The Retention of High-Quality Personnel in the U.S. Armed Forces

Michael P. Ward, Hong W. Tan

Rand

The research described in this report was sponsored by the Office of the Assistant Secretary of Defense/Manpower, Installations and Logistics under Contract MDA903-83-C-0047.

Library of Congress Cataloging in Publication Data

Ward, Michael Paul, 1947-
The retention of high quality personnel in the U.S. armed forces.

"February 1985."
Bibliography: p.
"R-3117-MIL."
1. United States—Armed Forces—Recruiting, enlistment, etc. 2. United States—Armed Forces—Personnel management. I. Tan, Hong W. II. Rand Corporation. III. Title.
UB323.W35 1985 355.2'23'0973 84-15959
ISBN 0-8330-0591-X

The Rand Publication Series: The Report is the principal publication documenting and transmitting Rand's major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation

R-3117-MIL

The Retention of High-Quality Personnel in the U.S. Armed Forces

Michael P. Ward, Hong W. Tan

February 1985

Prepared for the
Office of the Secretary of Defense/
Manpower, Installations and Logistics

Rand
1700 MAIN STREET
P.O. BOX 2138
SANTA MONICA, CA 90406-2138

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

PREFACE

This report was prepared under Task Order 82-II-1 as part of the research of The Rand Corporation's Defense Manpower Research Center. The Center is sponsored by the Office of the Assistant Secretary of Defense (Manpower, Installations, and Logistics), OASD/MIL.

Manpower issues have acquired an increasing importance in defense planning and budgeting. The goal of the Center is to develop both broad strategies and specific solutions for dealing with present and future defense manpower problems. This necessitates the development of new methodologies for examining significant classes of manpower problems, as well as certain problem-oriented research.

The purpose of this study is to address the question: Does the military retain the best of its first-term recruits? Using data from the 1974 Entry Cohort File developed by the Defense Manpower Data Center, the authors generate an index of job performance that combines entry-level attributes of recruits—Armed Forces Qualification Test scores and level of education—with first-term promotion histories. This index is used to assess quantitatively the relative importance of these characteristics and other unobserved "ability factors" for evaluating the military's success in retaining "high-quality" personnel.

This report, and the methods described here, should be of interest to OASD/MIL policymakers in several respects: first, in helping to document the extent of the quality-loss phenomenon; second, in developing policies to retain high-quality personnel; and, finally, in providing a framework for studying reenlistment standards.

SUMMARY

The retention of skilled enlisted personnel has been a major issue facing military planners since the advent of the All Volunteer Force. Based on the military's experience from 1974 to 1981, over one-third of first-term enlistees attrit before finishing their term of service, and another one-third do not reenlist after successfully completing their first term. The remaining one-third, who form the basis of the career force, have exceeded reenlistment standards based on on-the-job performance as well as on education and scores on generalized ability tests,[1] and, of course, they have voluntarily chosen to continue on in the Services. We ask whether, at the end of this screening process, the Services retain "higher quality" recruits? Are standards set by the Services successful in screening for the "best" of those eligible to reenlist? Do these standards correspond to military "performance"?

As each of these questions suggests, the principal difficulty in answering these important policy questions is in defining "ability," or "quality," or "performance." Reenlistment standards reflect this ambiguity, stressing variously attainment of a high school diploma, test scores exceeding a minimum percentile, or attainment of a grade. Often these criteria are used in conjunction with explicit trade-offs among them.

This report addresses these issues by developing a methodology for combining performance-based measures with entry-level and background characteristics to construct a single combined measure of "quality" which we call the quality index. In this report, the performance measures used are promotion speed to E4 and E5, whereas the entry-level characteristics employed are education and the Armed Forces Qualification Test (AFQT). For the methodology we develop, the list can be expanded to encompass more measures. We estimate indexes of quality for each of eight military occupations, examining the trade-offs between the various components of the index. At the end of the first term we examine the distribution of the quality index for three groups: those who reenlist, those eligible for reenlistment who leave the Service, and those who are ineligible to reenlist, including those who attrit.

The study is based on a sample of non-prior-service males who entered active duty during fiscal year 1974. From this FY74 Entry Cohort File, we constructed each recruit's promotion history over the

[1]Other standards, relating to past criminal or antisocial behavior, are not considered explicitly in this work. Reenlistment standards, based on ability and performance, have changed over time and differ across Services.

1974–1981 period, and determined whether he attrited, separated at Expiration of Term of Service (ETS), or reenlisted at the end of the first term. We selected eight occupations for analysis, representing, for each Service, jobs that were specialized and highly skilled on the one hand and arduous on the other.

We found that the military is, in general, successful in retaining high-quality enlisted personnel, as measured by the quality index. Those lost through attrition have the lowest quality. Those who separate at ETS have about the same average quality index as the average of those entering the military. However, except for the Air Force, those entering the career force have the highest performance of any group. By comparison, differences in AFQT and schooling among these groups were very small.

We were also able to assess the effects of entry-level characteristics (education and AFQT) on job performance, and to evaluate their importance relative to other unobserved "ability factors." AFQT scores had a uniformly small quantitative effect on the quality index across occupations when compared to the presence or absence of a high school diploma. The trade-off between AFQT and a high school diploma ranged from a low of 33 points of AFQT for Navy radiomen to a high of over 200 points for Marine Corps aircraft mechanics. More importantly, these entry-level characteristics were usually a minor component of the quality index. For some Navy occupations almost all of the variation in quality, as we measure it, was explained by "ability factors" other than AFQT and education. The Air Force occupations represented the other extreme, possibly reflecting the importance AFQT and schooling play in that Service's promotion process.

These conclusions do not challenge the usefulness of AFQT and education as *enlistment* standards. In fact, our results confirm that these characteristics do significantly predict subsequent performance as measured by promotion rates. However, by the end of the first term, we find that the weight associated with these entry-level characteristics has fallen so that their usefulness as predictors is greatly reduced. The actual track record of the recruit is, by the end of the first term, a much better indicator of quality.

These results suggest that Service-wide guidelines on reenlistment eligibility should not be heavily weighted toward either mental category or educational achievement. The rank attained at the end of the first term, and perhaps the speed with which that rank is achieved, should receive much larger weight in these decisions.

This work should be viewed as only a first step toward answering the important policy question of how the military can attract and retain high-quality recruits, and how reenlistment standards should be

designed. Although much work has been done on general issues of attrition and retention,[2] little is known about the quality of those who leave the military and of those who stay. More importantly, little is known about the effects of policy variables in attracting and in retaining high-quality enlisted personnel. This report and the methods described herein help to define the extent of the quality-retention phenomenon problem, will form a basis for future research into the causes and consequences of this problem, and will aid in designing reenlistment standards to address this problem.

[2]See, for example, the collection of studies in H. W. Sinaiko et al., 1981.

ACKNOWLEDGMENTS

The results in this report were made possible through the enthusiastic support and cooperation of the Defense Manpower Data Center (DMDC), Monterey, California. Robbie Brandeary, Jane Crotzer, and Dennis Van Langdon of DMDC were instrumental in constructing the data used in this analysis, and instructed us in the institutional background so important for analyzing military records. Without their assistance this work could not have been accomplished.

Lt. Col. Harry Thie, project monitor, helped us to understand the Services' promotion process, and along with Rand reviewers James Dertouzos and William Rogers provided us with thoughtful and constructive criticism of earlier drafts of this report.

Roger Madison and Karl Schultz assisted ably in data management and programming.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS ix

FIGURES AND TABLES xiii

Section

I. INTRODUCTION 1

II. DATA AND OVERVIEW OF THE 1974 COHORT 4
The Data 4
Promotion Times and Job Performance: The 1974 Entry Cohort 6

III. THE MODEL 21
General Characteristics 21
Some Caveats 25

IV. EMPIRICAL RESULTS 27
Estimates of the Quality Index 27
Quality and Reenlistment 34

V. CONCLUSIONS 39

Appendix: THE MODEL: TECHNICAL DETAILS 43

REFERENCES 57

FIGURES

1. Percent Promoted to E5: Army MOS11C 10
2. Percent Promoted to E5: Army MOS91B 11
3. Percent Promoted to E5: Navy NECBT 12
4. Percent Promoted to E5: Navy NECRM 13
5. Percent Promoted to E5: Air Force AFSC42 14
6. Percent Promoted to E5: Air Force AFSC20 15
7. Percent Promoted to E5: Marine Corps MC03 16
8. Percent Promoted to E5: Marine Corps MC60 17
9. Quality Viewed as a Latent Variable 22
10. Effects of Heterogeneous Promotion Probabilities 23
11. Index of Quality: Effects of AFQT and Schooling 28
12. Index of Quality: Effects of Promotion Times 29

A.1. Functional Form for the Hazard Function 50
A.2. The Effect of Quality on the Hazard Function 51

TABLES

1. Selected Occupations From the Four Services 6
2. Rates of Attrition, Separation, and Reenlistment 7
3. Characteristics of Personnel by Outcome 9
4. The Effects of Entry Characteristics on the Performance Index 31
5. Analysis of Variation in the Performance Index 32
6. Effect of the Quality Index on Reenlistment 35
7. Effects of the Performance Index on Reenlistment: A Decomposition 37
8. Average Characteristics by Disposition of Entry Cohort 38

I. INTRODUCTION

Each year between 1974 and 1981 about 300,000 men and women entered the Armed Forces as enlisted personnel. About one-third of these left the military before fulfilling their first-term obligation and another one-third did not reenlist after successfully completing their first term. At the end of the first term, both these individuals and the Services make separate decisions about the suitability of their continued relationship. Individuals leave the Services voluntarily for perceived better prospects outside the military. At the same time, the Services set minimum reenlistment standards, which lead to the rejection of otherwise willing applicants. The end product of this selection determines the "quality" of those entering the career force.

Much of our analysis is devoted to answering the question: What is quality? In addition to multifaceted job performance characteristics, the ability to work in a timely fashion, to exercise initiative, and to work cooperatively with one's associates all contribute to any comprehensive definition of quality. Further, one's ability in any of these dimensions depends on the job itself, so that two people who might be ranked unambiguously in one job may have their order reversed in another job. In this view there are no "good" workers but only "good matches" of workers with jobs. Rather than referring to the quality of the worker, we ought to refer to the quality of the match and to rank personnel by how well they match with their own job. We adopt this view by studying the promotion histories of first-term recruits and use, in effect, the military's own internal ranking of individuals (compared with their peers in the same job) as one measure of quality.

The speed of promotion as well as the rank ultimately attained by a recruit is a function of the job (Military Occupational Specialty, or MOS), the military's own manning requirements, time in grade, time in service, and technical competency, qualities of leadership, timeliness, and compatibility. For these reasons it is not always correct to conclude that someone promoted sooner is necessarily "better" than someone promoted later or that an earlier promotion signals a better match unless the people being compared are affected by the same external environment. Thus, persons with similar characteristics and performance may be promoted at different rates because they entered in a different year or because they chose or were given an occupation or specialty whose aggregate manning table called for a different number of E4s and E5s than another.

To control for service and specialty differences in promotion rates, as well as changes in manning tables in different years, our analysis is restricted to a single entry cohort of recruits (FY74) and our comparisons of promotions are made only *within* military occupations. We do not directly compare people in different occupations or people who went through their first term during periods of different aggregate promotion prospects.

To fix these concepts, consider enlistees in a single job who have the same education, Armed Forces Qualification Test (AFQT) scores and, possibly, other observable background characteristics. Suppose that the first recruit is a better match for the military in his job and thus is promoted to E4 in the 10th month of his first term. During his tenure as an E4, his performance continues to be superior and after 14 months he is promoted to E5. The second recruit is a worse match for the military and takes 15 months to make E4 and another 18 months to make E5. Finally, the third recruit, representing the worst match, makes E4 in 24 months and finishes his three-year term in this rank. In this example, it is a simple matter to rank-order the three recruits. The first attains each rank sooner than the other two, and the last recruit is not only the slowest but never gets to E5.

In this example, the ordinal rank is clear but the metric—the "distance" between them—is not. We have no idea of how far apart these three are. This problem is even more pronounced when the ranking among recruits is ambiguous. Suppose that the second recruit beat the first to E4 but fell behind in getting to E5. To rank these two we would need some measure of how far apart the two were in achieving the two grades. Our solution is, implicitly, to rank the recruits by how "unusual" the promotion times were. If promotion to E4 for the first recruit placed him in the top 5 percent of all who made E4, then this percentile ranking would give us one measure of where he stood in the whole population. If his subsequent promotion time to E5 placed him in the top 10 percentile, we would have two measures of an implicit test score. For someone who was never promoted to a rank (for example, someone whose term of service expired while he was still an E4), there is an additional problem. Although we have a measurement of percentile of promotion time to E4, we have only limited information on his promotion prospects to E5. What we do know is that he was "available" for promotion for all of his months of tenure as an E4. As his E4 peers are promoted, his percentile ranking falls. The longer he remains an E4, the lower his ranking will be.

To summarize, the promotion history of recruits gives information on both the ranking and the numerical "promotion scores" of recruits. But, for two individuals of equal quality, promotion histories may

differ. Because of this randomness, the promotion scores should be viewed as important measurements of quality. The other correlates of quality which we use—education and AFQT test scores—are also indicators of quality even though they are not "quality" themselves. That is, we can incorporate these background characteristics along with promotion histories in constructing a comprehensive measure of quality. For example, someone with a high school diploma and an AFQT of 80 might be found, on average, to have a speed of promotion to E4 and E5 that is in the top 10 percentile. This information is another indicator of the recruit's quality. Thus, the background characteristics of a recruit (AFQT and education) give us an average index value based on the promotion histories of other recruits with similar characteristics. We use the background characteristics to tell us the "expected quality" of a recruit. The actual promotion history (to E4 and E5), whether slower or faster than expected, gives the other two sources of information about quality. These three measures are merged to yield a single index of performance—the quality index.

At the enlistment point all that is known about the recruit is a set of test scores and level of education. As the term of enlistment progresses, more information accumulates—to superiors as they observe his performance and to outside investigators in the form of promotion speed. By the end of the first term, information accumulated on performance may well have swamped that obtained at the enlistment point. If so, this intervening performance information constitutes a better measure of quality at the end of the first term than the entry-level test scores and education. Our work is designed to answer such questions by constructing a reenlistment point quality index and to determine the relative importance of entry-level characteristics and subsequent promotion history.[1]

After constructing these measures of quality, we impute to each member of an entry cohort a scalar measure that reflects both entry-level characteristics and promotion history. At the end of the first term we look at the distribution of quality, as measured by our index, for those who reenlist and for those who do not. We also inspect the quality of those who leave the Service before the end of the first term. Along the way we are also able to measure the importance of entry-level characteristics in affecting overall quality. We ask, in effect, whether these measures are good surrogates for a recruit's subsequent success in the military.

