The Perils of Snapshot Selection Bias in the Workplace

John R. Michael and William R. Schucany
Southern Methodist University, Dallas, Texas

ABSTRACT

This study addresses how personnel staffing strategies affect the distribution of Promotability for incumbents in a job group (Group), where we define Promotability as a combination of traits that predicts job performance (Performance) as rated by management. Promotability can be regarded as a placeholder for the predictor, such as ability, for a criterion, here performance ratings (Ratings). Given the distribution of Promotability for new entries in the Group, and a history of promotion and demotion rates, we derive the distributions of Promotability and Ratings for incumbents, expected Tenure given Promotability, and expected Promotability given Tenure. Promotions and demotions deplete the tails of the distribution of promotability, reduce its standard deviation, and affect the correlation coefficient between Promotability and Ratings. We define three staffing strategies: Skimming, Dredging, and Trimming, which are based upon the relative magnitudes of promotion and demotion rates. We emphasize the distinction between a snapshot sample of incumbents and a representative sample of entries. The threat to validity posed by confusing the two is so common and so perilous that we give it a name: Snapshot Selection Bias. We discuss the consequences of this bias for analytic studies in general and age discrimination in particular.

KEY WORDS
dredging, employment discrimination, incumbents, industrial-organizational psychology, length-biased sampling, meta-analysis, snapshot selection bias

AUTHOR'S FOOTNOTE
John R. Michael is a statistical consultant who provides litigation support services, specializes in employment discrimination, and is concurrently a Research Professor in the Department of Statistical Science, Southern Methodist University, Dallas Texas 75275-0332 (e-mail: john@expertstatistician.com). William R. Schucany is Professor in the Department of Statistical Science, Southern Methodist University.

1. INTRODUCTION

The focus of this study is on the personnel processes that govern the manner in which employees flow through a Group based upon Promotability and the consequences regarding the distribution of Performance in the Group. The relationships among *cognitive mental ability* (Ability), Performance, Tenure, and Ratings have long been topics of interest by researchers in the social sciences, notably industrial-organizational (I/O) psychologists. We propose an elaboration of the conventional model of Performance by drawing upon methods and concepts in statistical science such as *sample selection bias* (SSB), *length-biased sampling* (LBS}, snapshot surveys, and renewal theory. Although these concepts are well known to some statisticians, they appear to be largely overlooked by other statisticians, social scientists, and other practitioners who interpret human resource data. In our consulting practice in the area of employment discrimination, we have encountered many lawyers, labor economists, statisticians, and I/O psychologists are not heedful that *incumbents in a Group* (Incumbents) are subject to LBS, and therefore are not a representative sample of employees who enter the Group.

We propose a bivariate statistical model that takes into account promotion and demotion rates as additional explanatory variables, and enables the relationships among Ability, Performance, and Tenure to be explored more fully. Importantly, this model provides the means for making an exact adjustment for the effects of double-range restriction, an objective that has eluded I/O psychologists for many years. We are hopeful that these results will be of interest to social scientists, notably IO psychologists, working in the areas of personnel selection, Performance, and meta-analysis. We encourage statisticians to elaborate our statistical model further, drawing upon their knowledge in such areas as survey sampling and renewal theory. The results should also be of keen interest to attorneys and expert witnesses in the area of labor law, who are often faced with the challenge of interpreting the relationships among Ability, Performance, and Ratings based upon snapshot samples. But perhaps the most important audience is comprised of the large numbers of employees in the workplace - subordinates, supervisors, managers, and human resource specialists - who frequently draw inferences from data based upon snapshot samples of employees. Since the potential audience for this paper is broad, we relegate the mathematical statistics to the Methods section. However, we illustrate the results beforehand, throughout the paper, all of which are derived in the Methods section.

Statistical scientists continue to advance statistical methods; however, the nature of these methods is frequently an elaboration of existing theory based upon mathematical statistics, as opposed to the adaptation and development of such methods to other disciplines. Case in point: meta-analysis. This is a powerful statistical technique widely used in the social sciences, where the unit of analysis is an individual study: the input data consist of summary statistics from individual studies, and the output is ideally a much more reliable estimate of the criterion variable.

Our observation is that some statisticians live in somewhat of a statistical vacuum, with little interest or involvement in content areas such as Performance in the workplace, or a statistical method such as meta-analysis that is not firmly grounded in the theory of mathematical statistics. On the other hand, social scientists have knowledge of a broad array of statistical methods, but relatively less expertise in the theory of mathematical statistics and how such theory might be brought to bear on the statistical problems they face in their content areas. The point here is that social scientists and statisticians have a lot to offer each other, and a lot to learn from each other. We hope this paper will encourage researchers-and practitioners alike to venture across inter-disciplinary boundaries where their collaboration can be synergistic.

1.1 The Setting

The setting for this study is one that is common in the world of work. The center of interest is a Group comprised of Incumbents who are similarly situated. Figure 1 illustrates the paths that employees take as they flow through the Group. We refer to these employees as *entries* (Entries) upon arrival, as incumbents during their residencies, and as Exits upon their departures. Incumbents periodically receive Ratings by managers that are likely to impact their jobs and their futures. Incumbents depart the group by way of three possible exit processes: promotion, demotion, and transfer, which we formulate (in the Methods section) according to the role that Performance plays in the reason for exit. Promotions and demotions are associated with good and poor Performance, respectively. We classify Transfers as departures for reasons unrelated to Performance, or for reasons unknown. Transfers also include deceased employees who are coded as active on the human resources database.

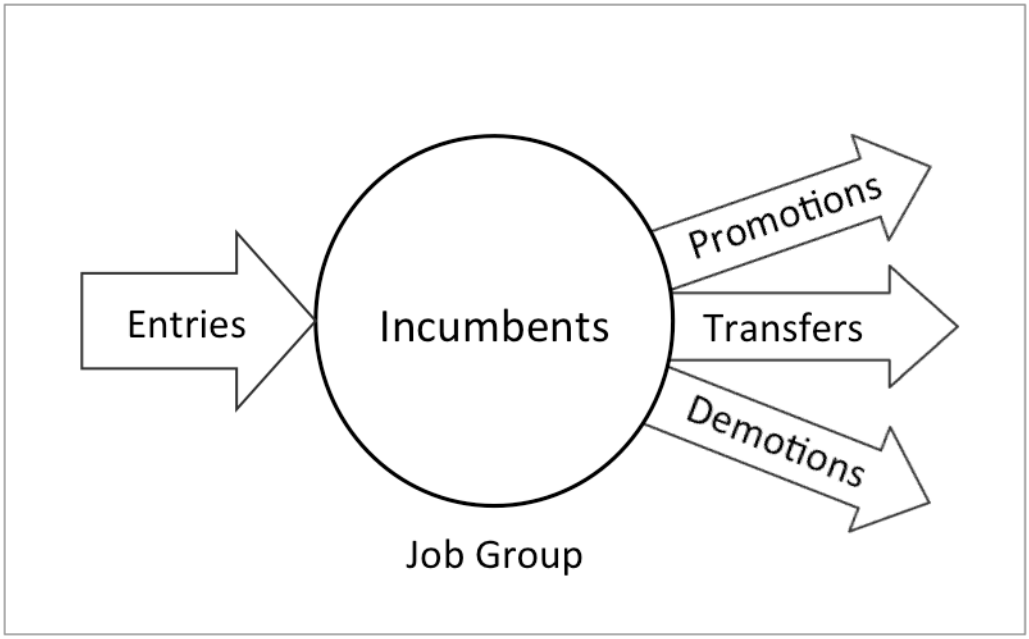

Figure 1. Diagram of the flow of employees through a job Group. We refer to employees as: Entries, upon arrival; Incumbents, during their residencies; and Exits, upon departure. Incumbents exit by way of promotion, demotion, and transfer.

1.2 Origin

The original motivation for this research was a lawsuit in which similarly situated plaintiffs alleged employment discrimination on account of age with regard to promotion. In the legal profession, "similarly-situated" is a term of art used to describe a group of individuals who are similar with regard to legal issues and circumstances, often for the purpose of joining them so they can be tried as a class. In an employment discrimination class action lawsuit, these circumstances typically include measures of organizational unit, job, and location. The temporal scope can range from a snapshot of Incumbents at a single point in time, to an entire organization over a period of years. In the legal setting, it is the court that defines the scope of the class. In a scientific research setting, when modeling Performance, it is researchers and analysts who must thoughtfully define the parameters that define the Group.

In the class-action case that motivated this research, experts for both Plaintiffs and Defendant agreed that the average promotion rate of older workers was significantly below that of younger workers. This disparity constituted prima facie evidence of age discrimination and shifted the burden of proof to the Defendant. Plaintiffs' expert bolstered his argument by noting that older workers had more relevant experience (which is not surprising). Defendant's expert countered that Plaintiffs had received significantly lower Ratings, which explained their relatively low promotion rate. As with many Performance appraisal systems, Ratings were based largely on subjective judgments made by supervisory personnel. The element of subjectivity raised the possibility that Ratings were "tainted" by a discriminatory animus toward older workers. For this reason, the court declined to accept Ratings as a valid explanation for the disparity in promotion rates, and Plaintiffs ultimately prevailed.

