) |

Notes: [a] Standard errors are enclosed in parentheses below change estimates.

[b] The number of population members active throughout the relevant period was 7,547 for 1972-1973, 6,603 for 1972-1974, 6,170 for 1972-1975 and 5,689 for 1972-1976.

[c] The procedure used to construct the performance rating percentile variables, ranking decile variables and two sets of corresponding salary variables on which the change estimates are based is detailed in Section II. The reported change estimates were obtained by regressing the change in the relevant variable for each stayer on a constant, change in the square of company service and change in the square of time in grade, then calculating the predicted change for a person with mean stayer characteristics.

percentile in any year could be written as functions of observable and unobservable characteristics of each managerial and professional employee:

$$12)\quad P_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{1t}^2 + \underset{\sim}{\beta_3} \underset{\sim}{X_{2t}} + \beta_4 A + \varepsilon_t$$

and

$$13)\quad S_t = \theta_0 + \theta_1 X_{1t} + \theta_2 X_{1t}^2 + \underset{\sim}{\theta_3} \underset{\sim}{X_{2t}} + \theta_4 A + \mu_t,$$

where $P_t$ represents within-grade performance percentile as of time t, $S_t$ represents within-grade salary percentile as of time t, $X_{1t}$ represents length of company service as of time t, $\underset{\sim}{X_{2t}}$ represents a vector containing pre-company experience and its square, educational attainment dummies, region dummies and grade level dummies for period t, A represent uncaptured individual differences in factors such as innate ability which affect relative performance and relative earnings, and the $\beta$'s and $\theta$'s represent parameters to be estimated. Given the assumed model, both within-grade performance percentile and within-grade salary percentile should move in a predictable fashion with the passage of time:

$$14)\quad P_{t+1} - P_t = \beta_1 + \beta_2[(X_{1t} + 1)^2 - X_{1t}^2] + \varepsilon_{t+1} - \varepsilon_t$$

and

$$15)\quad S_{t+1} - S_t = \theta_1 + \theta_2[(X_{1t} + 1)^2 - X_{1t}^2] + \mu_{t+1} - \mu_t$$

for those individuals not changing their educational status, region or grade level. Note that differencing eliminates A, so that estimation of equations 14 and 15 should yield estimates (applicable to the population not changing educational status, region or grade level over the relevant time period) of the effect of the passage of time on both within-grade performance percentile and within-grade salary percentile which are not biased by uncaptured individual-specific differences that affect performance and earnings.[30] The results presented in Table 5 were derived from regressions of change

in within-grade performance percentile and change in within-grade salary percentile on a constant term plus change in the square of company service.

The major conclusion which emerges from the Table 5 analysis is that the passage of time affects within-grade performance percentile much differently than it affects within-grade salary percentile. Both the results at the top of the page (based on four different samples including everyone in the original populations not changing grade between 1972 and 1973, 1972 and 1974, 1972 and 1975, and 1972 and 1976, respectively) and the results at the bottom of the page (based on one sample including everyone in the original populations not changing grade between 1972 and 1976) imply that (for those satisfying the relevant sample selection criteria and with the appropriate sample mean company tenure) within-grade rating percentile does not change significantly with the passage of from one to four years' time. For both sample selection alternatives, the predicted changes at the mean level of company service in within-grade ranking percentile were insignificantly negative for the one and two year periods, more substantially and significantly negative for the three and four year periods. It is interesting that for sample members with mean company tenure the predicted change over time in within-grade rating percentile is indistinguishable from zero while the predicted change over time in within-grade ranking percentile (at least for the three and four year periods) is substantially and significantly negative; this is the opposite of what an argument that managers and professionals might become more productive over time without moving into higher rating categories would lead one to expect.

The salary percentile results contrast sharply with the performance percentile results; for both sample selection alternatives, and for both salary percentile variables (one analagous to the rating percentile variable and the other analogous to the ranking percentile variable) at the mean level of

company service, the passage of one, two, three or four years time is associated with a substantial and significant upward movement in relative within-grade salary position. With time, then, for those persons not changing grade, relative within-grade performance appears to remain stable or deteriorate while relative within-grade salary rises substantially.