---

[1]The statistical model that underlies this index is described in detail in the Appendix.

II. DATA AND OVERVIEW OF THE 1974 COHORT

THE DATA

Our data consist of the 1974 Entry Cohort File.[1] The cohort file tracks all FY74 accessions over the 1974–1981 period, and includes information both from the Military Enlistment Processing Command (MEPCOM) record and from the last quarterly master or loss file of each year. From the MEPCOM record, we find information on occupation (primary MOS), AFQT percentile score, highest level of schooling completed, date of accession, entry pay grade, term of enlistment, and personal characteristics such as race, sex, and age at entry. For each year that an enlistee is in active service, we know pay grade, date of promotion to that pay grade, date of last enlistment, and Expiration of Term of Service (ETS) date from the quarterly master. If the enlistee separates during a fiscal year, the loss file provides the separation date and inter-Service separation code. We can therefore construct that recruit's entire promotion history or separation from the military up to 1981.

In the analysis we looked only at non-prior-service males who enlisted with an entry pay grade of E1.[2] Considerable effort went into cleaning the data. We deleted individuals from the sample if they had missing values for AFQT and schooling, or if they had gaps in their quarterly master files.[3] We also corrected for inconsistencies between pay grade and pay date.

A set of indicator variables was created to describe the final disposition of this entry cohort: attrition before completing term of service or separation with less than an honorable discharge (ATTRIT); separation after reaching ETS (ETSEP); and reenlistment (REUP). As noted above, we are interested in the retention of high-quality personnel. We therefore want to distinguish between those who are eligible

[1]This cohort file was extracted from personnel records maintained at the Defense Manpower Data Center (DMDC).

[2]Since information on entry pay grade was not available for the Marine Corps we set this variable equal to E1 for all Marine recruits.

[3]Gaps in the data arose whenever matches could not be made from one year to the next, usually because of missing or miscoded social security numbers used to match a recruit's different files.

to reenlist (the ETSEP and REUP groups) and those who are not (the ATTRIT group).

To do this, we used several different pieces of information. From the inter-Service separation code we distinguished three kinds of separations:

"Good" separation—expiration of term of enlistment or early release (codes 1–8) and entry into officer programs (40–49)

"Nonprejudicial" separation—medical disqualification, dependency and hardship, death, and "other separations" (codes 10–33, 90–99)

"Bad" separation—failure to meet minimum behavioral or performance criteria (codes 60–87)

A recruit with a "bad" separation code was defined as an "attrit"; all others (those with "good" or "nonprejudicial" separations) were considered to have discharged their TOE (term of enlistment) obligation and were presumably eligible to reenlist.[4]

All recruits were then filtered through several screens to determine if they separated at ETS[5] or reenlisted. We did this by looking at their ETS date in the following year. If the date changed, reenlistment was presumed. A complication that arose in using this approach was that individuals could (and did) extend for up to two years. Thus, an individual "reenlisted" if his ETS date increased by more than two years, and "extended" if it increased by less. If he extended, we continued to track him until we determined if he attrited, separated, or reenlisted.[6] For someone who extended, TOE was also extended and our analysis proceeded as if the individual had originally enlisted for the longer term.

---

[4]On the advice of the Defense Manpower Data Center (DMDC), we opted to use the inter-Service separation code to determine reenlistment eligibility rather than the more problematic reenlistment code variable.

[5]This ETSEP group includes those with "good" or "nonprejudicial" separations who therefore may not have actually reached their recorded ETS dates.

[6]We also developed several rules to deal with missing data on both inter-Service separation codes and separation dates. First, we checked to see if the individual went into the reserves. If he did, we treated him as an ETSEP and set his separation date equal to the recorded ETS date. Otherwise, we dropped him from the sample. Those who reenlisted were assigned a "good" separation code and their separation date set equal to the date of last enlistment.

PROMOTION TIMES AND JOB PERFORMANCE: THE 1974 ENTRY COHORT

Table 1 shows the eight military occupations chosen for analysis. These occupations satisfied the principal restrictions we had set for our work: They have sample sizes between 1000 and 2000 and, within Services, represent jobs that are either specialized and highly skilled on one hand or arduous occupations on the other. Sample size restrictions were dictated by the computational expense of matching individuals across yearly files and by the necessity of gaining some degree of precision in the estimated parameters.

Table 1 also shows the percent in each MOS that would be classified as "high quality" (HQ) on the basis of their entry-level characteristics—those with at least a high school diploma and an AFQT score of 50 or higher.

Table 2 shows the percentages of the cohort who attrit (ATTRIT), who separate at the end of their first term (ETSEP), and who reenlist (REUP). The final column gives the reenlistment rate, defined as the fraction of those successfully completing the first term who reenlist. These breakdowns are stratified by whether entry-level characteristics would classify them as being "high quality" (HQ) or "not high quality" (NHQ).

There are a number of stories told in this table, not all with equal clarity. First, HQ recruits are uniformly less likely to leave before

Table 1

SELECTED OCCUPATIONS FROM THE FOUR SERVICES

Service	Occupation	Job Title	N	% HQ[a]
Army	MOS11C	Indirect Fire Infantryman	1380	28.8
Army	MOS91B	Medical Specialist	1076	45.1
Navy	NECRM	Radioman	1438	50.0
Navy	NECBT	Boiler Technician	1841	39.2
Air Force	AFSC20	Communications and Intelligence	1129	80.4
Air Force	AFSC42	Aircraft Mechanic	1712	46.0
Marines	MC03	Machine Gunner and Mortarman	1544	28.0
Marines	MC60	Aircraft Mechanic	1905	47.9

[a]HQ = high-quality personnel, defined as those having at least a high school diploma, and an AFQT score of 50 and above.

Table 2

RATES OF ATTRITION, SEPARATION, AND REENLISTMENT

Service/ Occupation	Quality	Percent that ATTRIT	ETSEP	REUP	Rate of Reenlistment
Army					
MOS11C	--	30.9	49.3	19.8	28.6
	HQ	27.6	45.7	26.6	36.8
	NHQ	32.2	50.7	17.1	25.2
MOS91B	--	21.8	61.2	17.0	21.7
	HQ	14.4	70.1	15.5	18.1
	NHQ	27.9	53.8	18.3	25.4
Navy					
NECRM	--	14.7	67.9	17.3	20.3
	HQ	12.8	73.8	13.4	15.4
	NHQ	16.7	62.0	21.3	25.6
NECBT	--	32.2	55.1	12.7	18.7
	HQ	23.1	64.0	12.9	16.8
	NHQ	38.0	49.3	12.7	20.5
Air Force					
AFSC20	--	13.2	55.0	31.8	36.6
	HQ	12.9	55.5	31.6	36.3
	NHQ	14.5	52.9	32.6	38.1
AFSC42	--	22.3	45.6	32.1	41.3
	HQ	18.9	49.4	31.7	39.1
	NHQ	25.2	42.3	32.5	43.4
Marine Corps					
MC03	--	20.0	67.0	13.0	16.2
	HQ	16.9	61.8	21.3	25.6
	NHQ	21.2	69.1	9.7	12.3
MC60	--	17.2	64.0	18.8	22.7
	HQ	12.3	66.9	20.8	23.7
	NHQ	21.8	61.3	16.9	21.6

NOTE: HQ = high-quality recruit; NHQ = not high quality recruit; ATTRIT = attrit; ETSEP = separate at ETS; REUP = reenlist. Reenlistment rate is calculated as REUP/(REUP+ETSEP).

ETS. Second, in three of the four "high-skilled" occupations the HQ recruits reenlist at a lower rate than their NHQ counterparts.[7] In the fourth, MC60, the difference is minor. Third, in the two combat arms MOSs (MOS11C and MC60) HQ recruits reenlist at a higher rate. Fourth, compared with the other Services, Navy HQ recruits reenlist at lower rates, whereas the Air Force has uniformly high reenlistment rates.

Table 3 tells the same story only in reverse: It shows the entry-level characteristics of recruits stratified by the disposition of their first term. Here we see that most of the distinction between HQ and NHQ is due to education. That is, AFQT differences are very small among the three groups and most of the quality difference is due to the distinction between high school and non-high school graduates.

How important are the differences that do exist? Is the loss of high-quality recruits among Navy radiomen (NECRM) (the largest loss rate of any of the eight occupations) a serious problem or is the distinction between HQ and NHQ unimportant for this occupation? To answer these questions we need to know how important are differences in entry-point characteristics for job performance. Figures 1–8 are a first approach to addressing these questions and introduce the notion of "unobserved" performance as it relates to promotion data.

Figure 1 contains four graphs, one for each of four classifications of individuals in Army MOS11C. The topmost graph, HQ:F, depicts the promotion rate to E5 for recruits who were "high quality" upon entry (high school diploma or more and AFQT greater than the 50th percentile) and who attained grade E4 *before* the median time to that grade. Along the horizontal axis is months in grade E4; the vertical axis measures the cumulative promotion rate to E5. Thus, for those who were HQ upon entry and who were promoted quickly to E4 (F), approximately 42 percent were promoted to E5 before the 20th month in E4 (the height of the HQ:F graph at 20 months). The three other graphs on Fig. 1 show the promotion rates for the other complementary groups: NHQ:F denotes those who were not in the HQ group but who made E4 before the median time; HQ:S those who were in the HQ group but who made E4 slower (S) than the median; and NHQ:S those who were not in the HQ group *and* who were promoted later than the median time to E4.

The ordering of the graphs is as might be expected: Those promoted quickly to E4 are also promoted quickly to E5 and those classified as HQ are promoted quickly to E5. Within the HQ group those promoted

[7]The high-skill occupations are Army M0S91B (Medical Specialist), NECRM (Radioman), AFSC20 (Communications and Intelligence), and Marine Corps MC60 (Aircraft Mechanic).

Table 3

CHARACTERISTICS OF PERSONNEL BY OUTCOME

Service/ Occupation	Attrit	Separate	Reenlist
A: Average AFQT Scores			
Army			
MOS11C	49.5	49.2	53.4
MOS91B	54.7	57.6	54.4
Navy			
NECRM	55.7	56.4	51.9
NECBT	56.1	57.5	55.3
Air Force			
AFSC20	68.4	69.5	69.1
AFSC42	50.6	52.7	51.5
Marine Corps			
MC03	57.3	50.9	57.7
MC60	68.5	66.8	66.3
B: Percent with High School Diploma			
Army			
MOS11C	44.1	50.9	56.9
MOS91B	44.3	72.5	67.8
Navy			
NECRM	65.6	84.2	77.5
NECBT	41.9	61.3	60.4
Air Force			
AFSC20	96.6	98.9	98.3
AFSC42	92.7	94.0	94.4
Marine Corps			
MC03	35.9	52.6	65.5
MC60	40.9	57.7	58.9