1.3 Tenure, Performance, and Ratings

In the lawsuit, Ratings tended to decrease as Tenure increased. This raised fundamental questions about the relationships among Tenure, Performance, and Ratings. Was the observed relationship due to employment discrimination, or was it because some older workers had "topped out," i.e., reached their levels of highest competence (ability to do the job), and had therefore stagnated in the Group, where they grew older? An irreverent version of the latter explanation was popularized in the satirical bestseller, the *Peter Principle* (Peter and Hull 1969) that posited: "in a hierarchy every employee tends to rise to his level of incompetence."

Having worked on similar employment discrimination cases, we had observed different approaches to personnel decision-making in different Groups. For example, in one Group, management might be reluctant to terminate Incumbents who had "topped out," and would control the size of the group by reducing hiring, promoting the best performers, and terminating the very poorest performers. The resulting imbalance in promotion and demotion rates would thereby induce a negative correlation between age and Performance and could give the false impression of Age Discrimination. On the other hand, we observed the opposite pattern in which only the better performers were retained. On the basis of these patterns, we reasoned that the historical promotion and demotion rates might be valuable explanatory variables when modeling the relationship between Tenure and Performance, and could also have value as explanatory values in meta-analyses of Performance.

The two latent variables (constructs) of primary interest in this study are Promotability and Performance. We define Promotability as an individual human trait, or combination of traits, that predicts Performance as rated by management. Being latent variables, neither Promotability nor Performance can be observed directly; however, they can be estimated, and correlated with other variables that can be measured. The criterion measure used in most meta-analytic studies of Performance is Ratings (Hunter and Schmidt 2004, p170). In the Methods section we introduce a bivariate statistical model of Promotability and Ratings upon which all the results presented in this study are based.

1.4 Snapshot Samples.

One of the most important principles we stress in this paper is that Incumbents are *not* representative of Entries. Rather, Incumbents constitute a *snapshot* (Snapshot) in time of employees who flow through the Group. In social research studies, it is common for inferences to be based upon Snapshot samples, as opposed to representative samples, of employees. The term Snapshot is well established in the lexicon of survey research and is in the common parlance of workers, researchers, and attorneys. "Social research depends heavily on surveys, which can provide a snapshot or moving picture of social trends (The American Statistical Association 2011). The term Snapshot is fitting because of the analogy with taking a photograph: first, a specific point in time is chosen; next, the "picture is framed" by focusing on certain positions of interest; and finally, a "snapshot" it taken of those employees who just happen to be in the field of view. Coincidentally, in the lexicon of survey research, the list of items from which a sample is selected is referred to as the sample "frame."

<u>Example 1. Five Incumbents</u>. To illustrate how different a Snapshot sample can be from a representative sample, we contrived an example that is intentionally simple – consisting of but one position. Figure 2 illustrates the completed Tenures of five Incumbents who successively occupied a single position over a period of forty months. Entry A held the position for 2 months, then was replaced by Entry B who held the position for 5 months, and so on. Average Tenure for Entries is equal to the simple average: (2+5+1+23+9)/5) = 40/5 = 8 months. Now consider what happens if the only data available to us is a Snapshot taken at some random point during the 40-month period, which, in the present case, is the Tenure of the one incumbent who filled the position at the time. Each month has the same probability of being selected. This implies that employee A is twice as likely to be selected as employee C, employee D is more likely to be selected than all other employees combined, etc. To estimate the expected value of Tenure from the sample, we take the probability of selection into account by computing a weighted average: (2x2 + 5x5 + 1x1 + 23x23 + 9x9)/(2 + 5 + 1 + 23 + 9) = 16 months, which is fully *twice* the value of the average of a random sample of Entries!

Employee	A	B	C	D	E	Total
Tenure	2	5		23	9	40

Figure 2. Timeline of five employees, A through E, who successively fill a single position over a period of 40 months.

The Inspection Paradox. Feller (1970, p.185) refers to this surprising result as the *Inspection Paradox* in the context of inspecting the lifetimes of batteries. He makes the following assumptions: lifetimes are continuous random variables, identically distributed, statistically independent, and remaining life does not depend on how long the battery has already survived. These assumptions characterize the exponential distribution, which Feller uses to model battery lifetimes. Feller (1970, pp.10-14) introduces the same phenomenon as the *Waiting Time Paradox*, which concerns the time a traveler spends waiting for a bus to arrive. The Inspection Paradox and the Waiting Time Paradox differ only with regard to the circumstances in which they are framed.

Feller observes that waiting times constitute a renewal process (Feller, 1970, p181-185; Vardi, 1988). He defines "spent" waiting time as the time the traveler has already spent waiting for a bus; "residual" waiting time as the additional time needed until the next bus arrives; and "total" waiting time as the sum of spent and residual waiting times. Feller presents some interesting, if not surprising, results that follow from these assumptions. He demonstrates that the statistical distribution of spent waiting time is identical that of residual waiting time, and the distribution of residual waiting time is identical to the unconditional distribution of total waiting time. He comments that many practitioners find these results surprising, even counterintuitive; nevertheless, they are all true, and they all follow from the memory-less property of the exponential distribution.

We make the same assumptions in this paper about the distribution of Tenure as Feller made about distribution of waiting times and battery lifetimes. In particular, we assume that Tenures are continuous random variables, identically distributed, statistically independent, and memory-less. Tenure, thus defined, can therefore be regarded as a renewal process: when one Incumbent departs, a replacement arrives, thereby renewing the incumbency for the vacated position. Analogous to Feller's conclusions, the distribution of current tenure (time already spent in the job) is the same as the distribution of remaining tenure; the distribution of remaining tenure is the same as the unconditional distribution of total tenure; and the expected value of current tenure is half that of total tenure. We address the reasonableness and limitations posed by these assumptions, and ways of overcoming them, in the Discussion section.

1.5 Length-Biased Sampling

Analysts sometimes fall into the trap of regarding a Snapshot sample of Incumbents as if it were a representative sample of Entries. This mistake poses a serious threat to the validity of inferences because *Incumbents are not representative of Entries*. When the probability of selecting a sample item is proportional to its length, the selection procedure is, *by definition*, referred to as *Length-Biased Sampling* (LBS) (Vardi 1982). The concept of LBS is in the everyday vocabulary of survey statisticians, but is foreign to most other practitioners of statistics and scientists in other fields. Snapshot samples selected using a LBS protocol are widespread, both in the workplace and in research settings, and are easily confused with representative samples, often with unfortunate results.

1.6 Snapshot Selection Bias

The threat to validity posed by interpreting a Snapshot sample as a representative sample is so common and so perilous that we give it a name: *Snapshot Selection Bias* (SSB). Incumbents will differ systematically from Entries with regard to Ability because Incumbents almost certainly under-represent the best and worst performers, who are relatively more likely to be promoted and demoted, have shorter Tenures, and be less likely to be intercepted by a Snapshot sample (Michael 1999, Hunter and Schmidt 2004). It therefore follows that Tenure for Incumbents will tend to be larger than the Tenure of new Entries. This is a direct consequence of LBS and the properties of the renewal process discussed above. Differences are perhaps most apparent with variables that are directly related to time, such as Tenure. However, differences are by no means limited to such variables. This is because a variable like Tenure is associated with a wide array of important worker and workplace characteristics. Tenure has been linked to job satisfaction (Mobley 1991) and other important job-related variables. Moving outward from Tenure in this fashion, the network of causal paths can extend to encompass many of the psychological constructs that are the focus of workplace studies. If the construct of interest is causally related to Tenure, either directly or indirectly, then there is a potential for SSB.

2. THREE STAFFING STRATEGIES: SKIMMING, DREDGING, AND TRIMMING

For expository purposes, we introduce three stereotypical *staffing strategies* (Strategies) that are found in the workplace: *Skimming*, *Dredging* and *Trimming*. We classify Strategies according to the relative magnitudes of promotion and demotion rates in the Group. For each Strategy we now present a hypothetical example and accompany it with an illustration that dramatizes the degree to which the Strategy can distort the probability density function (Density) of Promotability for Entries compared to that for Incumbents. We set the annual turnover rate to be the same for each strategy: 20%. We selected the promotion (demotion) rates to be 15% (5%) for Skimming, 5% (15%) for Dredging, and 10% (10%) for Trimming. Regarding the dependency of Ratings on Promotability, we adopted 0.60 for the value of the correlation (validity) coefficient between Ratings and Promotability. Figures 3 through 8 are discrete approximations to continuous density functions that were scaled to represent continuous probability distribution functions (PDFs), and we shall refer to them as such. The details are set forth in the Methods section.

2.1 Skimming

Skimming is the Staffing Strategy in which the promotion rate is substantially larger than the demotion rate. It occurs when the organization has a policy of promoting the best performers and retaining others, except for very poor performers. The term Skimming is taken from the personnel policy in which management promotes the "rising stars," thereby "skimming the cream."