Thus, the Table 5 results imply that for many employees in our economy, the growth in relative earnings with the passage of time cannot be explained by the growth of relative performance, contrary to the human capital explanation of the experience earnings profile. Moreover, they suggest that with additional labor force experience the relative performance of a substantial number of employees actually deteriorates.[31]

## IV Conclusions and Directions

This study has presented the results of an experiment designed to test whether or not all but a small fraction of experience-earnings differentials can be explained by experience-productivity differentials. Under our assumption that supervisors' ratings of their subordinates adequately reflect the subordinates' true relative productivity in the year of appraisal, the answer is clearly no, at least for white male managerial and professional employees in the four major corporations we have studied.

Our primary finding is based on both cross-sectional and longitudinal personnel data files, which include performance ratings done under a number of different procedures, and in one case, a ranking by current performance of similarly-situated employees done collectively by all supervisors who have observed the employees' work during the relevant year. Our key assumption concerning performance ratings is supported by a seemingly substantial amount of institutional and econometric evidence presented above and in Medoff and Medoff and Abraham.

If our experimental design is taken as sound and our data as adequate for their task, we must begin to formalize theories of experience-earnings differentials that do not have experience-productivity differentials at their heart, design experiments under which those theories might be falsified, and gather the data required for this experimentation. It is our belief that major steps in these directions can be taken through interaction with those who formulate company compensation policies and with those affected by these policies. In particular, interviews with the members of top management who are responsible for the outlines of a company's pay practices should be conducted; discussions with supervisors about how they determine the proper salaries for their

subordinates should be initiated; and the attitudes of employees toward different compensation schemes (e.g., relative earnings reflect relative productivity versus relative service) should be assessed. Moreover, we should seek data which would permit analysis of the impact of changes in the nature of firms' compensation practices on things such as productivity, quits, ability to attract new hires, absenteeism, and job satisfaction.

Unlike physical scientists, economists typically are not involved in the collection of the data they use, and unlike other social scientists, economists generally avoid having contact with their units of observation. As a result, the proper data for testing numerous important beliefs that many economists hold have not been gathered and the knowledge of those who are likely to really know what is going on has been ignored. We feel that this most certainly has been true in the case of the belief that experience-earnings differentials can be explained by experience-productivity differentials. Thus, we contend that our evidence stands at odds both with this belief and the way in which it has been uncritically accepted.

## Footnotes

[1] By rated performance we mean an overall evaluation of how well an employee is currently fulfilling the requirements of his job, where this evaluation takes the form of a performance rating assigned by the employee's immediate supervisor. We have assumed that the relative rated performance of employees performing comparable work conveys information regarding the relative current productivity of those same employees. This key assumption is defended in Medoff, in Medoff and Abraham and later in this study.

[2] The human capital model of investment in on-the-job training is laid out in detail in Becker, pp. 16-37.

[3] Mincer has seniority provisions under collective baragining agreements in mind when he makes this statement, but his logic applies equally well to other institutional settings.

[4] For good general discussions of performance ratings and further references, see Barrett, Bass and Barrett, Landy and Trumbo, and McCormick and Tiffin.

[5] Research assessing cross-rater variation in performance ratings includes Borman and Whitlock. Lawler refers to Whitlock's work and asserts that "other studies have shown that raters tend to agree upon the weight to be assigned to.... different behavior specimens; thus inter-rater reliability is possible" (p. 371) but unfortunately does not specify what other studies he has in mind.

[6] The existence of variation in supervisor's rating standards would imply that performance ratings captured true performance with some measurement error:

$$1) \quad P = P^* + v,$$

where P represents rated performance, P* represents true performance and v represents the measurement error. If employees with different observed characteristics were not systematically paired with supervisors applying different sets of rating standards, the measurement error would be of the classical variety (E(v/observed characteristics) = 0). Our semilog earnings function with grade level controls and performance controls can be represented:

$$2) \quad \ln(y) = \alpha + \underset{\sim\sim}{\beta X} + \underset{\sim\sim}{\gamma G} + \delta P + \varepsilon,$$

where y represents annual salary, the vector $\underset{\sim}{X}$ captures educational attainment, pre-company experience, and company service, $\underset{\sim}{G}$ is a vector of grade level dummies, $\alpha$, $\underset{\sim}{\beta}$, $\underset{\sim}{\gamma}$ and $\delta$ represent parameters to be estimated, and $\varepsilon$ represents the equation error. The presence of classical measurement error in our performance variable would cause downward bias in $\hat{\delta}$ and also downward bias in the estimated return to both pre-company experience and company service resulting from the negative partial correlation between observed performance and each of those labor force experience variables. Hence, if the measurement error in the performance variable were reduced causing $\hat{\delta}$ to increase, the estimated pre-company experience and company service coefficients would also increase. For measurement error in the performance variable to be a plausible explanation of why the experience coefficients in our ln(earnings) regressions did not move towards zero when performance controls were introduced, the arguable bias in the experience coefficients would have had to have been in the opposite direction.