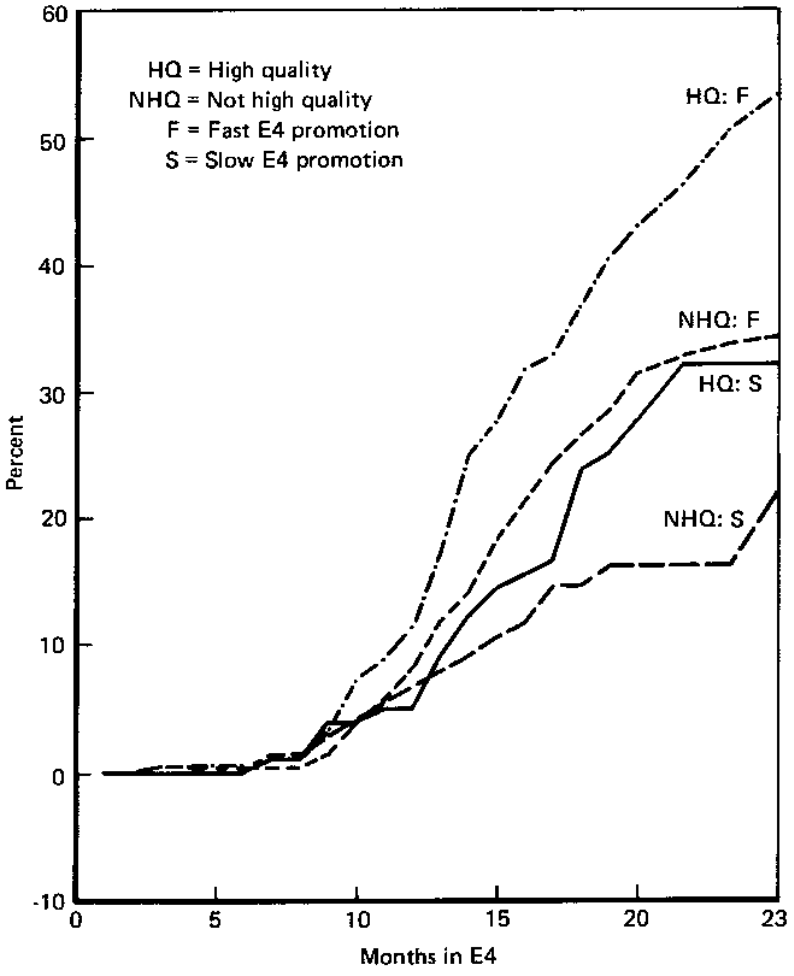


Fig. 1—Percent promoted to E5: Army MOS11C

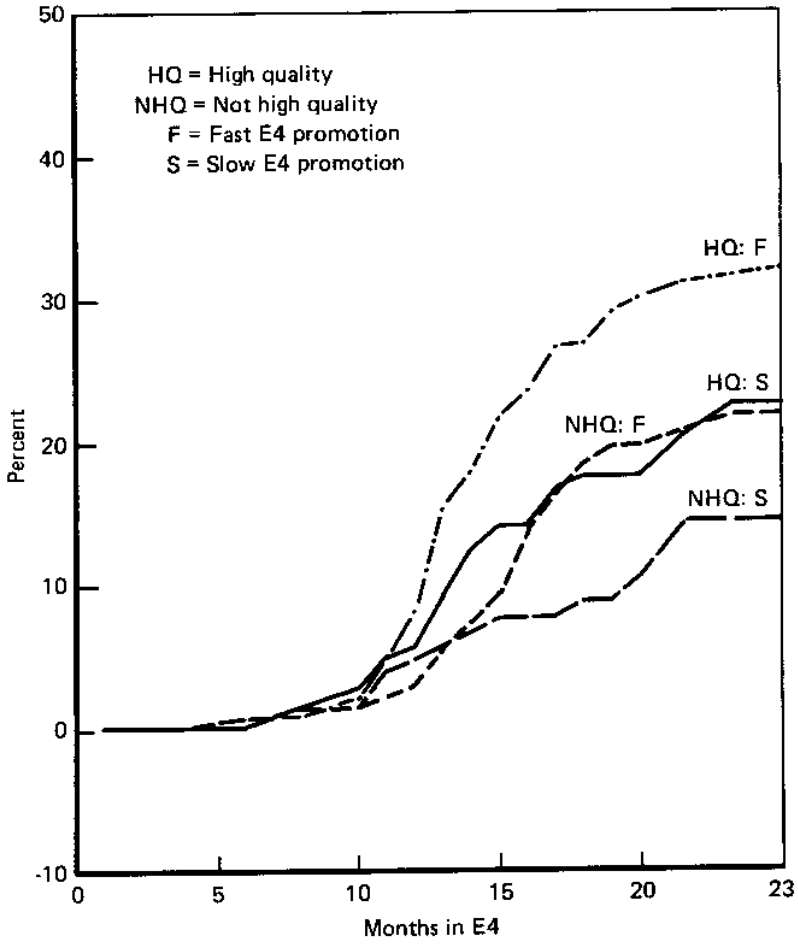


Fig. 2—Percent promoted to E5: Army MOS91B

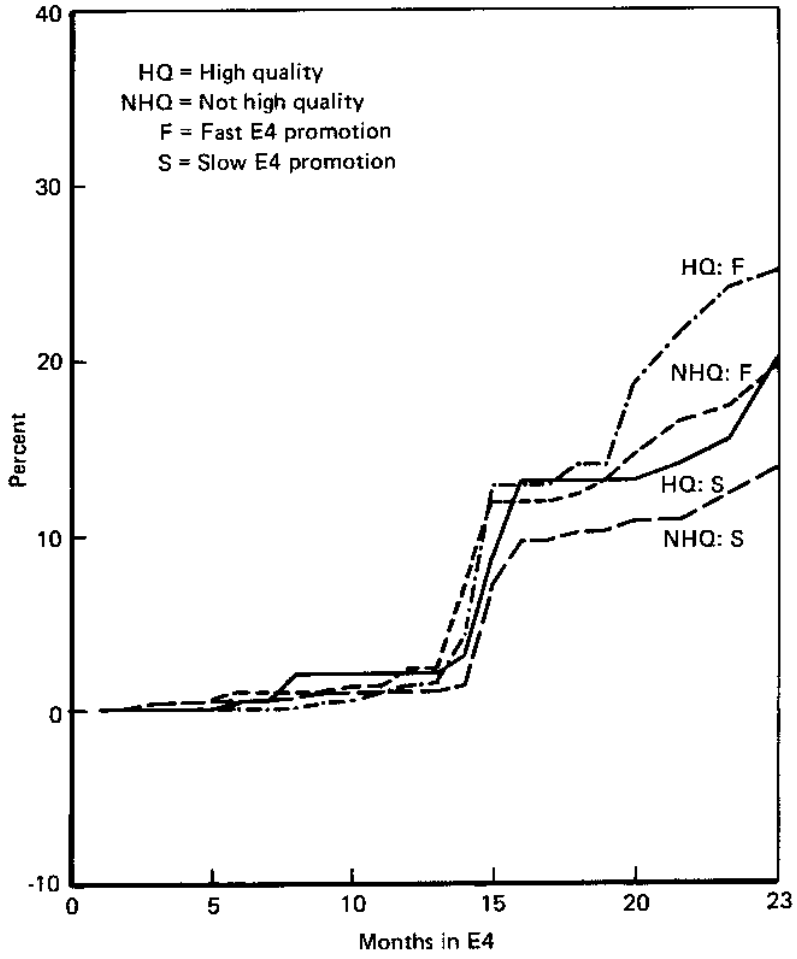


Fig. 3—Percent promoted to E5: Navy NECBT

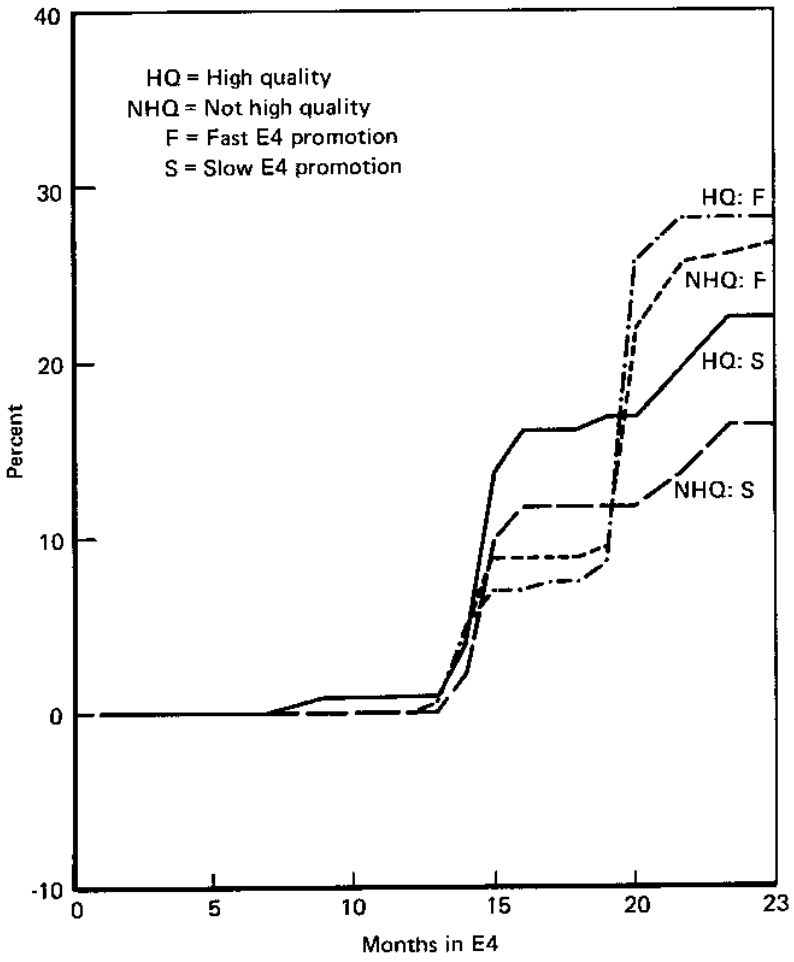


Fig. 4—Percent promoted to E5: Navy NECRM

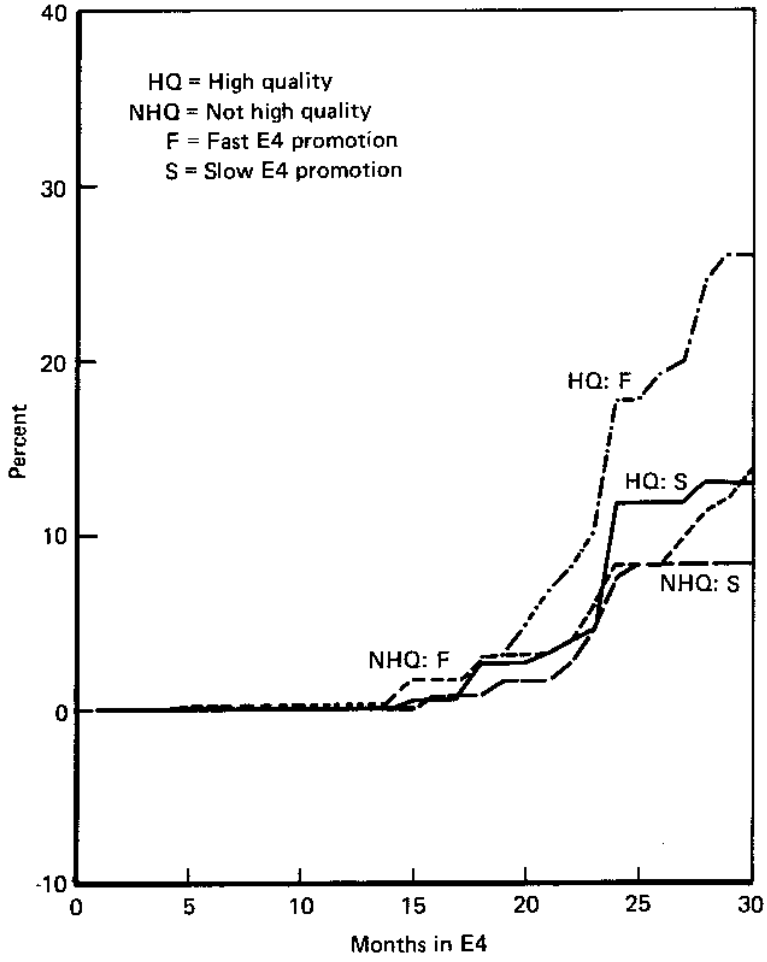


Fig. 5—Percent promoted to E5: Air Force AFSC42

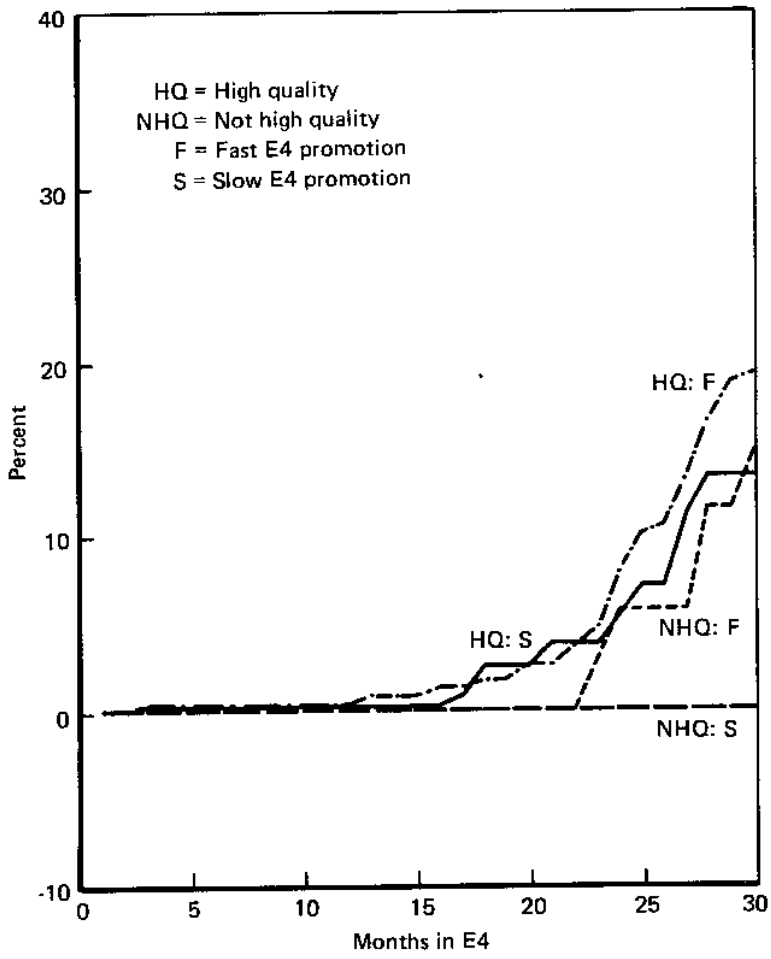


Fig. 6—Percent promoted to E5: Air Force AFSC20

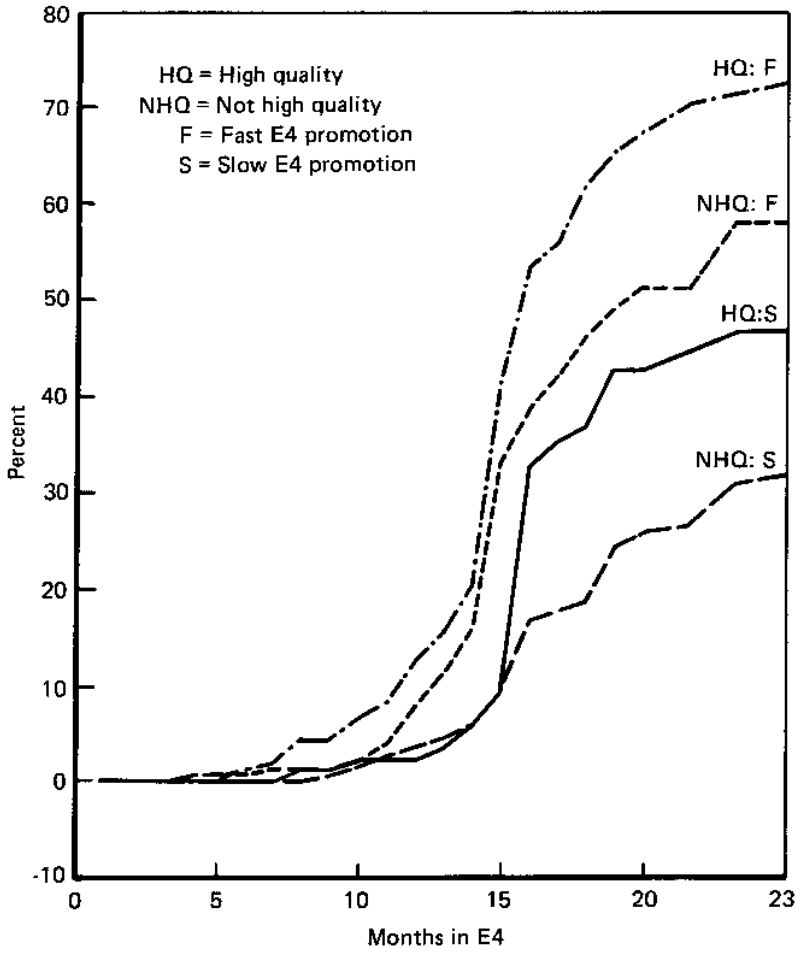


Fig. 7—Percent promoted to E5: Marine Corps MC03

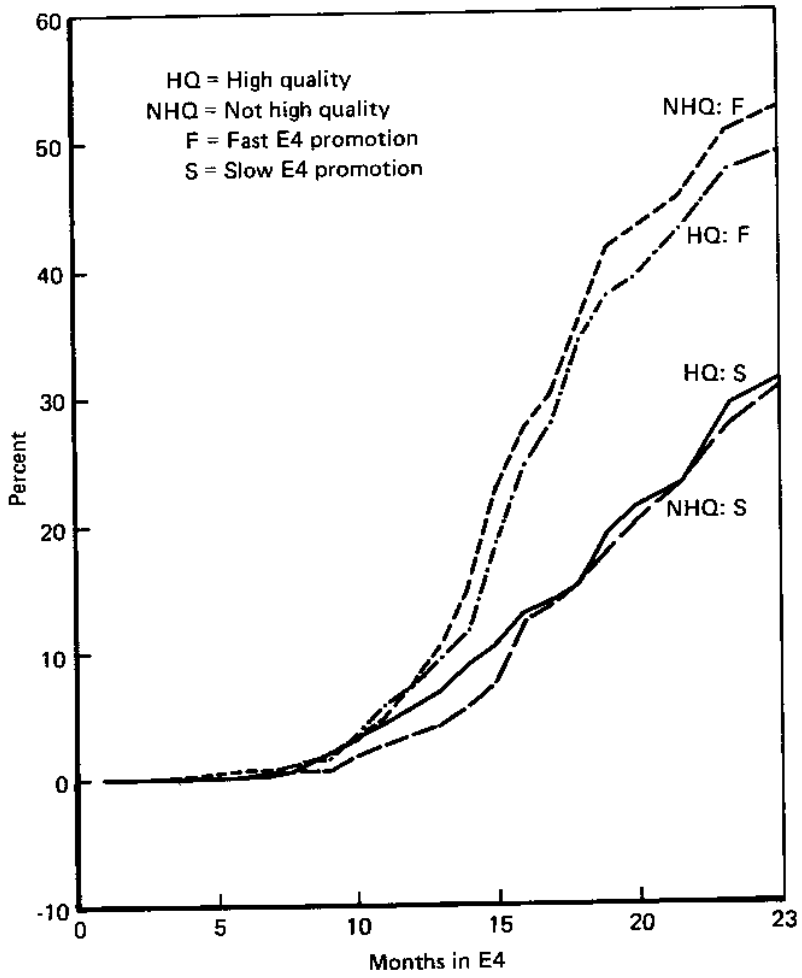


Fig. 8—Percent promoted to E5: Marine Corps MC60

fast to E4 get to E5 even faster. Likewise, those not in the HQ group but promoted fast to E4 get to E5 quicker than their "slow" counterparts. The middle two graphs show the trade-off between entry-level characteristics (HQ or NHQ) and performance as measured by promotion speed. In this occupation the promotion rate to E5 is almost the same for the NHQ:F as for the HQ:S.