Example 2: Skimming - Entry Level Clerk. Suppose you are a project manager and want to add an entry-level clerk to your project staff. You go to the administrative group and make your request. The personnel policies of the group are well known: getting hired is competitive; entry level clerks who perform satisfactorily are usually promoted within two or three years; median completed Tenure is three years; few clerks are terminated involuntarily; and poor performers tend to hold onto their jobs for many years. Accordingly, the promotion rate is relatively high compared to the demotion rate.

Suppose you are given your choice of two clerks: John and Jane. John has been a clerk for 8 years and Jane has been a clerk for 2 years. In summary:

- Getting hired: competitive
- Promotion rate: high
- Demotion rate: low
- Median tenure: 3 years
* John's tenure: 8 years
* Jane's tenure: 2 years

If you want to select the best performer for your project, whom would you choose? Since John's Tenure is far beyond the median Tenure for the job, it appears that he might have "topped out." Jane is the logical choice.

Illustration of Skimming. Figure 3 illustrates the theoretical distribution of Promotability for Incumbents (shaded) in a Group where the Staffing Strategy is Skimming: the promotion (demotion) rate is 15% (5%). For comparative purposes, it is superimposed over the distribution of Promotability for Entries. The arrows show how Skimming has depleted the upper tail of the distribution of Promotability for Incumbents relative to Entries. Poor performers are likely to be passed over for promotion and may stagnate in the Incumbent pool, where they age.

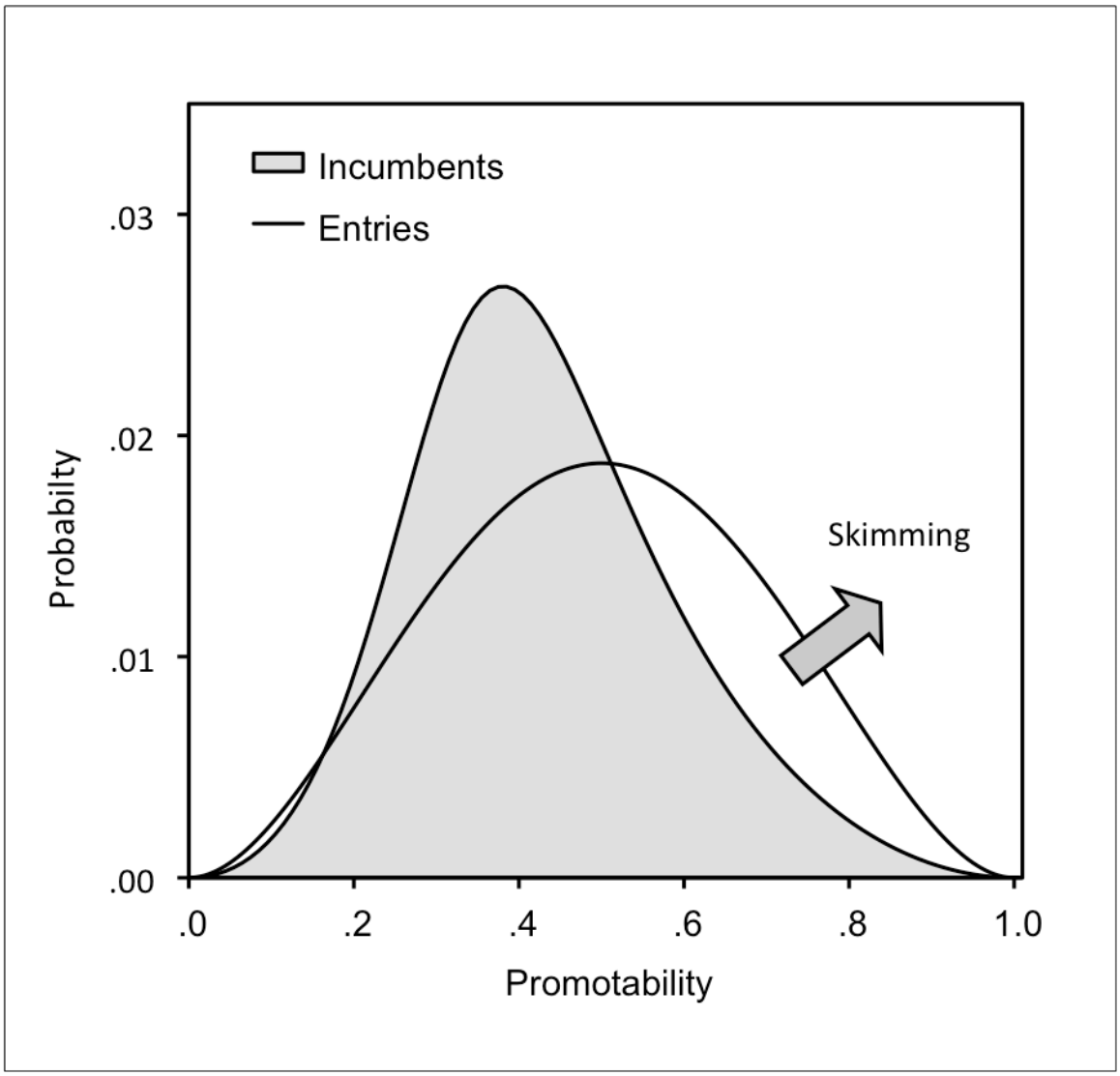

Figure 3. The effect of Skimming on the distribution of Promotability. The probability distribution of Promotability for Incumbents (shaded) is superimposed upon that for Entries to illustrate how Skimming depletes the upper tail of the distribution for Entries. The promotion (demotion) rates for this example are 10% (5%).

2.2 Dredging

Dredging is the Staffing Strategy in which the demotion rate is substantially larger than promotion rate. It occurs when management has a policy of terminating or demoting poor performers. The better performers are more likely to survive, prosper, and occasionally be promoted. The term Dredging is taken from statements sometimes heard in the workplace when management wishes to get rid of the "deadwood," thereby "dredging the silt."

Example 3: Dredging (Real Estate Agent). As a hypothetical example of Dredging, consider the position of real estate agent. Getting hired is relatively easy, since the cost of hire to the broker is minimal. New agents often find the work difficult, the hours long, and the pay low until and unless they begin to sell properties and earn commissions. The needed skills are varied and take time to acquire. Many new agents are dismissed or discouraged and leave within a year or two. The most capable agents are those who prosper and have many years of experience. Most of these better performers remain as agents for the duration of their careers. The one logical line of progression is to the position of broker, but this requires an additional set of managerial skills and an entrepreneurial spirit. There are relatively few such "promotions."

Suppose you go to a brokerage and are given your choice of two agents: John's has been an agent for 8 years, and Jane has been an agent for 2 years. In summary:

- Getting hired: easy
- Promotion rate: low
- Demotion rate: high
- Average tenure: 4 years
* John's tenure: 8 years
* Jane's tenure: 2 year

If you want to select the best agent to help you find a home, whom would you choose? The logical choice is John who, judging by his Tenure appears to have a proven track record and the advantage of more experience.

Illustration of the Effect of Dredging. Figure 4 illustrates the theoretical distribution of Promotability of Incumbents (shaded) in a Group where the Staffing Strategy is Dredging: the promotion (demotion) rate is 5% (15%). For comparative purposes, it is superimposed over the distribution of Promotability for Entries. Referring to Figure 3, for Skimming, the promotion and demotion rates are reversed. Accordingly, Figure 4 is the mirror image of Figure 3. The arrows shows how Dredging has depleted the lower tail the distribution of Promotability for Incumbents relative to that of Entries. The better performers are more likely to survive, remain in the Group, and have relatively high Tenures.