Classical measurement error in an ordinary least squares dependent variable would cause no bias in estimated ordinary least squares coefficients. It thus seems highly unlikely that measurement error in the performance variable used to create the categories employed on the left hand side of our multinomial logit performance equation could

have caused any substantial bias in our multinomial logit coefficient estimates.

[7]Two studies of blue collar workers, one by Rothe and one by Rundquist and Bittner, contain some evidence that the ratings of long service laundry workers and container inspectors tend to be inflated. A study by Ferguson of ratings given to assistant life insurance sales managers and another study by Stockford and Bissel of ratings given to first line supervisors at a manufacturing plant suggest that superiors tend to be more lenient in rating those whom they have known for a longer period of time.

[8]One might be tempted to look at the change over time in the hourly production of individual piece rate workers as a source of direct evidence regarding the shape of the experience-productivity profile among such workers. To the extent that factors other than a worker's experience which might affect productivity change with the passage of time (for example, workers might typically be given first claim on materials during periods of shortage or be allowed to choose better machines as they accrued more seniority) such an experiment would produce misleading results.

[9]Promotion equations with education dummies, pre-company experience and its square, company service and its square, region dummies and grade level dummies on the right hand side were estimated using data from one of the company files analyzed in Medoff and Abraham. The results implied that, for a person with sample mean characteristics as of July 1, 1974, an additional year of either pre-company experience or company service was associated with a 5 percent higher probability of receiving a promotion between July 1, 1975 and July 1, 1977, and that, for a person with sample mean characteristics as of July 1, 1976,

and additional year of either pre-company experience or company service was associated with a 6 percent higher probability of receiving a promotion between July 1, 1976 and July 1, 1977.

[10]The only imputed 1972 information used in the analysis presented in Section III was grade level and salary. The 1972 grade level imputation rested on the assumption that no Company C employee moved down in grade level or moved up more than one grade level between year-end 1972 and year-end 1973. Fewer than 5 percent of those included in both our 1973 and our 1974 populations moved down in grade level or moved up more than one grade level between year-end 1973 and year-end 1974. The 1972 salary imputation rested on the assumption that the two salary increases recorded for each Company C employee as of year-end 1973 captured all salary changes made during calendar year 1973. Beginning in 1975, as many as six salary increases were recorded for each employee. Fewer than 3 percent of those included in both our 1974 and our 1975 populations received more than two salary increases during calendar year 1975 and somewhat more than 1 percent of those included in both our 1975 and our 1976 populations received more than two salary increases during calendar year 1975. Thus, the assumptions made in order to construct the required fix on each employee's year-end 1972 status do not appear to have been unreasonable.

[11]It should be noted that grade level was not adjusted for those employees who changed grade level between year-end 1972 and year-end 1973 and also between June 30, 1972 and year-end 1972. Just over 1 percent of those included in both our 1973 and our 1974 populations changed grade level between year-end 1973 and year-end 1974 and also changed grade level between June 30, 1973 and year-end 1973. Adjusting the recorded grade level of those who changed grade during the last six months of a year by

subtracting one from end-of-year grade level seemed reasonable in light of the previously mentioned fact that very few employees moved down in grade level or up more than one grade level during a typical year.

[12]Complete performance information consisted of the employee's performance rating, the employee's rank among a group of his peers and the number of people in the group the employee was included in for ranking purposes. Employees with a rank group size of one were excluded from our populations; in 1976, this exclusion reduced the size of our population by approximately one tenth of 1 percent. The median rank group size in the 1976 population we used for analysis was 46.

[13]The Northeast region comprises Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island and Vermont; the North Central region Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; the South region Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia and West Virginia; and the West region Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah and Washington.

[14]The list of employees normally includes only persons in the same grade level, but may include persons in several adjacent grade levels if such broader coverage is needed in order to obtain a reasonably sized comparison group within any division.