In the last period for which we measure promotion rates (the 23rd month), the HQ groups are separated by about 21 percentage points in cumulative promotion rates (53 for the "fast" group and 32 for the "slow" group). Within the "fast" group, those who are HQ have cumulative promotion rates about 19 percentage points higher (53 versus 34) than those who are classified as being NHQ (similar comparisons apply to the "slow" group). We interpret this as showing the importance of performance measures that are unrelated to the HQ/NHQ distinction. If, for the moment, we take promotion to E5 as representing successful performance in the first term, then the distinction between being fast and slow to gain E4 is as important for success as the distinction between HQ and NHQ. Even among the HQ, the large difference between fast and slow shows the importance of performance factors completely unrelated to entry-level background characteristics.

We call these factors unobserved ability. It is unobserved to us, as investigators, although it may be perfectly obvious to the individual's supervisors. Individuals who are particularly energetic or competent, or both, will be recognized as such and promoted regardless of their formal education or cognitive test scores. The variation in promotion times, after controlling for entry-level characteristics, gives us some sense of the importance of this unobserved component of performance.[8]

Figures 2 through 8 repeat these promotion rate graphs for the other seven occupations that we analyzed. Figure 2 shows the results for Army MOS91B, medical corpsmen. The ordering of the graphs is the same as for Army MOS11C except that the middle two—HQ:S and NHQ:F—are reversed though very close to one another. Note that the overall rate of promotion for all groups in this MOS is lower than those in MOS11C.

Figure 3 shows results for Navy boiler technicians (NECBT). Although the terminal promotion rates at the 23rd month are ordered like those of the Army, the entire set of graphs is very congested—

[8]These graphs tell only part of the story. To be classified as "fast" or "slow" the individual must first have been promoted to E4, something that does not always occur in the first term. Thus, these graphs deal only with a subset of the entry cohort—the more successful subset—and cannot be interpreted as generalizations to be applied to all. The model we develop below tries to solve this problem as well as other statistical and data difficulties.

there is little to distinguish the four groups until after the 20th month. The jump in promotion that occurs between the 13th and 15th months occurs for all groups and at about the same rate. Figure 4, for Navy radiomen, shows a peculiar crossing-over of promotion rates between the 14th and 20th month. Those who were promoted fast to E4 are actually delayed slightly in their promotion to E5 compared with those promoted slowly to E4. Even so, their ultimate ranking at 23 months is like the other occupations.[9]

The two Air Force occupations—aircraft mechanic (AFSC42) and communication and intelligence (AFSC20)—are shown in Figs. 5 and 6. Both show promotion patterns resembling those in the other Services. However, the average time of promotion to E5 in the two Air Force occupations is much longer and the scale on the horizontal axis has been extended to 30 months in order to show differences between the four groups.

For Marine Corps occupations, promotion rates for machine gunners and mortarmen (MC03) and for aircraft mechanics (MC60) are shown in Figs. 7 and 8, respectively. For MC03 the graphs look very similar to the other combat arms occupation we analyzed, Army MOS11C. Promotion rates are very high and differentiated both between fast and slow, and between HQ and NHQ. For MC60 the graphs show an unusual insensitivity to the HQ/NHQ distinction, once speed of promotion to E4 has been (roughly) controlled for. In other words, promotion to E5 in this occupation is almost wholly explained by speed of promotion to E4.

While not a substitute for formal analysis, these figures show the central analytic features of the statistical model that we employ. First, there is a strong relationship between speed of promotion to E4 and speed of promotion to E5. This correlation suggests that there is a component of job performance that is specific to the individual throughout the first term, a fact that will be central to the formulation of the model. If speed of promotion to E4 were unrelated to speed of promotion to E5 because, for example, selection for the two grades was based on entirely different sets of characteristics, then our attempt to estimate an individual-specific performance measure would be frustrated. Second, these graphs show that this common component of job performance is due to matters other than AFQT and education. Even

[9]We do not attach any significance to this apparent anomaly although there are at least three possible explanations. First, the best radiomen may be sent back to school for further training. They may not be eligible for promotion while in school, but are elevated quickly upon completion. Second, time-in-service (as opposed to time-in-grade) requirements may also be responsible for delays in promotion to E5 for those who moved quickly to E4. Finally, the differences may simply be due to sampling variation.

within those classified as HQ or NHQ, those promoted fast to E4 are also promoted quickly to E5.[10]

Finally, the graphs show that the HQ/NHQ distinction *does* have an influence on promotion rates. This demonstrates another not-so-surprising result: Individuals with greater entry-level "quality" perform better in the military. The interesting question, not answered in these graphs, is the relative importance of "unobserved" ability versus the observable entry-level characteristics.

---

[10]It might be argued that within the HQ group there is sufficient variation in AFQT to explain the association of promotion times. That is, people with AFQTs of 75 and above perform better than those with AFQTs between 50 and 75. This would give rise to an association of promotion times within the HQ class that is wholly explained by background characteristics. We are able to refute this hypothesis in the statistical model developed below.

III. THE MODEL

GENERAL CHARACTERISTICS

Figure 9 is a schematic of our model. We are attempting to identify the latent unobserved variable we call "quality." Consistent with our preceding discussion, we see this person-specific variable as composed of cognitive ability (AFQT), education, and a catch-all collection of other factors called "unobserved ability." Performance is viewed as affecting the outcome variable, speed of promotion. We concentrate here on the timing of promotion, but this same model could as well be applied to other measures of performance. For example, supervisor ratings or skill test scores could be combined with promotion speed to construct an even more comprehensive index of performance. We do not have these measures in the data set we analyze although, in principle, there is no difficulty in incorporating this information.

What allows us to identify the contribution of unobserved ability to performance is the common effect of this latent variable on promotion speed. As shown in Fig. 9, someone endowed with a high level of unobserved ability would be distinguished by fast promotion times to both E4 and E5 even after controlling for AFQT and education. The econometric specification corresponding to this schematic is given by the three equations below:

$$Q = \beta_0 + \beta_1 AFQT + \beta_2 LHS + \beta_3 GHS + Q \tag{1}$$

$$Prob[E4 \mid time\ in\ E3] = f_4\,(Q + U_1, time\ in\ E3) \tag{2}$$

$$Prob[E5 \mid time\ in\ E4] = f_5\,(Q + U_2, time\ in\ E4) \tag{3}$$

Equation (1) relates quality (linearly) to AFQT and to two variables indicating if the individual does not have a high school diploma (LHS) or if he has at least some college education (GHS). If we could observe the quality variable Q, Eq. (1) is the regression we would be estimating. The errors in that regression are denoted Q, representing the estimated value of the unobserved component of performance (i.e., unobserved ability).

Although we do not observe Q directly, we do, given the structure of the model, observe it indirectly through promotion speed. In our

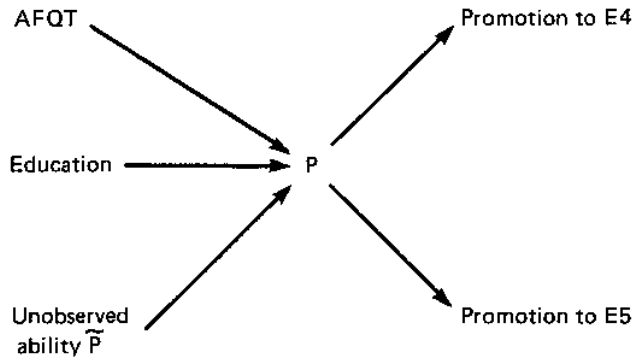
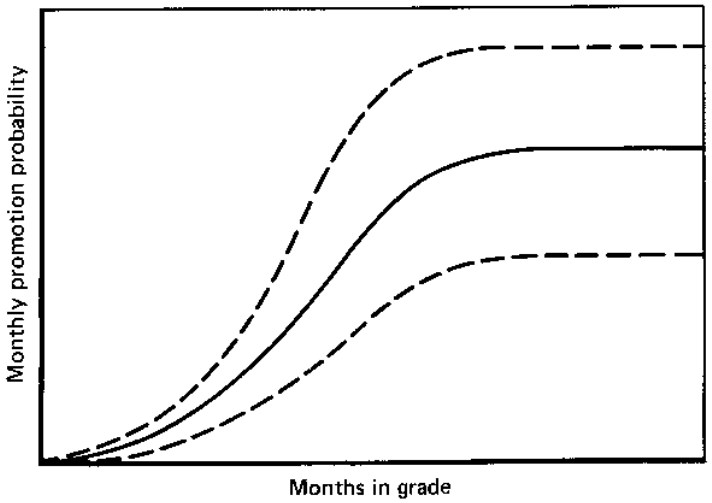


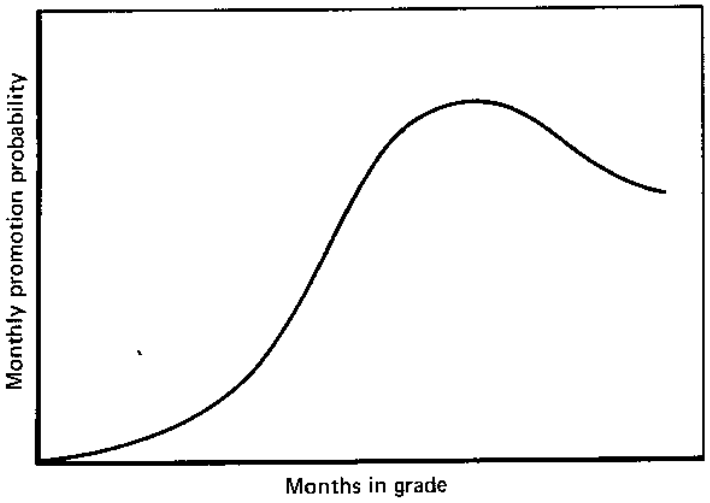
Fig. 9—Quality viewed as a latent variable

model, we do not use promotion speed directly because it is observed only for the subsample of those who are promoted. Rather, we estimate a model that describes the probability of being promoted in any month of "exposure" to that grade. For example, given data on Q we would estimate a model describing the probability of being promoted to E4 in the 15th month of service as a function of Q. The data would be composed of those who could have been promoted in that month, i.e., those still in the Service and not yet promoted to E4. Similarly, we could estimate a model describing promotion in the 16th month using the sample in the 15-month model, less those who were promoted in the 15th month as well as those who left the Service at the end of the 15th month. The same model could be repeated for any month of exposure for promotion to E4, from the first month to the last in the data. Similar models could also be estimated for E5. Equations (2) and (3) describe such a sequence of models. What distinguishes these equations from a completely general sequence of models is that we constrain the effect of Q on promotion probabilities to be proportional from one month to the next. Thus, rather than estimating an equation of promotion for each month, we estimate only one equation and parameterize the effect of time in grade on the promotion probability.

Figure 10(a) shows the way in which Q is constrained to affect promotion probabilities, as well as the way in which promotion probabilities change with months in grade. The middle (solid) line shows the course of monthly promotion probabilities for a "typical" individual. At first, this probability is close to zero, reflecting the Services' policies



(a) Individual promotion probabilities



(b) Average promotion probabilities for survivors

Fig. 10—Effects of heterogeneous promotion probabilities

of time-in-grade requirements.[1] The curve rises rapidly around the normal promotion time and then stays fixed at some level (more on this below). For someone with a higher Q, the curve describing monthly promotion probabilities is constrained to rise proportionally, i.e., by an equal percentage change in each month. Thus the curve rises by (say) 10 percent everywhere. Because the monthly promotion probability is close to zero in the very early months of tenure in grade, a 10 percent change is very small. Beyond some limit, when the probability of promotion has hit a peak, the 10 percent change is arithmetically much larger. The top (dashed) graph depicts such a case, whereas the lower dashed line shows the opposite, the promotion probabilities for someone with a lower level of Q. Our model estimates both the effect of Q on promotion probabilities as well as the shape of the typical time path of promotion probabilities.

Although we have constrained the model of monthly promotion probabilities (for a given Q) to have the shape shown in the top panel, the empirical shape is quite different. The bottom panel (Fig. 10(b)) describes the typical empirical shape for these probabilities. It would be calculated, as described earlier, by dividing the number of promotions that occur in a given month of a grade by the total number of people who could have been promoted in that month. As the months progress, this calculated promotion probability rises to a peak and then declines. However, since this probability is calculated using data on those who have not been promoted, the sample becomes more heavily weighted with those who have very low prospects of ever being promoted. These individuals will draw down the average promotion probability since those with high promotion probabilities will have already been promoted out of the sample. Heterogeneous probabilities will give rise to such curved shapes even though they may not represent the path of promotion probabilities for any given individual.[2]

---

[1]Although all the Services have time-in-grade as well as time-in-service requirements for promotion, these requirements can be waived. Thus, exceptional cases can merit promotion faster than the established minimum time limitations. On the other hand, having passed these minima is no guarantee of promotion since they are but some of many requirements that must be met.