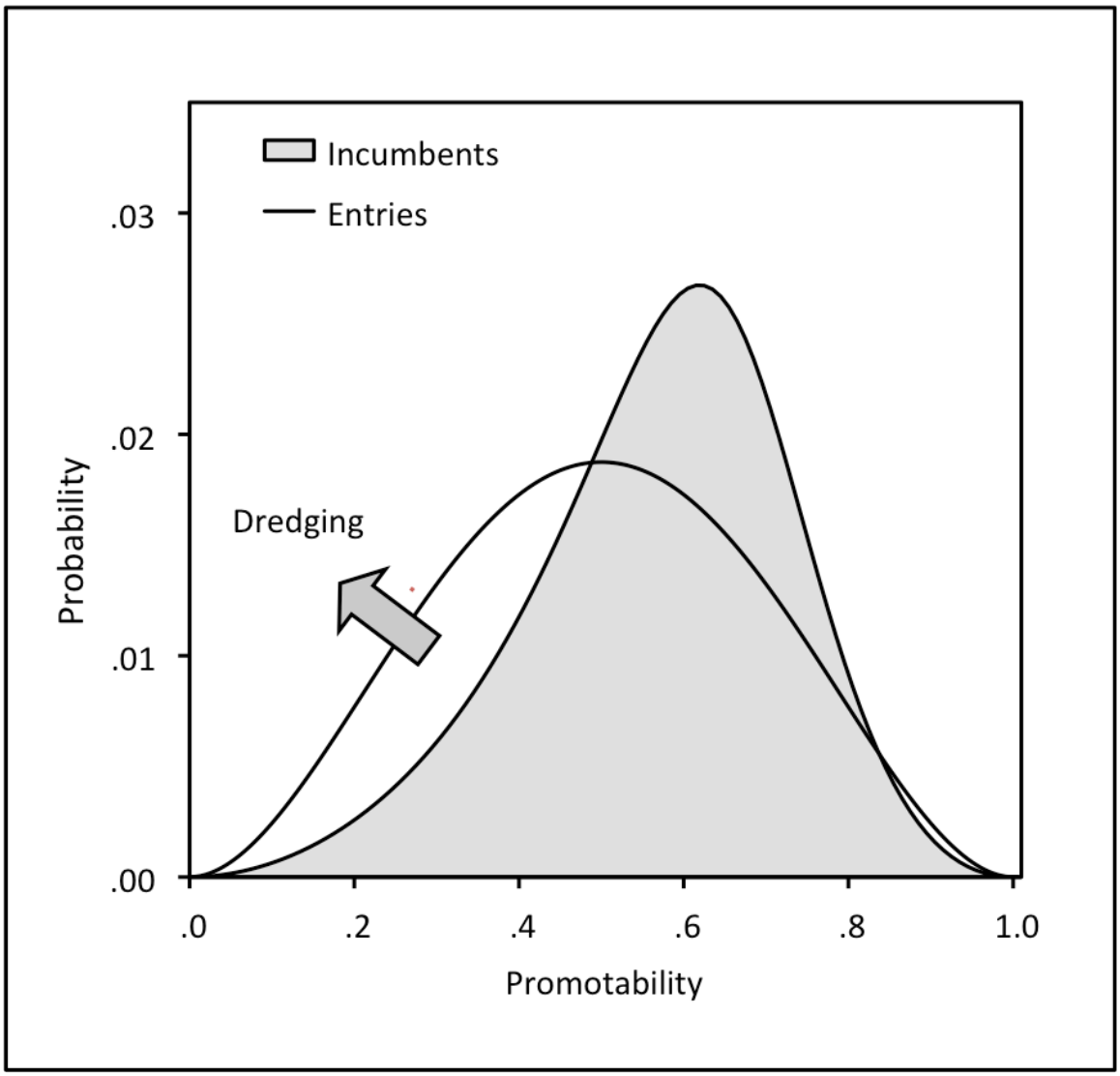

Figure 4. The effect of Dredging on the distribution of Promotability. The probability distribution of Promotability for Incumbents (shaded) is superimposed upon that for Entries to illustrate how Dredging depletes the lower tail of the distribution for Entries. The promotion (demotion) rates for this example are 5% (15%).

2.3 Trimming

Trimming is a Staffing Strategy in which the promotion and demotion rates are both substantial and approximately equal. In the workplace, the phrase "up-or-out" might be used to describe this Strategy.

Example 4: A Minor League Baseball Team. As a hypothetical example of Trimming, consider a minor league baseball team. An important function of the team manager is to identify players with major league potential. Better players are sent to the majors, when the need arises. Conversely, players with low potential are released to make room for other promising candidates. Within several years, about half of new recruits will either be sent to the majors or are cut from the team.

Illustration of the Effect of Trimming. Figure 5 illustrates the theoretical distribution of Promotability for Incumbents (shaded) in a Group where the Staffing Strategy is Trimming: the promotion and demotion rates are both equal to 10%. For comparative purposes, the density function is superimposed on the distribution of Promotability for Entries. The arrows show how Trimming has depleted both the upper and lower tails of the distribution of Promotability for Incumbents in equal amounts, relative to that of Entries.

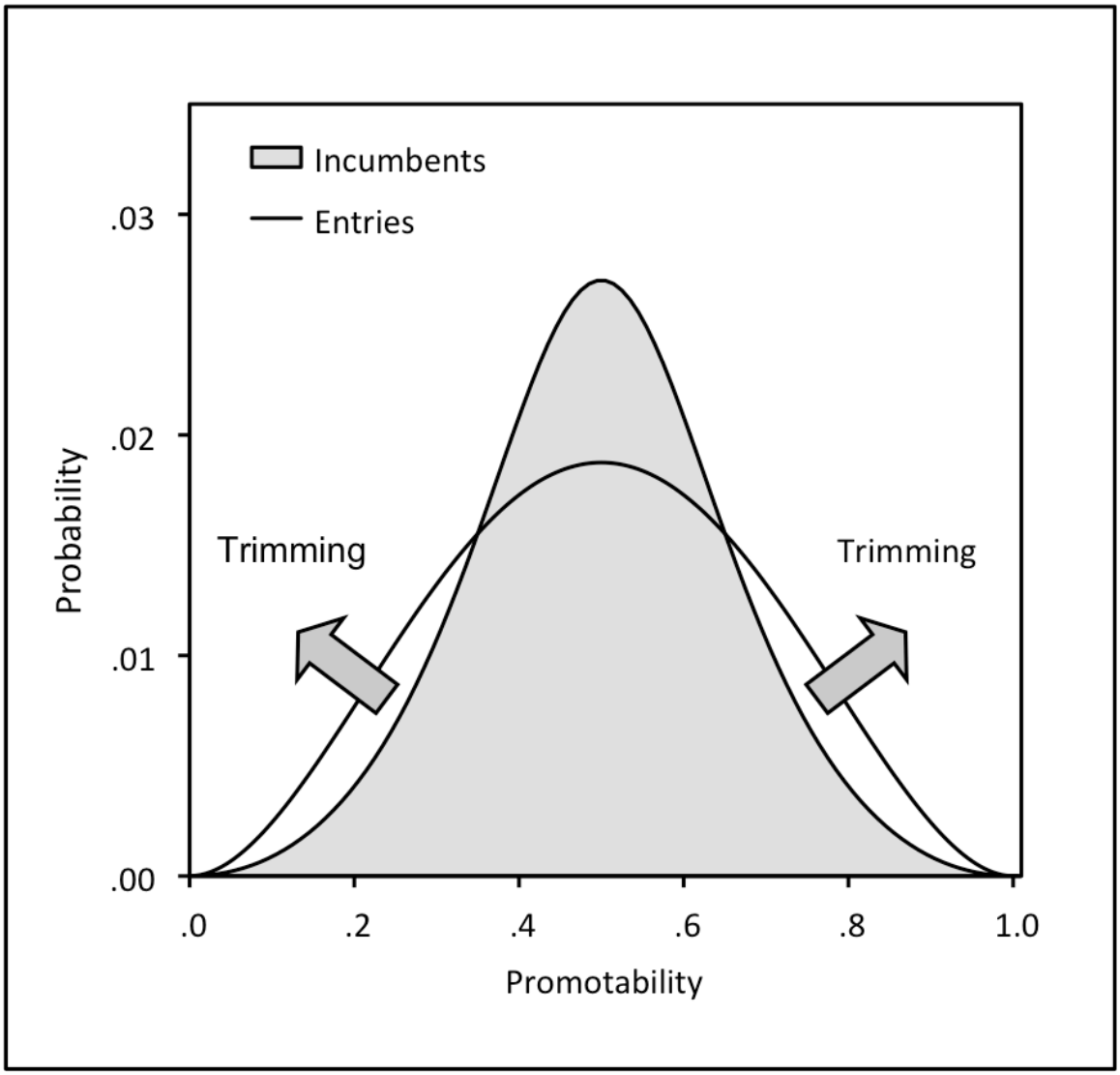

Figure 5. The effect of Trimming on the distribution of Promotability. The probability distribution of Promotability for Incumbents (shaded) is superimposed on that for Entries to illustrate how Trimming depletes both tails of the distribution for Entries. The promotion (demotion) rates for this example are 10% (10%).

Comparison of All Three Staffing Strategies. Figure 6 illustrates the distributions of Promotability for Entries and all three Staffing Strategies so they all can be visually compared. With a 20% turnover rate, the effect of each of Staffing Strategy is to deplete the tails of the distribution of Promotability for Entries and make the density of Promotability for Incumbents more peaked, thereby reducing its standard deviation.