[15]Straightforward calculations show that for a person with the mean number of years of pre-company experience, the within-grade return to an additional year of pre-company experience is roughly three times as large as the total return to that same year of pre-company experience. This implies that, all else equal, the additional year of pre-company experience tends to be

associated with holding a job in a lower grade level than would otherwise be predicted.

[16]For an excellent discussion of omitted variable bias, see Appendix C in Griliches and Ringstad.

[17]The Table 1 results imply the existence of a substantial and significant total return to education among managers and professionals at Company C, most of which is captured by the introduction of grade level dummies into the ln(earnings) regression; only 20 percent of the earnings differential between non-high school graduates and bachelors degree holders, 7 percent of the differential between high school graduates and bachelors degree holders, 24 percent of the differential between masters degree holders and bachelors degree holders, and 22 percent of the differential between doctorates and bachelors degree holders remains uncaptured by the grade level dummies. The estimated return to education in the equation with grade level controls does not move towards zero when first crude and then finer performance controls are introduced which suggests that this return cannot be explained by within-grade differences in rated performance.

[18]The results presented in equations 2 through 5 of Table 1 are not true within-grade results since the specification used permits only the regression intercept and not the coefficients on the education, experience, performance and region variables to vary across grades. True within-grade equations comparable to regressions 2 through 5 were estimated separately for each of the twelve grade levels represented in our population. In all twelve cases, having either an additional year of pre-company experience or an additional year of company service beyond the overall mean amount for the relevant variable was associated with a reasonably sized positive increment to earnings; in no case did

the estimated within-grade return to experience move towards zero when performance controls were introduced.

[19]Appendix C to Griliches and Ringstad discusses omitted variable bias.

[20]The difference $\hat{\beta}_1-\hat{\Theta}_1$ is a biased estimator of $\beta_1-\Theta_1$ if it is not true that unobserved ability affects relative performance in the same way as it affects relative salary so that $\beta_3$ and $\Theta_3$ are not equal. While the existence of a divergence between $\beta_1$ and $\Theta_1$ thus does not necessarily imply that experience affects performance differently than it affects salary, it does imply that either experience or ability affects relative performance differently than relative salary and hence calls into question theories which explain all earnings differentials in terms of productivity differentials.

[21]Controlling for labor force experience and grade level, education has a pronounced positive effect on within-grade salary percentile and a pronounced negative effect on within-grade performance percentile. Both the estimated positive effect of education on salary percentile and the estimated negative effect of education on performance percentile are stronger when finer measures of the dependent variables are employed.

[22]The equations presented in Table 2 are not really within-grade regressions since the inclusion of a set of grade dummies among the independent variables only permits a different intercept and not different effects of education, experience and region across grade levels. Separate within-grade equations comparable to the six regressions in Table 2 were estimated for each of the twelve grade levels represented in our population. In each of the twelve cases, the regression coefficients implied that, using any of the three sets of dependent variables, for a person with the overall mean number of years of both pre-company experience and company service, an additional year of either pre-company experience

or company service was associated with lower relative within-grade performance but higher relative within-grade salary.

[23]The promotion probability results were obtained using a maximum likelihood logit procedure. The logit model is discussed in Nerlove and Press.

[24]The high mean percentage salary increase observed during calendar year 1974 (roughly 19 percent) resulted from the granting of two across-the-board cost-of-living increases totalling approximately 12 percent during that year. Very similar results were obtained when percentage salary increase during calendar year 1975 (having a mean of 9 points) was regressed on year-end 1974 independent variables.

[25]Results based on one of the data sets analyzed in Medoff and Abraham indicating that more experienced employees in any grade level were likely to be promoted than similarly-educated less-experienced employees in the same grade level are discussed in footnote 8. Given the zero or negative partial correlation between rated performance and labor force experience at Company C, the promotion equation results presented in Table 3 imply that, controlling for education, region, and grade level but not for rated performance, more experienced Company C employees are significantly less likely to be promoted than similarly educated but less experienced Company C employees. The coefficient estimates obtained from equations of the same form as those in Table 3 but without performance rating controls implied that, for a person with the appropriate sample mean characteristics as of year-end 1973, an additional year of pre-company experience was associated with a 5 percent lower probability of receiving a promotion between year-end 1973 and year-end 1974 and with a 4 percent lower probability of receiving a promotion between year-end 1974 and year-end 1977; an additional year of company service was associated with a 4 percent lower probability of receiving a promotion between year-end 1973 and year-end 1974 and with a 4 percent lower probability of

receiving a promotion between year-end 1974 and year-end 1977.