[2]In estimation, we use the shape given in the top panel—one that reaches a maximum asymptote without turning down. We do this based on *a priori* information about the promotion policy of the Services as well as on some preliminary data analysis. In attempting to find functional forms for the shape of monthly promotion probabilities, we found that this simple characterization combined with estimates of the heterogeneity of promotion probabilities fit the data much better than more complicated functional forms that ignore heterogeneity. Although we would have preferred to have carried out a more formal set of tests to determine the appropriateness of this form, we are satisfied that our choice of functional form is reasonably close to the actual one for Q-constant individuals.

The estimation technique follows from Eqs. (1) through (3) above, with the technical details of the model given in the Appendix. Since we do not observe Q, and hence cannot estimate Eq. (1) directly, we can substitute Eq. (1) into Eq. (2) giving a model relating promotion probabilities to AFQT and education. The model is structured to have one Q-factor common to both promotion functions, so that the effect of the background variables should be proportional to one another in each promotion equation. In other words, if a 10-point decline in AFQT lowers promotion probabilities to E4 by as much as the lack of a high school diploma, this proportional relationship is constrained to hold for promotion to E5. We are, in effect, estimating an index of background characteristics that has the same relative weight attached to each of the entry-level characteristics. The effect of this common index on promotion is, however, allowed to be different for each promotion grade.

Having estimated the effect of this index on promotion probabilities in the two grades, we can next net-out the influence of these background characteristics on promotion. We effectively "handicap" each individual according to the value of his index. Someone with a high school diploma and a high AFQT would be expected to be promoted faster than average based on the estimated effect of these characteristics in the promotion equation. If this person was promoted at a slower pace than would be expected (though perhaps still faster than average overall), we would estimate that the unobserved component of his Q was low. Similarly, someone without a high school diploma and with low AFQT might be estimated to have a low promotion rate. If an individual exceeded this prediction, he would be estimated to have a high level of unobserved Q. For each individual, then, we can calculate a Q-index at the end of the first term or when E5 is attained, whichever comes first.[3]

SOME CAVEATS

Before discussing the empirical results (Sec. IV), the restrictions and limitations implicit in the model should be noted. First, the quality index that is estimated pertains to an ability and performance factor that is *common* to activities as both an E3 (and under) and an E4. That is, *relative quality* is constrained to be constant over time. If there are dimensions of quality specific to E3 activities and a different

[3]We do not account for promotions to E6, which are very rare in any case. Thus, for the purpose of estimating the model, information on these individuals stops once they make E5.

set specific to E4 activities, our technique will not estimate these different kinds of quality. Rather, the methods we employ estimate only the *common* set of quality factors.

Second, the specification we use relating quality to entry-level characteristics (Eq. (1)) is linear in AFQT. This means that two people who are 10 points apart in AFQT are quality differentiated to the same extent no matter what the absolute level of AFQT that separates them, i.e., 30 versus 40 is the same as 80 versus 90. It might be argued, though, that within some *lower* range higher AFQT is very important for quality, whereas in a higher range AFQT matters very little. The average person with an AFQT of 25 might have very low performance compared with the average person of 50 AFQT. But that difference at 25 points might not significantly separate the performance of persons of 75 versus 50 AFQT. If such nonlinear relationships exist, our linear assumption will lead to an estimate of the effect of AFQT that is an *average* of the effects across the whole range. In the examples discussed, we would underestimate the effect of AFQT in the lower AFQT range and overestimate its effect in the upper range.

Third, the model represented by Eqs. (1) through (3) and Fig. 9 assumes that the correlation between speed of promotion to E4 and to E5 is due to the common quality component in these two performance measures. If, for example, there is a *direct* effect of early promotion on subsequent promotion rate *apart from ability*, then we will tend to overestimate the relative importance of *unobserved* quality differences as compared to entry-level characteristics (*observed* quality differences) in the construction of the quality index. Our estimation technique assumes independent evaluations for promotion, and any dependency in these evaluations will mimic common ability factors.

Finally, since the quality index we construct is based solely on first-term performance and entry-level characteristics, we cannot directly evaluate its usefulness as a predictive tool for second-term performance. For example, responsibilities of leadership, organization, and command that would be required of an E6 might not be closely related to the set of factors that make a good E3 or E4. Further research is needed to verify the predictive content of the index for second-term performance.

IV. EMPIRICAL RESULTS

ESTIMATES OF THE QUALITY INDEX

Figures 11 and 12 introduce the simulation results based on our estimates of the model. They show the way in which the quality index combines information about AFQT, education, and promotion speeds to arrive at a single scalar measure of performance. The horizontal axis in each of these graphs measures total time in service in the first term. The vertical axis, representing the Q-index, is centered on zero, the normalized average performance of the entire entry cohort for this particular occupation. The scale of the vertical axis is therefore in "standardized" units, so that a value of 1.0 indicates that performance is one standard deviation higher than the mean, corresponding to the top 16th percentile.

The middle graph in Fig. 11 corresponds to an enlistee with mean AFQT scores and a high school diploma who is promoted to E4 at 14 months and to E5 at 34 months. The top graph in Fig. 11 corresponds to the estimated index for someone promoted at the same times but with 20 more points of AFQT; similarly, the lower graph corresponds to the Q-index of someone with like promotion times but without a high school diploma. The distances between these graphs are derived from parameter estimates for Army MOS11C.

First, consider the middle graph. At entry, the model would predict that this "typical" recruit has a performance index of zero, the average, since information about his subsequent promotions is not as yet available. As months in service increase, the index at first remains constant because no new information is accruing about the relative performance of this individual. Between the 10th and 14th months, the index begins to bend down because some members of the entry cohort are being promoted to E4 whereas the hypothetical person we are following has not yet been promoted. Because the average performance index is zero for the cohort and because those who were promoted will be estimated to have higher than average performance, it must be true that those left behind are of below average performance—at least based on the limited information available up to the 14th month.

Subsequently, this person is promoted to E4 at the 14th month and the performance index jumps to about 0.6. For the reasons discussed above, there is at first little change in the Q-index and only as time in

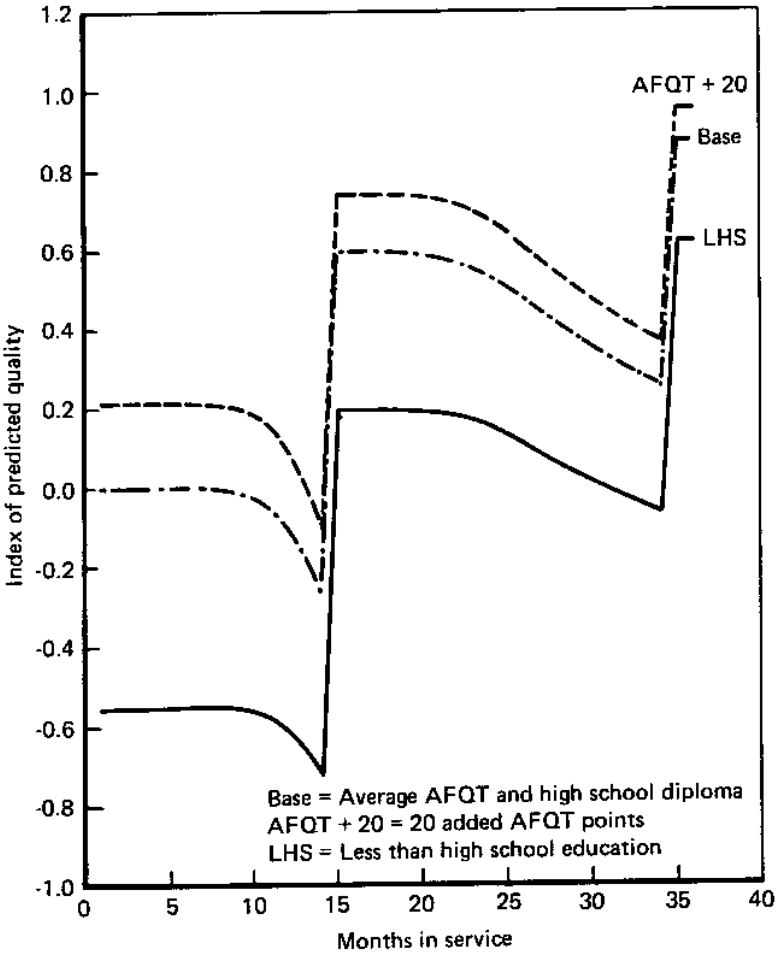


Fig. 11—Index of quality: effects of AFQT and schooling

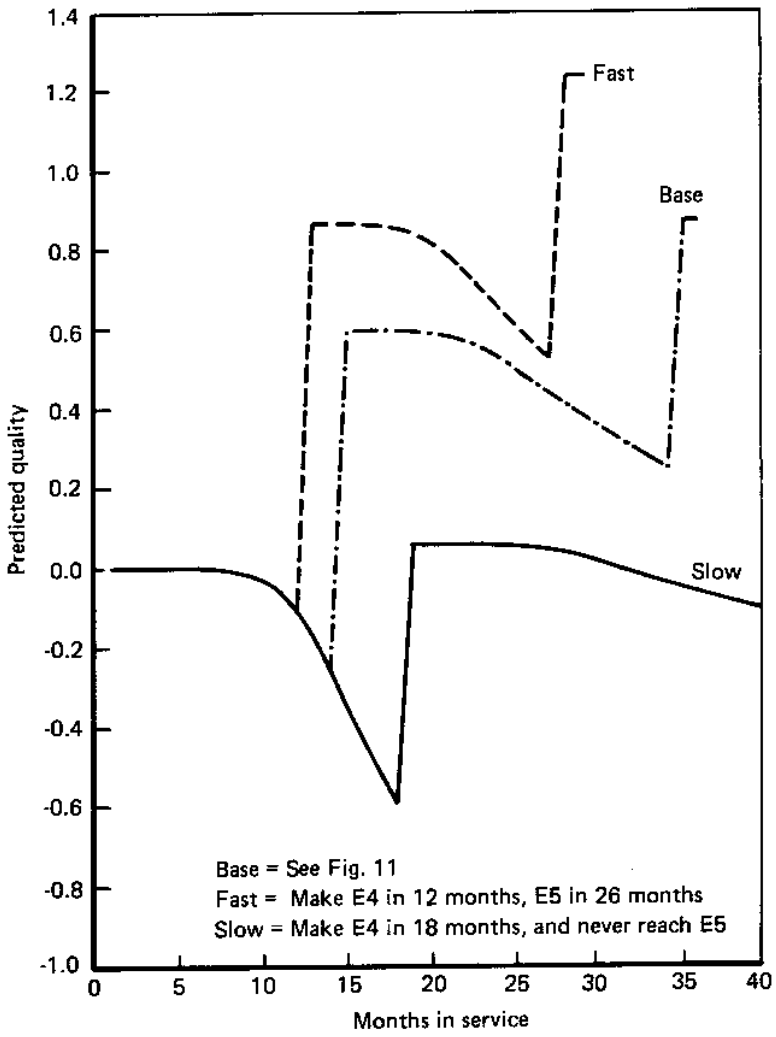


Fig. 12—Index of quality: effects of promotion times

E4 increases toward the 30th month does the index again fall slightly. This reflects the promotion of some E4s to E5 during the second year in the first term. The fact that some are being promoted, while our "typical" person is not, leads to a slight revision downward in the promotion index. Finally, at the 34th month he is promoted to E5 and the index jumps again, this time to around 0.9. The time path of this index is our estimate of the average performance index of individuals, through their first terms, with mean AFQT scores and high school diplomas, who were all promoted at the times indicated. Since there is no other information available about individuals other than their entry-level attributes and their promotion histories, the *terminal* value of the Q-index by the time ETS is reached represents our best guess of the job performance of any individual with those characteristics.

Figure 12 shows the time path of the performance index for three recruits with the same AFQT and education but with *different* promotion times. The top two graphs represent "fast" and "baseline" promotions to E4 and E5, whereas the bottom graph is for someone who remains an E4 at ETS (the "slow" case). Jumps in the performance index correspond with promotions to E4 and to E5. The size of the jump in the performance index is related to the speed of promotion. Because a promotion to E4 after only 12 months in service is very unusual, the index rises most for the "fast" case shown in the top graph, and rises least for the "slow" case in the bottom graph. After promotion to E4 and before promotion to E5 there is a convergence of the indexes. If the promotions to E5 had not occurred for the two top cases, their graphs would eventually converge toward that of the slow case. This convergence is caused by dilution of the effect of the promotion to E4 as time goes on and as these individuals look more and more similar.

Table 4 gives numerical estimates of the effects of background characteristics on the quality index. In the earlier graphs, these coefficients were used as weights in combining AFQT scores and education variables to form our background index. In this table, they show the effects of these variables on the performance index, measured in standardized units. Thus, for Army MOS11C, a 10-point change in AFQT raises the performance index by 0.0977 standardized units. Recall that a 20-point change in AFQT raised the Q-index by about 0.2 in the top graph of Fig. 11. For this same MOS, the absence of a high school diploma lowers the index by 0.519, the drop in the level of the lower graph. Finally, education in excess of a high school diploma raises the index by 0.115, although this is statistically insignificant.

This table also shows that, although usually statistically significant, AFQT has a very small quantitative effect on the index when compared

Table 4

THE EFFECTS OF ENTRY CHARACTERISTICS ON THE PERFORMANCE INDEX

Occupation/ Service	AFQT	Less than High School	More than High School
Army			
MOS11C	0.00977[a]	-0.519[a]	0.155
MOS91B	0.01084[a]	-0.698[a]	0.323[a]
Navy			
NECRM	0.00249[a]	-0.084[a]	0.168
NECBT	0.00505[a]	-0.444[a]	0.290
Air Force			
AFSC20	0.02830[a]	0.452	0.244
AFSC42	0.01572	-0.793[a]	2.451[a]
Marine Corps			
MC03	0.01280[a]	-0.569[a]	0.285[a]
MC60	0.00286	-0.595[a]	0.010

[a]Statistically significant at the 5 percent level.

to the presence or absence of a high school diploma. In Army MOS11C, an enlistee without a high school diploma would need to have about 53 more AFQT points to achieve an index equal to that of a high school graduate. This trade-off ranges from a low of 33 for Navy radiomen (NECRM) to a high of 208 for Marine Corps aircraft mechanics (MC60).[1]

Table 5 presents an analysis of variance of the model. In the column under "Total" is the pseudo-R^2 associated with the model. If performance were observable this measure would show the ability of the background index (AFQT and education) and the promotion time rankings to explain the unobserved latent variable, quality, and its interpretation would be exactly like that of a traditional R^2 concept.