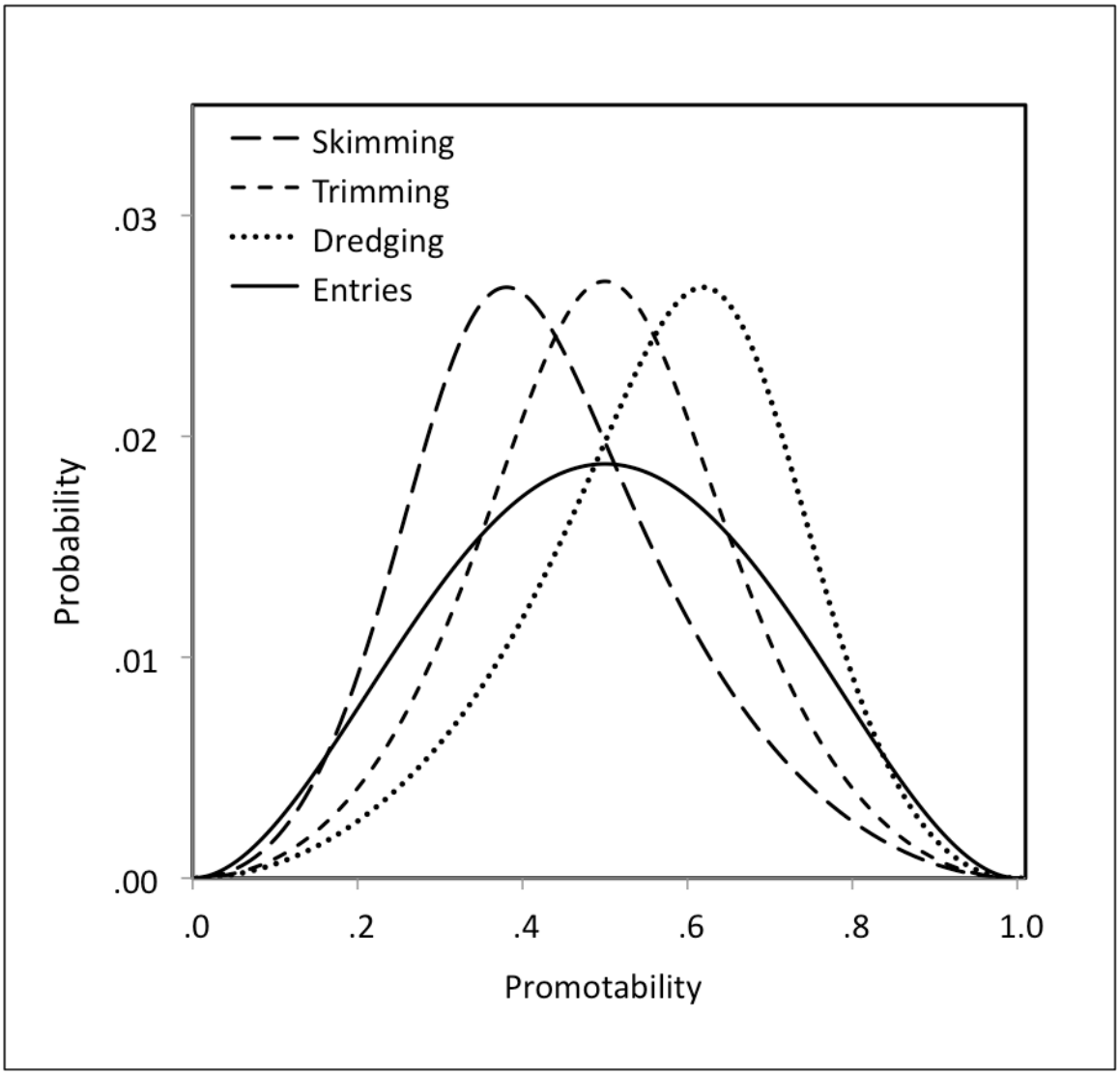

Figure 6. The effects of Staffing Strategies on the distribution of Promotability. Probability distributions of Promotability for all three Staffing Strategies are superimposed on that for Entries so their effects can be compared. The promotion (demotion) rates are 15% (5%), 10% (10%), and 5% (15%).

2.4 Expected Tenure Given Promotability

Once the density of Promotability for Incumbents is obtained, the derivation of Expected Tenure is fairly straightforward. (See the Methods Section.) Expected Tenure given Promotability is illustrated in Figure 7 for all three Staffing Strategies. Their shapes are similar to the distributions in Figure 6; however, they have heavier tails because promotion and demotion decisions are made on the basis of Ratings, which is a stochastic variable, and constitutes an additional source of variation. The maximum expected Tenure of about 8 years is reasonable, considering the setting and Staffing Strategies described above.

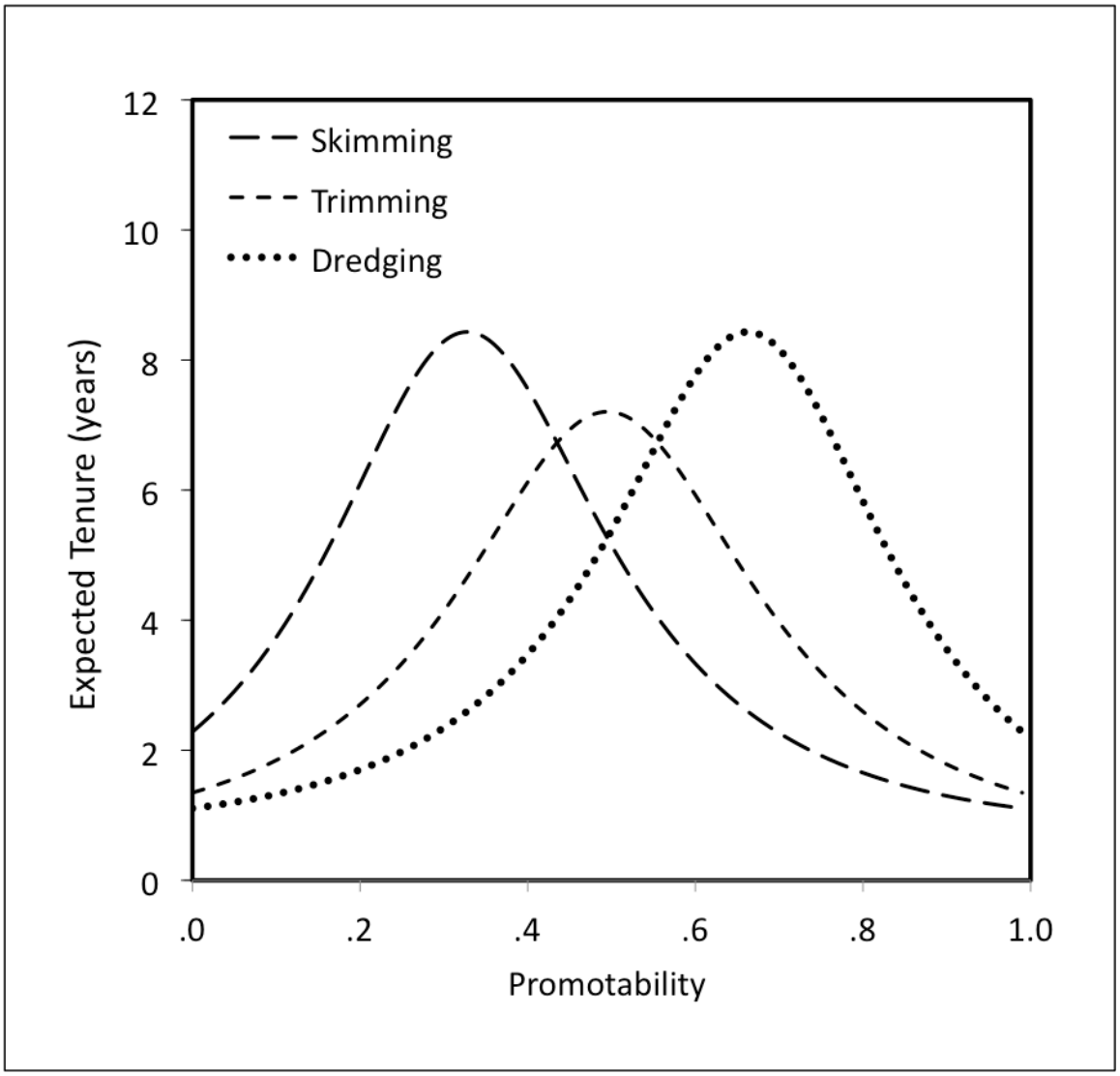

Figure 7. Expected Tenure given Promotability. The relationship of Expected Tenure to Promotability for all three Staffing Strategies are superimposed so their effects can be compared. The promotion (demotion) rates for this example are 15% (5%), 10$ (10%), and 5% (15%).

2.5 Expected Promotability Given Tenure

Figure 8 illustrates expected Promotability as a function of current Tenure. This is an interesting relationship because it provides insight into that important unobservable latent variable, Promotability, given something that can be observed and objectively measured, current Tenure. One might ask the following question: if an Incumbent has occupied the position for, say, 6 years, what inferences can be made regarding the Incumbent's Promotability? Judging from Figure 8, the answer depends upon the Staffing Strategy for the Group. With Skimming, the larger the Tenure, the smaller the expected Promotability. The converse is true with Dredging. With Trimming, current Tenure is not correlated with an Incumbent's Promotability; however, a large value of Tenure increases one's confidence that the Incumbent's Promotability is near the average Promotability for the Group (see Figure 6).