[26]Equations 1 through 3 were estimated using the same maximum likelihood logit procedure as was used to estimate the promotion equations presented in Table 3.

[27]The formula for $\rho^2$ is given on page 121 of the McFadden article:

$$(1) \quad \rho^2 = 1 - \frac{L(\hat{\Theta})}{L(\hat{\Theta}^H)}$$

where $L(\hat{\Theta})$ is the log likelihood of the unrestricted model and $L(\hat{\Theta}^H)$ is the log likelihood of the restricted model.

[28]Other promotion and ln(earnings equations were estimated using a set of dummy variables capturing each employee's potential grade level rather than ln(mean annual salary as of December 1976 in employee's potential grade level) as in the equations reported. Results very similar to those included in Table 4 were obtained. The results actually presented were chosen for expositional reasons.

[29]Very large decreases over time in the share of on-the-job training costs borne by employees would be required if the explanation just offered were to account for the cross sectional results we obtained. There seems to be no good reason to believe that changes in the necessary direction and of the necessary magnitude have occurred.

[30]One cannot draw inferences regarding absolute change over time in either performance or earnings based on observed change over time in either within-grade performance percentile or within-grade salary percentile, since there are likely to be shifts over time in the underlying within-grade productivity and earnings distributions on which the performance and salary percentile measures are based. However, it seems likely that any

major shifts in the underlying within-grade productivity and earnings distributions would occur slowly.

[31]As explained in the text, the change in performance percentile and the change in salary percentile figures presented in Table 5 were predicted values derived from equations estimated with change in percentile position as the dependent variable and a constant term plus change in the square of company service on the right hand side. Change in educational status dummies and change in region dummies were not included in the prediction equation because we did not really know educational status and region as of year-end 1972. Equations based on 1973 through 1976 data with change in education dummies and change in region dummies included as explanatory variables produced results very similar to those reported in Table 5.

Bibliography

Barrett, Richard S. Performance Rating. Chicago: Science Research Associates, Inc. 1966.

Bass, Bernard M. and Gerald V. Barrett. Man, Work and Organizations: An Introduction to Industrial and Organizational Psychology. Boston: Allyn and Bacon, Inc., 1972.

Becker, Gary S. Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education. New York: National Bureau of Economic Research, 1964.

Borman, Walter C. "Effect of Instructions to Avoid Halo Error on Reliability and Validity of Performance Evaluation Ratings," Journal of Applied Psychology, Vol. 60, No. 5, 1975.

Dalton, Gene W. and Paul H. Thompson. "Accumulating Obsolescence of Older Engineers," Harvard Business Review, September-October 1971.

Ferguson, Leonard W. "The Value of Acquaintance Ratings in Criterion Research," Personnel Psychology, Vol. 2, No. 1, 1949.

Griliches, Zvi and V. Ringstad. Economics of Scale and the Form of the Production Function. Amsterdam: North Holland, 1971.

Landy, Frank J. and Don A. Trumbo. Psychology of Work Behavior. Homewood, Illinois: The Dorsey Press, 1976.

Lawler, Edward E. "The Multitrait-Multirater Approach to Measuring Managerial Job Performance," Journal of Applied Psychology, Vol. 51, No. 5, 1967.

McCormick, Ernest J. and Joseph Tiffin. Industrial Psychology. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1974.

Medoff, James L. "The Earnings Function: A Glimpse Inside the Black Box," mimeo, 1977.

Medoff, James L. and Katharine G. Abraham. "Experience, Performance, and Earnings," mimeo, 1978.

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

Nerlove, Marc and S. James Press. Univariate and Multivariate Log-Linear and Logistic Models. Santa Monica, California: The Rand Corporation, 1973.

Rothe, H. F. "The Relation of Merit Ratings to Length of Service," Personnel Psychology, Vol. 2, 1949.

Rundquist, E. A. and Bittner, R. H. "Using Ratings to Validate Personnel Instruments: A Study in Method," Personnel Psychology, Vol. 1, 1948.

Stockford, Lee and H. W. Bissell. "Factors Involved in Establishing a Merit-Rating Scale," Personnel, Vol. 26, No. 2, 1949.

Whitlock, Gerald H. "Application of the Psychophysical Law to Performance Evaluation," Journal of Applied Psychology, Vol. 47, No. 1, 1963.