The basis of this statistic can be thought of as the correlation of promotion rankings in the two grades (see details in the Appendix). If these two rankings were perfectly correlated, someone promoted in the

[1]This calculation ignores the anomalous positive (but insignificant) coefficient for the "less than high school" variable in AFSC20.

Table 5

ANALYSIS OF VARIATION IN THE PERFORMANCE INDEX

Service/ Occupation	Percent Explained			Fraction Explained by Unobserved Ability
	Total	by AFQT and Education	by Unobserved Ability	
Army				
MOS11C	52	13	39	0.75
MOS91B	65	22	43	0.66
Navy				
NECRM	87	0	87	1.00
NECBT	51	5	46	0.90
Air Force				
AFSC20	37	34	3	0.08
AFSC42	22	18	4	0.18
Marine Corps				
MC03	56	15	41	0.73
MC60	71	9	62	0.87

top 5 percentile to E4 would also be promoted in the top 5 percentile to E5, and the ordering of promotions would be preserved from grade to grade. If they were uncorrelated, someone promoted in the top 5 percentile to E4 might as likely be promoted in the bottom 5 percentile to E5. In fact, as was shown in the earlier figures, people promoted quickly to one grade are also promoted quickly to the next, although with something less than a perfect association.

Given some correlation between promotion rankings, we can then ask to what extent this is due to the fact that education and AFQT are common elements for each promotion decision. That is, are promotion rankings correlated because the same person has the same entry-level characteristics throughout his first term and thus is promoted either quickly or slowly to each grade because of the influence of these permanent characteristics? Alternatively, if background characteristics hardly affect promotion prospects, are the promotion rankings correlated because of other "unobserved ability" attributes? That is, even after "handicapping" individuals, as described above, is there still a strong association between promotion rankings? The (estimated) answer to these questions is given in the remaining columns of Table 5.

For Army MOS11C, the percent of total variation explained by the model is 52, indicating, as was apparent in Fig. 1, that there is a fair amount of correlation between promotion rankings. The second column for this occupation shows that only 13 of the total 52 can be attributed to background characteristics, whereas 39 (the difference) is attributable to unobserved ability. The last column expresses this decomposition as a fraction, showing the proportion of explained variance that is attributable to unobserved ability. For this occupation, 75 percent of the performance index is attributable to matters other than entry-level characteristics.

Except for the Air Force, the other Services and occupations show similar breakdowns. In each case, entry-level characteristics are a relatively minor component of the index. The two extremes among the occupations are Navy radiomen (NECRM) and Air Force aircraft mechanics (AFSC42). In NECRM all of the variation in the index (within rounding) is explained by components other than AFQT and education. Table 4 also shows this in that the effect of entry-level characteristics on the index is very small. In addition, 87 percent of the variation in the index is explained by the model. These two statistics taken together describe a job in which it is fairly easy to distinguish someone with the ability to perform his job (so as to explain the high correlation) and in which that ability is not related to either education or to general cognitive ability (as measured by the AFQT).

The opposite extreme is apparent for Air Force aircraft mechanics. In this case, the low (22) value for the total percent explained shows fairly low association of promotion rankings, but the association that does exist is almost entirely explained by AFQT and education. The same description would be true of the other Air Force occupation analyzed, communications and intelligence (AFSC20). Since the Air Force uses, as part of its promotion process, tests that measure general cognitive ability, it is perhaps not surprising that these characteristics play a more important role here than in the other Services. Table 4 confirms this. The coefficient of AFQT is higher in the Air Force than in any of the other Services.[2]

The conclusion to be drawn from this decomposition of the quality index is that, except for the Air Force, entry-level characteristics are relatively unimportant as a predictor of first-term quality.[3] By the time

[2]Even though the coefficient of AFQT for AFSC42 is insignificantly different from zero at the 5 percent level, its arithmetic magnitude is large compared to the other Services.

[3]This is not to say, for example, that education is a poor predictor of attrition. It is, in fact, a good predictor. What we mean is that education has relatively little weight in predicting the ranking of individuals as measured by their promotions.

ETS is reached, the importance of these background characteristics is greatly reduced. This suggests that rank achieved at the end of the first term, and perhaps the speed with which that rank is achieved, should receive much larger weight in reenlistment decisions than entry-level characteristics.

QUALITY AND REENLISTMENT

Based on each recruit's promotion record during the first term, his education, and AFQT, we have estimated his performance index value. In this subsection, we investigate the relationship between these values and the reenlistment decision, using the subsample of those who successfully completed their term of service.

Table 6 shows the results of a logistic analysis relating the value of the performance index to the probability of reenlisting. The performance index is measured in standard deviation units so as to facilitate cross-occupation comparisons.[4] The estimated effect of 12.32 for Army MOS11C means that someone with a quality index one standard deviation higher than the mean has a probability of reenlisting 12.32 percentage points higher than the mean. The mean reenlistment rate for this occupation is 28.6 (see Table 2) so that this change represents an increase of over 50 percent. For the other occupations, the effects are smaller, although all except AFSC20 are highly significant.

In Table 7 we decompose the effect of the quality index on reenlistment into components attributable to AFQT, education, and the unobserved ability component of the performance index. These results are based, again, on a logistic model, but this time include both the quality index and other background characteristics as explanatory variables. The first column, under "Q-index," shows the effect of changes in this index controlling for AFQT and education. Thus, it shows the effect of the changes in the *unobserved* component of this index. In all cases, the effect of this component is larger than the effect of the index as a whole (presented in Table 6).

The other variables in the model, AFQT and the two education classifications, show the effects of these variables *controlling* for the Q-index. Because the Q-index incorporates these background variables, the coefficients in Table 7 must be read as partial effects. For example, someone with 10 more AFQT points will necessarily have a higher Q-index. In order for the Q-index to be held constant, the unobserved

[4]The logistic coefficients have been transformed into derivatives of the reenlistment probability by evaluating them at the mean reenlistment probability.

Table 6

EFFECT OF THE QUALITY INDEX ON REENLISTMENT

Service/ Occupation	Percentage Effect on the Probability of Reenlistment
Army	
MOS11C	12.32[a]
MOS91B	4.68[a]
Navy	
NECRM	7.79[a]
NECBT	5.90[a]
Air Force	
AFSC20	2.88
AFSC42	5.76[a]
Marine Corps	
MC03	16.98[a]
MC60	8.08[a]

NOTE: Logistic coefficients are evaluated at the mean probability of reenlistment. The performance index is measured in standardized units.

[a]Statistically significant at the 1 percent level.

component of the *Q*-index must fall by a like amount. Thus, for Army MOS11C, the coefficient of −0.02 for AFQT means that a compensated increase in AFQT of 10 points reduces the probability of reenlistment by 0.2 percentage point. For the other occupations, this reduction is much larger. In MOS91B the same 10-point increase in AFQT would result in a 2.2 percentage point reduction in the probability of reenlisting. A similar result holds for education. Again using the first occupation in the table, someone without a high school diploma, but with a compensating higher value of the unobserved component, would be more likely to reenlist—by 0.68 percentage point.

We interpret these results, which are similar in sign but vary in magnitude for the other occupations, as showing that the reenlistment process selects for those whose comparative advantage is in the military. If we interpret AFQT and education as being associated with

Table 7

EFFECTS OF THE PERFORMANCE INDEX ON REENLISTMENT: A DECOMPOSITION

Service/ Occupation	Q-index	AFQT	Less than High School	More than High School
Army				
MOS11C	14.03[a]	-0.02[a]	0.68[b]	-0.55[b]
MOS91B	10.54[a]	-0.22[a]	0.99[b]	-21.69[b]
Navy				
NECRM	8.38[a]	-0.31[a]	8.37[b]	2.20
NECBT	6.84[b]	-0.23[b]	4.03	15.83
Air Force				
AFSC20	36.47[a]	-1.69[a]	-18.14[a]	-20.23[a]
AFSC42	33.30[a]	-1.24[a]	55.35[a]	-163.90[a]
Marine Corps				
MC03	18.82[a]	-0.11	5.01[b]	-9.27
MC60	9.14[b]	-0.07	4.88[b]	-5.98

NOTE: Logistic coefficients are evaluated at the mean probability of reenlistment. The Q-index is measured in standardized units.

[a]Statistically significant at the 1 percent level.

[b]Statistically significant at the 5 percent level.

generalized ability for both civilian and military jobs, whereas the Q-index measures military specific ability, then the coefficients in Table 7 tell a story consistent with assortive job matching. That is, for a given level of civilian job ability, those with higher military job ability are staying in the military. Likewise, for a given level of military job ability, those with greater civilian job ability would be expected to leave the military.

While this comparative advantage sorting is taking place, it is also true that the military is keeping those with higher absolute military ability (the coefficient of the Q-index in Table 6). Thus, although the economic theory of comparative advantage receives support in Table 7, the operationally important result, for the Services, is the gross effect in Table 6.[5]

[5]Other results in Table 7 deserve some comment. All of the statistically significant coefficients in this table have the "correct" sign in that they conform with our prior

Table 8 summarizes these results by comparing the average value of the Q-index, the background index (a weighted average of AFQT and education only), and its components for different groups of enlistees according to their disposition: those who either leave before ETS or separate with something less than an honorable discharge (ATTRIT), those who separate at ETS (ETSEP), and those who reenlist for a second term (REUP). The Q-index column reflects what we have seen in previous tables: Except for the Air Force, those who reenlist have a *much higher* performance index than those who either attrit or separate. By comparison, differences between the average values of the background index (which are in the same units as the Q-index) are *very small.* Not surprisingly, the differences in the components of this background index are also small. Results for the Air Force show little difference among those who attrit, reenlist, or separate. This may be due to the relative homogeneity of Air Force enlisted personnel, who almost universally have high school diplomas and high AFQT scores as well.

---

notions concerning sorting according to military or civilian comparative advantage. The coefficients for the "more than high school" variable, which indicates that the individual had at least some college, are very unstable, possibly because of the small number of people in this schooling group. In retrospect, we would probably have grouped them with high school graduates.

The results for the Air Force are also quite different from those in the other Services. Although the effect of the Q-index was very low for this Service in Table 6, that is, without controls for background characteristics, its coefficient in Table 7 is the largest. Part of the explanation lies in the fact that most of the variation in the Q-index for the Air Force was due to variation in AFQT and education. Thus, having controlled for these variables in Table 7, there is relatively little variation in the Q-index remaining. Thus, to be one standard deviation above the mean in the Q-index would require that this individual have extremely quick promotion times relative to his peers. Such people reenlist at a very high rate.

Table 8

AVERAGE CHARACTERISTICS BY DISPOSITION OF ENTRY COHORT

Occupation/ Disposition		Q-Index	Background Index	AFQT	More than High School	Sample Size
Army						
MOS11C						
	ATTRIT	-0.460	-0.041	49.5	0.442	426
	ETSEP	0.078	-0.004	49.2	0.509	680
	REUP	0.521	0.075	53.4	0.569	274
MOS91B						
	ATTRIT	-0.632	-0.200	54.7	0.443	235
	ETSEP	0.124	0.077	57.6	0.725	658
	REUP	0.365	-0.022	54.4	0.678	183
Navy						
NECRM						
	ATTRIT	-0.416	-0.013	55.7	0.656	212
	ETSEP	-0.020	0.006	56.4	0.843	977
	REUP	0.434	-0.011	51.9	0.775	249
NECBT						
	ATTRIT	-0.335	-0.063	56.1	0.419	592
	ETSEP	0.105	0.033	57.5	0.613	1014
	REUP	0.389	0.018	55.3	0.604	235
Air Force						
AFSC20						
	ATTRIT	-0.085	-0.027	68.4	0.967	149
	ETSEP	-0.021	0.012	69.5	0.989	621
	REUP	0.072	-0.010	69.2	0.983	359
AFSC42						
	ATTRIT	-0.095	-0.041	50.6	0.927	382
	ETSEP	-0.024	0.021	52.7	0.940	780
	REUP	0.101	-0.001	51.5	0.944	550
Marine Corps						
MC03						
	ATTRIT	-0.417	-0.036	57.2	0.360	309
	ETSEP	-0.063	-0.019	50.9	0.526	1035
	REUP	0.970	0.156	57.7	0.655	200
MC60						
	ATTRIT	-0.667	-0.085	68.5	0.409	328
	ETSEP	0.052	0.016	66.8	0.577	1219
	REUP	0.435	0.022	66.3	0.590	358

NOTE: The *Q*-index is derived from promotion histories as well as AFQT and education. The background index is derived from AFQT and education only.

V. CONCLUSIONS

We started out by asking the question: Have the Armed Forces been successful in attracting the best enlisted personnel into the career force? On average, are those who reenlist of higher or lower "quality" than those who either fail to complete their first term or fail to reenlist? We answered this question by constructing an index of performance that combined entry-level characteristics of recruits and information on their promotion histories to grades E4 and E5. Having developed this index, we addressed a related set of questions: Are the traditional measures of enlisted personnel quality—AFQT scores and education—meaningful predictors of job performance? Are these background characteristics of equal importance across different occupations and different services? How important are other unobserved factors in determining military-specific job performance?

Although it is somewhat risky to draw general conclusions, having looked at only a small number of occupations for a fairly old (1974) entry cohort, the uniformity of the results gives us some confidence. We believe that, when measured by relative performance in the first term, the military retains much better people in the career force than it loses. Those lost through attrition have the lowest performance. Those who separate at ETS have about the same performance as the average of those entering the military. However, those retained in the career force have the highest performance of any group. In contrast, when personnel are ranked by educational attainment or AFQT scores, the results are mixed and no uniform picture emerges. These background characteristics matter very little in predicting the success of the first-term enlistee, except perhaps in the Air Force. In any case, reliance on these measures alone to assess the quality retention "problem" is inappropriate.

We also found evidence that the reenlistment process selects those whose comparative advantage is in the military. That is, for a given level of general ability (as measured by background characteristics) those with higher military job ability tend to stay on. Likewise, those with higher AFQT and education are less likely to reenlist when their military performance is controlled for. The operationally important result is, however, the *combined* effect of both background characteristics and other "ability factors." By this measure, the military is successful in retaining those with higher absolute military ability.