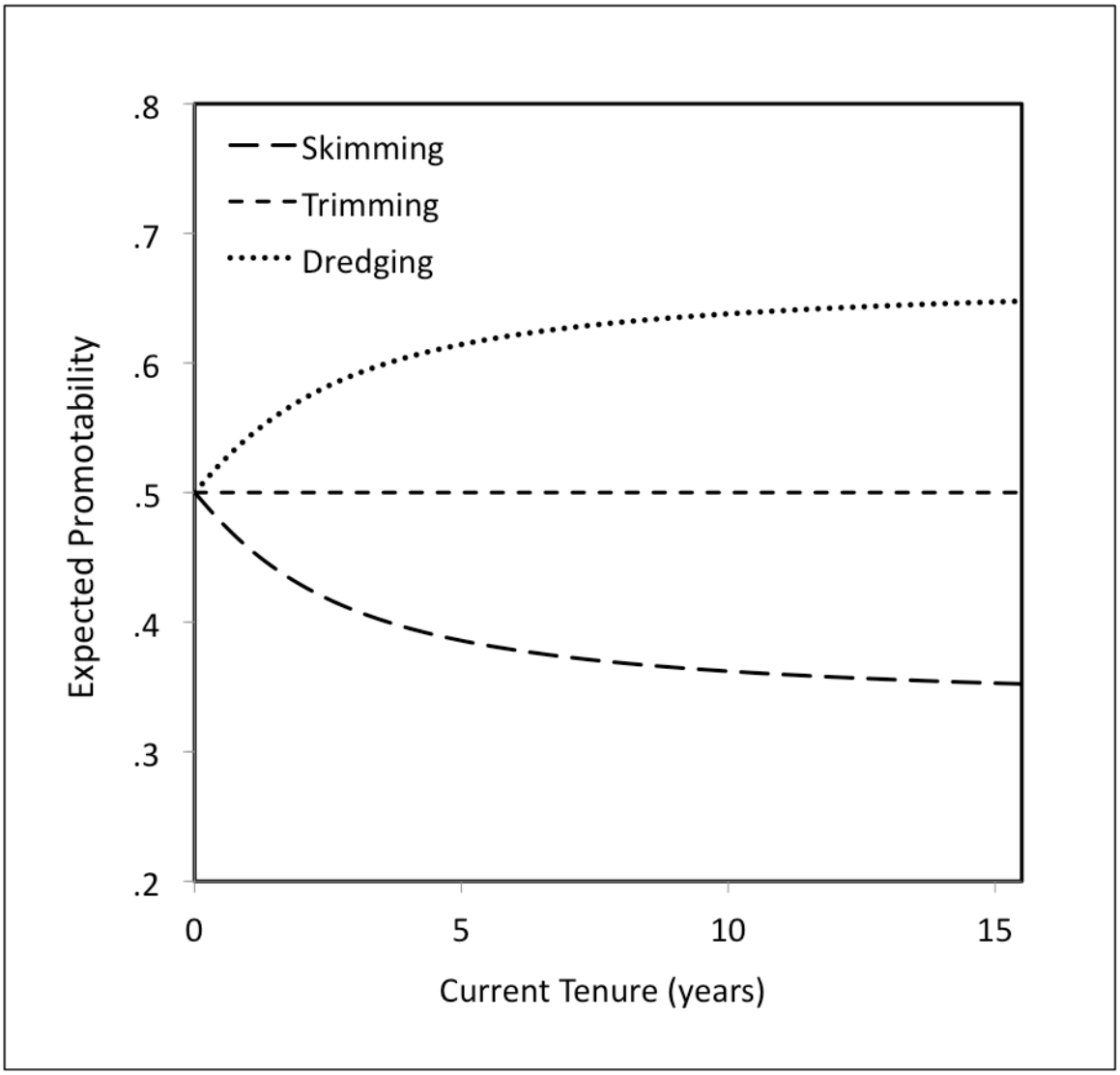

Figure 8. Expected Promotability given Current Tenure. Expected Promotability given Current Tenure for all three Staffing Strategies are superimposed so their effects can be compared. The promotion (demotion) rates for this example are 15% (5%), 10% (10%), and 5% (15%).

3. THE THEORY OF JOB PERFORMANCE

The modeling of Performance falls largely within the purview of researchers in the social sciences, notably I/O Psychologists. Social scientists make heavy use of statistical methods, typically applying them to observational data using statistical software. (Observational data are data that already exist, as compared to experimental data that is obtained by conducting a controlled experiment.) A frequent objective is to establish causality. To this end, social scientists have advanced, and relied upon, specialized statistical methods such as path analysis (or structural modeling) and meta-analysis. Establishing causality is vital objective, but it is often difficult to establish with confidence because of the multitude of threats to the validity of inferences when the data are observational.

Social scientists, notably I/O psychologists, have studied the many potential threats to validity to a far greater extent than have statistical scientists (Cook, Campbell, and Peracchio (1990). Shadish, Cook, and Campbell (2002) discuss 32 different threats to validity grouped into four categories: statistical conclusion validity, internal validity, external validity, and construct validity. The obvious purpose of studying such threats is to help ensure that the practitioner guards against plausible threats, and reports potential threats that could not be controlled. Although many of these threats are couched in statistical terms, statisticians have yet to make substantial contributions to this important area. The *Ethical Guidelines For Statistical Practice* (American Statistical Association, 1999) mandates that the practitioner "Clearly and fully report the steps taken to guard validity," and "Report the limits of statistical inference of the study and possible sources of error." In the opinion of these authors, this practice is infrequently followed.

General Mental Ability. Over 100 years ago Spearman (1904) introduced the psychological construct of General Mental Ability (Ability) and postulated that it played a central role in human cognition and learning. Over the last 30 years, I/O psychologists demonstrated that Ability is an important predictor of Performance for essentially all jobs (Hunter 1980, Schmidt, Hunter and Pearlman 1981, Hunter and Hunter 1984, Ones, Viswesvaran, and Schmidt 1993, Schmidt, Law, Hunter, Rothstein, Pearlman and McDaniel, 1993, Ones, Viswesvaran, Schmidt and Hunter, 1998; Schmidt and Hunter 2004; Schmidt, Shaffer, and In-Sue, 2008). This landmark result, commonly referred to as "validity generalization," created quite a stir in the psychological community, and in the private sector as well, where tests were routinely administered for the purpose of making personnel decisions, because such tests had an adverse impact on certain protected groups (Rushton and Jenson, 2005). In recent decades there has been a resurgence of interest in Ability, Performance, and the predictive validity of Ability for Performance. Today, workplace issues are prominent in the social science literature. Almost all of the articles in the *Journal of Applied Psychology*, a leading publication of the American Psychological Association, are concerned with the world of work. Other scholarly journals, such as *Personnel Psychology*, are devoted exclusively to workplace issues.

Ability and Performance. Schmidt & Hunter (2004) offer evidence that Ability alone is an excellent predictor of both occupational level and Performance. They present a path model that explicates the relationships among Ability, Performance, *Job Knowledge* (Knowledge), and Ratings, and assert that Ability compares favorably to any other trait or disposition when predicting Performance. Schmidt, Shaffer and In-Sue (2008) estimate the true score correlation (validity) coefficient between Ability and Performance to be 0.734.

3.1 Connection with Meta-Analysis

The aforementioned results are most impressive and add immeasurably to our understanding of Performance and personnel processes in the workplace. Moreover, they were made possible by the application of meta-analytic methods, which have been advanced largely by the I/O community, particularly in the last two decades, and are now commonplace in the social science literature. In a meta-analytic study, the unit of observation is an individual study and the variable of interest is the effect size, such as the true correlation coefficient between Ability and Performance. The judicious aggregation of results across studies provides a substantial increase in statistical power, largely through the increase in the effective sample size. However, special care must be taken because there are many other study "artifacts" that must be carefully taken into account so the aggregation of the results across studies is properly adjusted and not sullied by erroneous, biased, or unreliable factors. An "artifact" is an unwanted source of variation that typically increases the variation of the dependent or independent variable.

A sound meta-analytic study requires careful adherence to statistical principles and data collection protocols, such as those that have been painstakingly developed by survey statisticians. Despite the clear relevance of survey science and the need for rigorous data collection protocols, it does not appear that statisticians have made important contributions to the development of meta-analytic methodology. We view meta-analysis as an area where there are rich research opportunities for statisticians to make such contributions by applying their skills in mathematical statistics, survey science, and data collection. We believe the same is true in the area of path analysis (or structural equation modeling), which is statistical in nature.

3.2 Statistical Artifacts in Meta-Analysis

Hunter and Schmidt (2004) present a lengthy review of Meta-Analysis that includes an exhaustive discussion of statistical "artifacts" and data collection protocols for the inputs to a meta-analytic study. Most artifacts, notably the sample size, attenuate (reduce) the estimate of correlation coefficient, ρ. Correction for artifacts in a given study requires auxiliary information about the study, such as sample sizes, means, standard deviations, etc. (Hunter and Schmidt 2004). A succinct summary of meta-analysis can be found in Schmidt, Le and Oh (2009). The authors list 10 artifacts that typically reduce the value of the observed study correlation coefficient, ρ_0, relative to the true correlation coefficient, ρ, which is the parameter that one often wishes to estimate. Most artifacts cause a systematic attenuation (reduction) of the correlation.

Suppose the study correlation is affected by the M artifacts, $\{a_1, a_2, a_3, a_M\}$. Their joint effect is multiplicative. Therefore the relationship between thestudy correlation coefficient, ρ_0, and the true correlation coefficient, ρ, is

$$\rho_0 = (a_1 a_2 a_3, a_M)\, \rho.$$

The value of the i-th artifact is equal to the square root of the reliability coefficient for the i-th step in the measurement process, which is equal to the ratio of the true variation to the variation of our fallible measurement (Snedecor and Cochran 1967). The value of the i-th artifact is equal to the ratio of the standard deviation of the study measurement to that of true measurement for the i-th step in the process.

In the present study we focus on the correlation coefficient between Promotability and Ratings. As noted earlier, Promotability can be thought of as a placeholder for the independent variable the user wishes to correlate with Ratings. For example, the user might choose to define Promotability as Ability, or some combination of personality traits that might include Ability and the so-called Big 5 personality traits (Digman 1990). When estimating the true correlation coefficient, the user would then have to adjust for the artifacts associated with all the factors between, say, Ability and Ratings, such as the measurement error for Ability, the variability of Performance, and Rater unreliability.

3.3 Restriction In Range

In the last decade, a nagging problem in the development of meta-analytic theory has been how to adjust for double-range restriction, which is the joint occurrence of two artifacts that jointly restrict the ranges of the dependent and independent variables. An example of attrition is the loss of poor performers who quit their jobs, or are terminated for poor Performance, and are not captured by the sample. The flip side of this is attrition due to the loss of the better performers who are promoted, or depart for better jobs (Michael 1999, Hunter and Schmidt 2004). The conventional artifacts in meta-analytic studies that represent losses due to promotion and demotion are not multiplicative, because the distributions of these losses are not independent. This renders the use of simple multiplicative adjustment factors inaccurate. Improved adjustment methods have been developed, and these are valuable; however, they are approximations made at the parametric level. Our bivariate model is founded in mathematical statistics and takes these artifacts into account, exactly. We see meta-analysis as an area where statisticians can expand their horizons into in the social sciences and make important contributions.

3.4 Reduction In The Standard Deviation

As noted for Figure 6, with a turnover rate of 20%, the promotion and demotion processes make the densities of Promotability for Incumbents noticeably more peaked, which implies a reduction in the standard deviation for Incumbents relative to that for Entries. The relevant artifact here is the ratio of the standard deviation in the study population (Incumbents) to that in the reference population (Entries).

One of the outputs of our analysis is the standard deviation of Promotability for both Entries and Incumbents for each of the three Staffing Strategies. Relevant statistics are given in Table 1, below. Skimming, Dredging, and Trimming induce artifacts equal to 83%, 83%, and 80%, respectively.

Table 1. Artifacts for Restriction In Range Induced by Staffing Strategies. Attenuation Artifacts are listed for each of the three Staffing Strategies. They are computed as the ratio of the standard deviation of Promotability for Incumbents (0.189) to the that of Promotability for Entries.

Staffing Strategy	λ_p Prom Rate	λ_d Demo Rate	λ_e Exit Rate	StDev Ratings	Attenuation (Artifact)
Entries	—	—	20%	0.189	n/a
Skimming	15%	5%	20%	0.156	0.829
Dredging	5%	15%	20%	0.156	0.801
Trimming	10%	10%	20%	0.151	0.829

The effect of promotions and demotions is to reduce the standard deviation of Ratings from 0.189 for Entries to 0.156 with the Staffing Strategies of Skimming and Dredging, and to 0.151 with Trimming. The corresponding values of the Attrition Artifacts are 83%, 83%, and 80%. This exemplifies the manner in which attrition artifacts can reduce the correlation coefficient in the study population relative to that in the reference population. The bivariate model offers a solution to the long-standing problem double-range restriction. The interested analyst can use the model to compute an adjustment factor for a single study of interest, or run simulations to develop formulas that estimate the value of attrition artifacts as a function of validity, promotion rate, and demotion rate.

4. METHODS

A Bivariate Model. We now introduce the bivariate statistical model used to derive the relationships among Promotability, Ratings and Tenure that are illustrated in Figures 3 through 8. We used a standard algorithm to compute the cumulative Binomial and Beta distributions. Although the derivations in this section are straightforward, the notation can become confusing since we found it necessary to define both continuous and discrete versions of Promotability for both Entries and Incumbents. Therefore, as an aid to the reader, in Table 2 we summarize seven *Random Variables* (RVs) that we discuss in this section.

Table 2. Notation summary for seven random variables.

RV	Distribution	Density	Description
U	beta(α,β)	e (u; 3,3)	Promotability, Entries
K	binomial(k;n)	h (k; 9,u)	Mixing Variable
V	beta(α,β)	p (v; k,3,3)	Supervisor Ratings
X	discretized U	f (x; 3, 3)	Promotability, Entries
Y	discretized V	g (y; 3, 3)	Supervisor Ratings
T	exponential	s (t; λ)	Tenure
Z	(derived)	(derived)	Promotability, Incumbents

The Mixture Method. One of the requirements of model is to be able to control the correlation (validity) coefficient between Promotability and Ratings. The theory of the *Mixture Approach* for simulating bivariate distributions with specified correlation coefficients, introduced by Michael and Schucany (2002), is pivotal in this regard. The authors present the theory and illustrate the method using several bivariate distributions, including the beta, gamma, and uniform. Among all statistical distributions, the normal is perhaps the most familiar to practitioners. Its popularity might be due more to its familiarity and ease of use than to its appropriateness for the problem at hand. In the present case, the normal is unsatisfactory for our purposes because of its symmetry, and because distributions encountered in practice are typically non-normal, e.g., skewed or relatively heavy-tailed.

4.1 The Bivariate Beta Distribution

We selected the bivariate beta distribution to model the joint distribution of the RVs for Promotability for both Entries U and Ratings V. One reason was because the beta has two shape parameters, alpha (α) and beta (β) (in addition to upper and lower bounds for the domain, which are typically 0 and 1, as in this case). Two shape parameters permit a rich variety of distributional shapes to be represented (Johnson and Kotz 1970b), and fit to a given set of data. This is particularly useful if the user opts to transform the data to the unit interval for subsequent analysis. (See the Discussion section regarding the utility of this strategy). Figure 9 illustrates the beta density with different choices of the parameters alpha (α) and beta (β). When $\alpha=\beta$ the distribution is symmetric. The beta (1,1) is the familiar uniform distribution, which is routinely used to simulate random numbers on the unit interval.