We were able to quantify the relative importance of the different components of job performance, at least for the eight occupations analyzed. We found that AFQT scores have a uniformly small quantitative effect on the performance index across occupations when compared to the presence or absence of a high school diploma. The tradeoff between AFQT and a high school diploma ranged from a low of 33 points of AFQT for Navy radiomen to a high of over 200 points for Marine Corps aircraft mechanics. Further, these background characteristics are usually a minor component of the performance index. For Navy radiomen almost all of the variation in job performance, as we measure it, is explained by "ability factors" other than AFQT and education. The Air Force occupations represent the other extreme, possibly reflecting the importance that AFQT and schooling play in that Service's promotion process.

The difference between our approach and other military research bears emphasis. Most studies accept the traditional definition of AFQT and education as "quality." In retention research, then, the problem becomes one of measuring quality loss solely in terms of AFQT and education, or one of devising strategies to retain personnel with more of these entry-level attributes. This same analytic approach also underlies accession research such as the studies by Hiatt and Sims (1980) and Greenberg (1980). They attempted to assess the effects of miscalibrated Armed Services Vocational Aptitude Battery (ASVAB) scores on job performance by validating enlistment standards against alternative measures of job performance.[1] Given their definition of "quality," the research objective became one of finding measures of performance that correlated highly with AFQT and education. In our study, on the other hand, these entry-level characteristics are treated as one component of job performance. Their importance as predictors of job performance is a proposition we examined empirically.

The implications of these results are that, in reenlistment policy, Service-wide guidelines on eligibility to reenlist should not be heavily weighted toward either mental category or educational achievement, and should in any case allow waivers of standards that vary by military occupation. The rank attained at the end of the first term, and perhaps the speed with which that rank is achieved, should receive much larger weight in these decisions. Further, given the small measured weight of AFQT in our quality index, we speculate that the well-publicized misnorming of ASVAB scores had minor detrimental effects

[1]Between 1976 and 1980, the ASVAB miscalibration inflated the test scores of many enlistees who would not have been admitted otherwise under correctly normed standards. OASD first brought this problem to light in the report, *Aptitude Testing of Recruits*, July 1980.

on overall military job performance. More definitive conclusions must await the application of this methodology to a broader range of military occupations.

Appendix

THE MODEL: TECHNICAL DETAILS

The formal structure of our model is that of a multiple indicator multiple cause (MIMIC) model with discrete outcomes. In our case, the latent variable underlying the structural model is the unobserved quality of a first-term recruit as manifested in his promotability. In the text, we call this measure the "quality index." We think of quality as being determined, in part, by observable background variables—education and AFQT test scores—and in part by unobservable, person-specific variables. Job match quality is observed indirectly through the promotion rates of the recruit. That is, if a good match has been made we expect to see a higher likelihood of promotion. Because we see more than one promotion event it is possible, given sufficient restrictions on the model, to identify important features of this structural model.

The formal structure of the model stems from the work of Jöreskog and Goldberger (1975), extended to the case of discrete outcomes by Muthen (1979). Our work represents a further extension in that the discrete outcomes which we observe—promotions—are time dates. We draw from the literature on survival models (Cox (1972)), and from Kalbfleisch and Prentiss (1980) in formulating our specification of promotion times. These latter models are most closely related to those of Flinn and Heckman (1982), and Heckman and Singer (1982), whose analysis of time-dated events strongly influenced our work.

Our starting point is a model for promotion times. The discussion which follows assumes that quality, q, is observed. After developing the promotion time model conditioned on q we proceed to develop the full, joint distribution of q and promotion times.

Let T be a nonnegative continuous random variable representing the time of promotion to a given rank in a homogeneous (conditional on q) population. The distribution function is defined by

$$1 - F(t;q) = P(T \geq t;q) \qquad 0 < t < \infty$$

and gives the probability that the promotion time occurs later than t. The probability density of T is

$$f(t;q) = \frac{dF(t;q)}{dt}$$

The hazard function gives the instantaneous rate of failure at t conditional on survival to time t, i.e.,

$$\lambda(t;q) = \frac{f(t;q)}{1 - F(t;q)}$$

The model we adopt here is the proportional hazard form of the hazard function,[1] where

$$\lambda(t;q) = \lambda_0(t) \exp(\gamma q)$$

In this case the variable q acts multiplicatively on the base-line hazard $\lambda_0(t)$. Assuming that γ is positive, those with higher q will have proportionally higher instantaneous probabilities of being promoted. Since

$$1 - F(t;q) = \exp\Big(- \int_0^t \lambda(u;q)du\Big)$$

$$= \exp\Big(- exp(\gamma q) \int_0^t \lambda_0(u)du\Big)$$

$$= [1 - F_0(t)]^{\exp(\gamma q)}$$

where

$$1 - F_0(t) = \exp\Big(- \int_0^t \lambda_0\,(u)du\Big)$$

the proportional hazard form restricts the effect of q on the probability of nonpromotion to be of the form

$$\ell n[1 - F(t;p)] = \exp(\gamma_q\, \ell n[1 - F_0(t)])$$

[1]We chose this form for computational convenience. In addition, the translation from continuous to discrete time is a natural one for the proportional hazard model, a property not held by other forms.

For example, if the hazard were constant, $\lambda_0(t) = \lambda_0$, the model would reduce to

$$\ell n[1 - F(t;q)] = \exp(\gamma_q + \lambda t)$$

THE DISCRETE TIME MODEL

The principal advantage of continuous time models over discrete time models is that the former are not sensitive to the time unit used in empirical applications (Flinn and Heckman (1982)). In general, the parameters of a discrete time model will change as the time units involved in the empirical analysis are changed.[2] In our empirical work we use monthly data, but we integrate the proportional hazard model to monthly units. A promotion in a given month means that the promotion occurred sometime during that month.

A straightforward use of the continuous time density function when the data are generated by time aggregation can lead to many cases of "ties." That is, the continuous time model expects that the number of failures in a given time interval will approach zero as the interval shrinks to width zero. If we ignore the time aggregation of our data, then a promotion at, say, 12 months in service is different from one at 11.8 months in service, even though the empirical data cannot distinguish between the two. Thus, from the perspective of a continuous time model the empirical hazard assumes a very strange form, having positive mass at integer values of time (in months) but zero mass between these points.[3] To avoid these difficulties we explicitly aggregate the continuous time model. That is, the discrete event we observe corresponds to

$$P(T\epsilon(a-1,a) \mid T > a-1) = \frac{\int_{a-1}^{a} f(u;q)du}{1 - F(a-1;q)}$$

But

[2]For example, a monthly promotion probability model will generate parameter estimates that are not necessarily consistent upon time aggregation with a quarterly promotion model. Different investigators using the same data but different time units would arrive at different results even though their models and functional forms were identical.

[3]Breslow (1974) and Peto (1972) document the effect of this "heaping" of times to integer values when continuous time models are used.

$$\int_{a-1}^{a} f(u;q)du = [1 - F(a - 1;q)] - [1 - F(a;q)]$$
$$= [1 - F_0(a - 1)]^{\exp(\gamma q)} - [1 - F_0(a)]^{\exp(\gamma q)}$$

Also,

$$1 - F_0(a) = [1 - F_0(a - 1)] \exp\left(-\int_{a-1}^{a} \lambda_0(u)du\right)$$

so that

$$\int_{a-1}^{a} f(u;q)du = [1 - F_0(a-1)]^{\exp(\lambda q)}\left(1 - \left[\exp\left(-\int_{a-1}^{a} \lambda_0(u)du\right)\right]^{\exp(\lambda q)}\right)$$

Thus, the discrete hazard is

$$P(T\epsilon(a - 1,a) \mid T > a - 1) = 1 - \lambda_{a-1}^{\exp(\lambda q)}$$

where

$$\lambda_{a-1} = \exp\left(-\int_{a-1}^{a} \lambda_0(u)du\right)$$

The parameter γ is thus invariant to the time unit. Below, we will write $q = X'\beta + \epsilon$ so that q is a linear function of both observable and unobservable variables. These parameters are also invariant to the time unit chosen.

FORMULATION OF THE LIKELIHOOD FUNCTION

Conditional on q, our data take the form of a sequence of promotion times during the first term of enlistment. We measure these promotion times from the onset of exposure to promotion to a grade. Thus, for promotion to E4 time is measured from the date of enlistment until promotion. If promotion never occurs, the observation is censored because of attrition or expiration of term of service. For a single promotion event, the likelihood for an individual promoted in month T_4 is

$$\ell n_i = \prod_{\tau=1}^{T_4-1} \lambda_\tau \exp(\gamma q)\,(1 - \lambda_{T_4}^{\exp(\gamma q)})$$

that is, the product of probabilities of nonpromotion for the first $T_4 - 1$ months multiplied by the promotion probability in month T_4. On the other hand, for someone who is never promoted to E4 and exits the military at month S_4, the likelihood is

$$\ell_i = \prod_{\tau=1}^{S_4} \lambda_\tau^{\exp(\gamma q)}$$

We treat the second event—being promoted to E5—in a similar fashion. Exposure time to E5 begins in the month after attaining E4 and runs until the individual is either promoted or censored. The full likelihood for both promotion times is, assuming independence, the product of the above two terms. The independence assumption is not implausible since we are conditioning on q. The likelihood of promotions to both E4 and E5 is

$$\ell_i = \prod_{\tau=1}^{T_4-1} \lambda_{4\tau}^{\exp(\gamma_4 q)} \left(1 - \lambda_{4T_4}^{\exp(\gamma_4 q)}\right) \prod_{s=1}^{T_5-1} \gamma_{5s}^{\exp(\gamma_5 q)} \left(1 - \lambda_{5T_5}^{\exp(\gamma_5 q)}\right)$$

With promotion to E4 but censoring at S_5, for someone who is never promoted to E5,

$$\ell_i = \prod_{\tau=1}^{T_4-1} \lambda_{4\tau}^{\exp(\gamma_4 q)} \left(1 - \gamma_{4T_4}^{\exp(\gamma_4 q)}\right) \prod_{s=1}^{S_5} \lambda_{5s}^{\exp(\gamma_5 q)}$$

In principle, maximization would be over the set $\{\lambda_{4\tau}, \lambda_{5s}\}$ as well as γ_4 and γ_5. The number of parameters involved would, however, be prohibitively large. We have instead constrained the λ's to lie on a function $\lambda_\tau = h(\Theta)$, where the number of Θ parameters is small.[4]

[4]The proportional hazard model has the attractive feature of allowing condensation of the likelihood function with respect to the incidental parameters $\{\lambda\}$, so that the γ's may be estimated without reference to the λ's. Given observations on q, we could follow this approach. When q is not observed, however, (simple) condensation is not possible (see below).

HETEROGENEITY AND THE STRUCTURAL MODEL

Because "performance" of a given recruit (the i-th) is not observed, we parameterize it as a linear function of background characteristics:

$$q_i = X_i B + \epsilon_i$$

Heterogeneity occurs because of both observed (X_i) and unobserved (ϵ_i) differences among individuals. In this model we consider both of these sources of heterogeneity as time invariant. We have in mind a process wherein natural cognitive ability, disposition for military life, and both the ability and willingness to perform one's duties in an occupation are relatively fixed over the first term.

This performance index is not observed perfectly by the recruit's superiors (much less by us as investigators), so that some errors result in the relationship between "true" and predicted performance. We can write this relationship as

$$q_{it} = q_i + U_{it}$$

where the error U_{it} is assumed to be serially and cross-sectionally independent but not necessarily of constant variance. As information is gathered about the recruit during his training and thereafter, we would expect the variance of U_{it} to decline over time.

The promotion decision rule for a given individual is of the following form: promote if $q_{it} > L_t$. That is, the best estimate of a recruit's quality is made and is tested against a cut-off limit. This limit L_t is determined by manpower requirements for that year and will, generally, be different for each military occupation. We assume that within a given entering cohort (FY74) and within a given MOS, the time path of L_t is common to all recruits.[5] Thus, a promotion occurs if

$$q_{it} \geq L_t$$

or if

$$U_{it} \geq L_t - q_{it}$$

When $h(\cdot)$ is the density of U_{it}, the probability of a promotion is

$$P[promotion \mid q_i, t] = \int_{L_t - q_i}^{\infty} h(Z)\, dZ$$

[5]Random variations in L_t can be subsumed in U_{it}.

This function is approximated by the proportional hazard form where the index t subsumes the critical cut-off level L_t.

Because we have an incomplete measure of q_i in the function $X_i\beta$, promotion time data will be subject to unobservable heterogeneity in recruit quality (ϵ_i). This will affect estimates of the time-shape of promotions. Those recruits who have particularly high ϵ_i (unobserved quality) will be more likely to be promoted, even after controlling for the observables ($X_i\beta$). As these individuals are promoted, those who remain behind will have progressively lower and lower average q_i. The time path of promotion probabilities would exhibit a "humping," first rising with time in grade and then falling. The falling portion is, we would argue, indicative only of progressive selection for lower and lower quality recruits.

We follow the traditional approach to this problem of unobserved heterogeneity by integrating the likelihood with respect to the individual-specific unobservable ϵ_i. As an example, for someone promoted to E4 and censored before being promoted to E5 we calculate

$$\ell_i(\Theta,\beta) = \int_{-\infty}^{\infty} \ell_i(\Theta,\beta,\epsilon_i)\, g(\epsilon_i)\, d\epsilon_i$$

where

$$\ell_i(\Theta,\beta,\epsilon_i) = \prod_{\tau=1}^{T_4-1} \lambda_{4\tau}^{\exp[\gamma_4(X_i\beta+\epsilon_i)]} \left(1 - \lambda_{4T_4}^{\exp[\gamma_4(X_i\beta+\epsilon_i)]}\right)$$

$$\prod_{s=1}^{S_5} \lambda_{5s}^{\exp\gamma_5(X_i\beta+\epsilon_i)}$$

For all of our estimates we constrain the parameters ($\lambda_{4\tau}$, $\lambda_{5\tau}$) to lie along a logistic-type function,

$$\lambda_{4\tau} = \alpha_4[1 + exp(-(1 - \beta_4)/\sigma_4)]^{-1}$$

and similarly for $\lambda_{5\tau}$. This functional form constrains the hazard to follow a path depicted in Fig. A.1. We arrived at this form after considerable experimentation and consideration of military promotion practices in the first term. DoD guidelines in 1974 specified that promotion to E4 could take place, without a waiver, only after 1.75 years in the service. With a waiver, up to 35 percent of Service-wide promotions to E4 could occur at shorter durations of up to four months. For

E5, the restriction is 2.75 years without a waiver. Up to 10 percent of promotions to E5 could occur with as little as one year's service with a waiver. Promotion to E4 and E5 is more or less automatic *if* certain minimum performance standards are met.

Empirically, the hazard looks very much like that depicted in Fig. A.1 except that, having attained a maximum, the empirical hazard turns down. This is precisely what we would expect with heterogeneous promotion probabilities. Moreover, our interpretation of the first-term promotion process suggests that a declining hazard is unlikely. Since this profile is for a recruit of *given* quality, the promotion probability should not decline, even though it may peak at a very low value. With the proportional hazard framework, the hazards of two individuals i and j will be proportional to each other such that for any duration τ, the ratio of promotion probabilities (say, to E4) is given by

$$\frac{\lambda_{4\tau}^{i}}{\lambda_{4\tau}^{j}} = \lambda_{\tau} \exp((X_i - X_j)\beta + \epsilon_i - \epsilon_j)$$

The hazards for two different individuals are shown in Fig. A.2, with that of the average path.

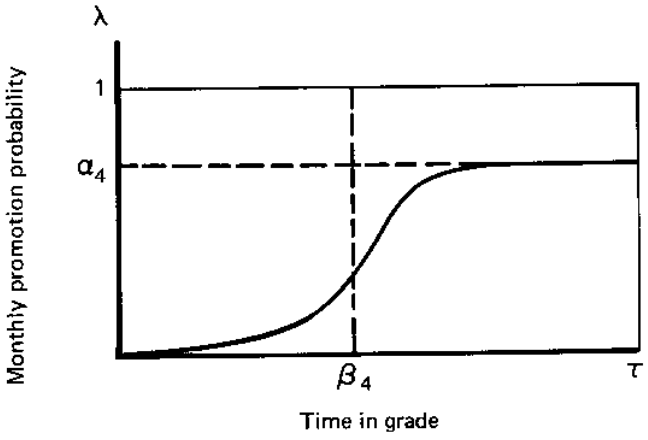


Fig. A.1—Functional form for the hazard function

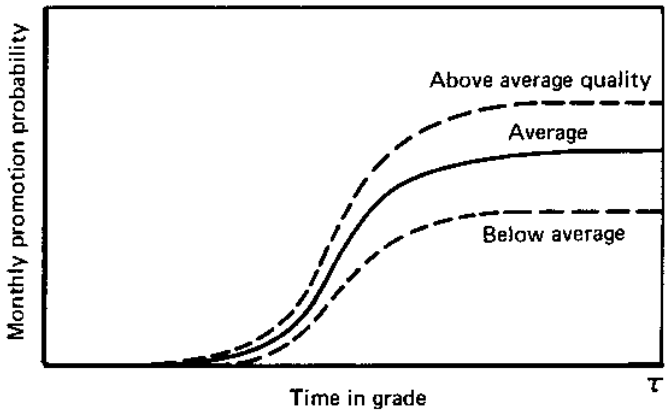


Fig. A.2—The effect of quality on the hazard function

ESTIMATING THE QUALITY INDEX

In the framework of this model, information about recruit quality comes from AFQT, education, and promotion times. If no censoring occurred and all recruits were eventually promoted to E4 and E5, the data would consist of the two promotion times and background variables. In the example below, we suppose that consistent estimates of all parameters could be had from a linear model.[6] If there were only one background variable, the form of the model would be

$$q = \beta X + \epsilon \tag{A.1}$$

$$T_4 = \gamma_4 q + u_4 \tag{A.2}$$

$$T_5 = \gamma_5 q + u_5 \tag{A.3}$$

where all observations are now in deviations from means, and X is a scalar background variable. Consistent with the earlier discussion, we maintain the assumption that u_4 and u_5 are independent.

[6]With uncensored data and some restrictions on the form of the densities of unobservables, a regression model would lead to consistent, though inefficient, estimates. These methods are not available to us because our data are censored. However, they do provide a convenient pedagogical device.

Further, because q is not observed, the scale of this unobservable is arbitrary and we set $\sigma_\epsilon = 1$.[7]

By substituting (A.1) into (A.2) and (A.3), the reduced form model can be written as:

$$T_4 = \gamma_4\beta X + \gamma_4\epsilon + u_4 \tag{A.4}$$

$$T_5 = \gamma_5\beta X + \gamma_5\epsilon + u_5 \tag{A.5}$$

Least squares on each equation yield consistent estimates of $\gamma_4\beta$, $\gamma_5\beta$, and $\gamma_4\gamma_5$. The last derives from the covariance of residuals in this regression, as well as the assumption that u_4 and u_5 are independent and that $\sigma_\epsilon = 1$. Note that there is a sign ambiguity in γ_4 and γ_5, which must be resolved *a priori*. Consistent with our interpretation of q as measuring performance in the military, we restrict these to be negative. Thus, consistent estimates of γ_4 and γ_5 and β are found as

$$(\gamma_4)^2 = \frac{COV(T_4,X)}{COV(T_5,X)} \cdot COV(T_4,T_5 \mid X)$$

$$(\gamma_5)^2 = \frac{COV(T_5,X)}{COV(T_4,X)} \cdot COV(T_4,T_5 \mid X)$$

$$(\beta)^2 = COV(T_4,X)\ COV(T_5,X)\ COV(T_4,T_5 \mid X) VAR^{-2}(X)$$

where the sign of β is the negative of the sign of $COV(T,X)$.

These three parameters represent the effect of recruit performance on promotion times (γ_4,γ_5) and the effect of the background variable X on quality. To get an intuitive sense of the meaning of these parameters suppose, for the moment, that X had no effect on recruit quality. We would see this in a zero (insignificant) regression of T_4 and T_5 on X. Since X does not affect promotion times, both $\gamma_4\beta$ and $\gamma_5\beta$ would be insignificantly different from zero. In this case, the covariance of T_4 and T_5 conditioned on X would still be positive if there were unobserved variations in quality. The unconditional covariance of promotion times tells us that there is a common factor relating the two promotion times, whereas the conditional (on X) covariance tells us what part is due to unobservables rather than to observables.

---

[7]Equation (A.1) could be multiplied by any constant without changing the value of the likelihood, i.e., without changing the sum of squared errors in the reduced form regressions.

More explicitly, note that

$$COV(T_4,T_5) = \gamma_4\gamma_5\ VAR(p)$$

$$= COV(T_4,T_5 \mid X)\ [\beta^2\ VAR(X) + 1]$$

If all of the quality variation were due to unobservable variables rather than to the observed X, then

$$\frac{COV(T_4,T_5 \mid X)}{COV(T_4,T_5)} \rightarrow 1, \text{ from below}$$

Conversely, if the opposite were true and all quality variation were due to X, then this ratio would approach zero. In other words, if a large proportion of the observed correlation between T_4 and T_5 is because each has X in common, then we will attribute (appropriately) a large proportion of the variation in q to X.

The model provides us with three independent measures of q that can be combined to yield an efficient estimate of this unobserved variable:

$$\hat{q}_1 = X\beta$$

$$\hat{q}_2 = \frac{T_4}{\gamma_4}$$

$$\hat{q}_3 = \frac{T_5}{\gamma_5}$$

We can develop an estimate of q either by conditioning on estimated parameter values as well as on data (the tack we follow below and also in the full model) or by conditioning only on the data. The latter would incorporate the variation in β, γ_4, and γ_5 as well as the error terms ϵ, u_4, and u_5. This approach weights the estimates of q according to the inverse of the sampling variation in the parameter estimates as well as the unreducible variation of measurement. We have not followed this approach due to its inherent complexity in the full nonlinear model. We know of no simple technique for doing so. Our approach is, however, asymptotically equivalent to the more complex one since the sampling error variances will converge to zero as the sample size grows while the measurement error variances (ϵ, u_1, and u_2) will not.

Because the measurement errors are independent, we can combine the three measures of q by weighting them inversely to their measurement error variance, i.e.,

$$VAR(q - \hat{q}_1) = \sigma_\epsilon^2 \equiv 1$$

$$VAR(q - \hat{q}_2) = \sigma_{u4}^2 \hat{\gamma}_4^{-2}$$

$$VAR(q - \hat{q}_3) = \sigma_{u5}^2 \hat{\gamma}_5^{-2}$$

and

$$\hat{q} = W_1\hat{q}_1 + W_2\hat{q}_2 + W_3\hat{q}_3$$

where

$$W_1 = \frac{1}{D}$$

$$W_2 = \frac{\sigma_{u4}^{-2}\,\gamma_4^2}{D}$$

$$W_3 = \frac{\sigma_{u4}^{-2}\,\gamma_5^2}{D}$$

$$D = 1 + \sigma_{u4}^{-2}\,\gamma_4^2 + \sigma_{u5}^{-2}\,\gamma_5^2$$

so that

$$VAR(\hat{q} - q) = \frac{1}{D} \quad .$$

$\hat{q}$ will be imperfectly correlated with q because of measurement errors in $\hat{q}_1$, $\hat{q}_2$, and $\hat{q}_3$. Note, however, that if more and more promotion times were added, D would grow and the error variance of q would decline. This means that we eventually obtain a zero variance estimate as the number of measurements of q increase.

For the proportional hazard model, the calculations are somewhat different although the statistical principles are the same. Conditioning on the estimated parameters, the unintegrated likelihood contribution for a single individual attaining both E4 and E5 is of the form $\ell(\Theta \mid X, T_4, T_5, \epsilon)$, where Θ includes all parameters of the model for both

promotion times. We can factor this likelihood as

$$\ell(\Theta \mid X,T_4,T_5,\epsilon)h(\epsilon) = \ell_1(\Theta_1, \mid q,T_4,T_5)\ \ell_2(q,\Theta_2 \mid X),$$

where Θ_1 includes the parameters of the q-conditional hazard, Θ_2 contains the parameters of the q-index, and $h(\epsilon)$ is the marginal distribution of ϵ. We seek the density of q conditioned on the estimated parameters and the data (T_4, T_5, and X). This will be proportional to the above density, i.e.,

$$g(p \mid \Theta,X,T_4,T_5) \ \alpha \ \ell_1(\Theta_1 \mid q,T_4,T_5)\ell_2(q,\Theta_2 \mid X),$$

where α is the factor of proportionality. The right-hand side is now viewed as a function only of the data and of the parameter vector Θ.

To illustrate, consider the case of someone promoted to both E4 and E5. The unintegrated likelihood is[8]

$$\prod_{\tau}^{T_4-1} \lambda_{4\tau}^{\exp\,[\gamma_4(X\beta+\epsilon)]}\left(1 - \lambda_{4T_4}^{\exp\,[\gamma_4(X\beta+\epsilon)]}\right) \cdot$$

$$\prod_{s-1}^{T_5-1} \lambda_{5s}^{\exp\,[\gamma_5(X\beta+\epsilon)]}\left(1 - \lambda_{5T_5}^{\exp\,[\gamma_5(X\beta+\epsilon]}\right)h(\epsilon) =$$

$$\prod_{\tau-1}^{T_4-1} \lambda_{4\tau}^{\exp\,[\gamma_4 q]}\left(1 - \lambda_{4T}^{\exp[\gamma_4 q]}\right) \cdot$$

$$\prod_{s-1}^{T_5-1} \lambda_{5s}^{\exp\,[\gamma_5 q]}\left(1 - \lambda_{5T}^{\exp[\gamma_5 q]}\right)h(q - X\beta) \equiv$$

$$k(q \mid \Theta,X,T_4,T_5)$$

where $k(q \mid \cdot)$ is the unnormalized conditional density of q viewed as a function of Θ and the data. We then estimate q as

---

[8]Throughout we take $h(\epsilon)$ to be the standardized normal, consistent with our normalization rule, $\sigma_\epsilon = 1$.

$$q_i = \frac{\int_{-\infty}^{\infty} q \;\; k_i(q \mid \bullet) dq}{\int_{-\infty}^{\infty} k_i(q \mid \bullet) dq}$$

where the integration is numerical with a 10-point quadrature for each individual in the data. These estimated q's form the data for our subsequent analysis of performance and reenlistment.

REFERENCES

Breslow, N. E., "Covariance Analysis of Censored Survival Data," *Biometrics,* No. 30, 1974, pp. 89–99.

Cox, D. R., "Regression Models and Life Tables," *Journal of the Royal Statistical Society,* No. 34, 1972, pp. 187–220.

Flinn, C., and J. Heckman, "Models for the Analysis of Labor Force Dynamics," *Advances in Econometrics,* Vol. I, 1982, JAI Press, Inc., Greenwich, Conn.

Greenberg, I. M., "Mental Standards for Enlistment: Performance of Army Personnel Related to AFQT/ASVAB Scores," *Implementation of New Armed Services Vocational Aptitude Battery and Actions to Improve the Enlistment Standards Process,* a Report to the House and Senate Committees on the Armed Services, Office of the Assistant Secretary of Defense (MRA&L), December 1980.

Heckman, J., and B. Singer, *The Identification Problem in Econometric Models for Duration Data,* Economic Research Center, NORC Discussion Paper No. 82-6, August 1982.

Hiatt, C. M., and W. H. Sims, "Armed Services Vocational Aptitude Battery (ASVAB) and Job Performance," Center for Naval Analyses, CNA Working Paper 80-3121, November 1980.

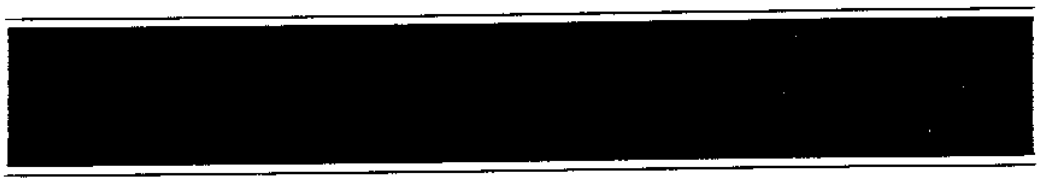
Jöreskog, K., and A. Goldberger, "Estimation of a Model with Multiple Causes of a Single Latent Variable," *Journal of the American Statistical Association,* Vol. 74, 1975, pp. 631–639.

Kalbfleisch, J. D., and R. L. Prentiss, *The Statistical Analysis of Failure Time Data,* Wiley, New York, 1980.

Muthen, B., "A Structural Probit Model with Latent Variables," *Journal of the American Statistical Association,* Vol. 74, 1979, pp. 807–811.

Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics), *Aptitude Testing of Recruits,* July 1980.

Peto, R., "Contributions to the Discussion of Paper by D. R. Cox," *Journal of the Royal Statistical Society* B, No. 34, 1972, pp. 205–207.

Sinaiko, H. W., P. R. Chatelier, C. A. Cook, J. R. Hosek, and G. T. Sicilia (eds.), *Military Personnel Attrition and Retention Research in Progress,* Office of Naval Research, Report No. TR-10, October 1981.

RAND/R-3117-MIL