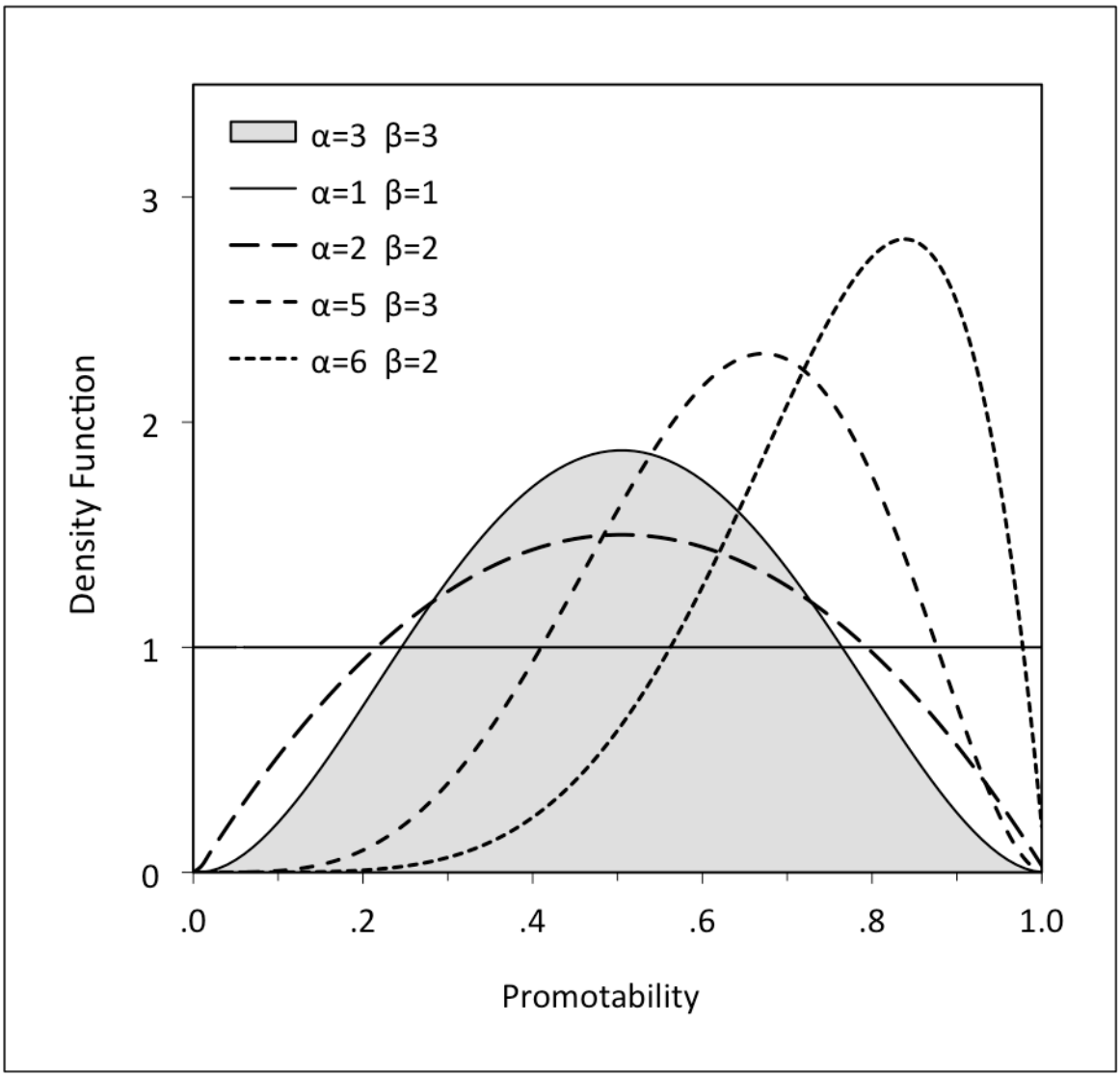

Figure 9. Beta densities with shape parameters (α, β) = (1,1), (2,2), (6,2), (5,3), and (3,3) are superimposed. For the density for Promotability of Entries in this study, we chose the beta (3,3), which is shaded for emphasis.

4.2 A Beta Distribution for the Promotability of Entries, *U*

For the distribution of Promotability for Entries, *U*, we selected $\alpha=\beta=3$ because the shape of the corresponding distribution is symmetric and somewhat similar to that of the familiar normal distribution, as well as other empirical distributions encountered in the workplace. More importantly, the choice of a symmetric distribution permits departures from symmetry induced by Skimming and Dredging to be easily perceived by the eye, as was illustrated in Figures 3–6. The density, mean, and standard deviation of *U* are $30u^2(1-u)^2$, $\mu_u = 1/2$, and $\sigma_u = 0.189$, respectively, where $0 \le u \le 1$.

4.3 The Distribution of Ratings, V

Michael and Schucany (2002) demonstrate that, given the beta RV, U, one can construct a second beta RV, V, as a binomial mixture of different betas based upon U,. Most importantly, one can precisely control the correlation coefficient between U and V through the choice of the binomial sample size, n. The formula for the correlation coefficient reduces to the simple expression $\rho=n/(n+\alpha+\beta)$. Inverting this formula we obtain $n=(\alpha+\beta)\rho/(1-\rho)$. When we set the parameters to the values of choice for this study, $\alpha=\beta=3$, and $\rho=0.60$, we obtain $n=9$.

Let K be a RV with probability function $h(k; n,u)$, where k is an integer outcome that ranges from 0 to n=9, and u is the probability of success. According to the theory of the Mixture Approach, the conditional density function of Ratings, V given $U=u$ is the following binomial mixture of beta RVs:

$$p(v|u; n,\alpha,\beta) = \sum_{k=0}^{n} h(k; n,u)\, e(v; \alpha+k, \beta+n-k). \tag{1}$$

The density in equation (1) is a mixture of 10 beta densities with weights equal to binomial probabilities that sum to 1.0. Therefore, if we replace the beta density, e, with the beta CDF, E, we obtain the conditional CDF of Ratings given Promotability $U=u$, as follows:

$$P(v|u; n,\alpha,\beta) = \sum_{k=0}^{9} h(k; n,u)\, E(v; \alpha+k, \beta+n-k). \tag{2}$$

The authors also show that, remarkably, the marginal distribution of V in equation (2) is identical to the marginal distribution of U, which is also Beta(3,3). As with U, the density, mean, and standard deviation of V are $30v^2(1-v)^2$, $\mu_v = 1/2$, and $\sigma_v = 0.189$, respectively.

4.4 Discretizing the Densities of the Promotability of Entries, *X*, and Ratings *Y*

The intractability of the incomplete Beta function in equation (2) creates difficulties when one attempts to calculate statistics, not only for specific values of Promotability, but summary statistics across the entire domain of Promotability, such as expected Tenure. We overcome this difficulty by approximating the continuous random variable U with its discrete analog X. We partitioned the unit interval into 100 intervals of width 0.01, and defined the domain, x, as the midpoints of these intervals: x_1=0.005, x_2=015,... , x_{100}=.095. We used the beta distribution for U to compute the probability that X falls in each interval. The precision of the approximation can be increased to the level the practitioner desires by increasing the number of intervals. The density, mean, and variance of X and Y reduce to $30x^2(1-x)^2$, $\mu_X = 1/2$, and $\sigma_x^2 = 1/28$, respectively.

4.5 Variables and Cutoff Scores

We found it helpful to classify variables and other measures into two categories: *Vector Variables* defined for the domain of *X*, and *Scalar Measures* defined for the Group as a whole. They are interdependent, so we introduce them first, and then discuss how they are connected.

Vector Variables. A Vector Variable is defined for all 100 points in the domain of *X*. We list all eight Vector Variables in Table 3.

Table 3. Vector Variables.
We maintained eight Vector Variables in an 8x100 matrix. Brief definitions are given here along with an exemplar for x=0.605 in a Group where the Staffing Strategy is Skimming. The Appendix presents derivations of the values of the exemplar.

Variable	Definition	Exemplar
x	Domain value	0.605
f_X	Pr(Promotability), Entries	0.017
g_X	Pr(Rating)	0.017
λ_{dx}	Demotion Rate	0.005
λ_{px}	Promotion Rate	0.295
λ_{ex}	Exit Rate	0.300
t_x	Expected Tenure, Entries	3.331
z_x	Pr(Promotability) Incumbents	0.011

Group Measures. Group Measures are scalars that apply to the Group as a whole. In Table 4, we list five selected Group measures by Staffing Strategy.

Table 4. Five Group Measures: definition and values listed by Staffing Strategy.

Measure	Description	Skimming	Trimming	Dredging
λ_d	Demotion Rate	0.050	0.100	0.150
λ_p	Promotion Rate	0.150	0.100	0.050
λ_e	Exit Rate	0.050	0.100	0.150
C_d	Promotion Cutoff	0.655	0.738	0.827
C_d	Demotion Cutoff	0.173	0.262	0.345

Transition Rates. Transition rates have a stochastic link to the Promotability of Entries X through annual supervisory Ratings Y. Supervisors use Cutoff scores C_d and C_d to select Incumbents for promotion and demotion. An Incumbent is promoted when $Y > C_d$ and demoted when $Y < C_d$. We can formulate these decisions as:

$$\begin{aligned} \lambda_d &= \text{Group Demotion Rate} &&= \Pr(Y < C_d), \\ \lambda_p &= \text{Group Promotion Rate} &&= \Pr(Y > C_d), \text{ and} \\ \lambda_e &= \text{Group Exit Rate} &&= \lambda_P + \lambda_d. \end{aligned} \tag{1}$$

Tenure. For this study, we assume that Tenure, T, has the exponential distribution; however, other distributions can be used to model Tenure, provided they have positive support. The PDF of the exponential distribution is $s(t) = \exp(-t/\lambda)/\lambda$, where λ is the transition rate and $1/\lambda$ is the mean. For the Group, the turnover rate for all three Staffing Strategies is 20%, which implies that average Tenure for the Group, λ_e, is 5 years.

Cutoff Scores. In our study, Group transition rates were specified in advance, but corresponding Cutoff scores were not. Among the Vector Variables in Table 3 are the transition rates λ_{dx}, λ_{px}, and λ_{ex}, defined for each of the 100 intervals. Each such rate depends upon the values of the Cutoff scores (See Appendix). We determined the values of the Cutoff scores by equating Group transition rates to the corresponding quantities we derived from the Vector Variables. For example, consider the Demotion rate for a Group where the Staffing Strategy is Skimming. The specified Group rate, $\lambda_d = 0.005$, must equal the expected value derived from the Vector Variables. We solved the equation $\lambda_d = 0.005 = \sum_x \lambda_{dx} f_x$, by varying C_d and C_d until equality was achieved. We refer the reader to the Appendix where we derive the values of the exemplar in Table 3 for all eight Vector Variables.

Derivation of the Promotability of Incumbents, Z. We make special note here of the derivation of the Promotability of Incumbents, Z, because it illustrates an important principle that we stress in this paper. In order to take LBS into account, one retains Entries as Incumbents in proportion to their expected length (Tenure) relative to that of the Group. At x=0.605, expected Tenure for an Entry is 3.331 years, compared to 5 years for the Group as a whole, the proportion being 0.666. Therefore, at x=0.605, where the Promotability of an Entry is 0.017, the Promotability of an Incumbent is two-thirds of this amount, or 0.011. One can observe this difference in Figure 10, where these two values have been added to the theoretical distribution of Promotability for Incumbents (shaded) in a Group where the Staffing Strategy is Skimming, and the promotion (demotion) rates are 10% (5%).

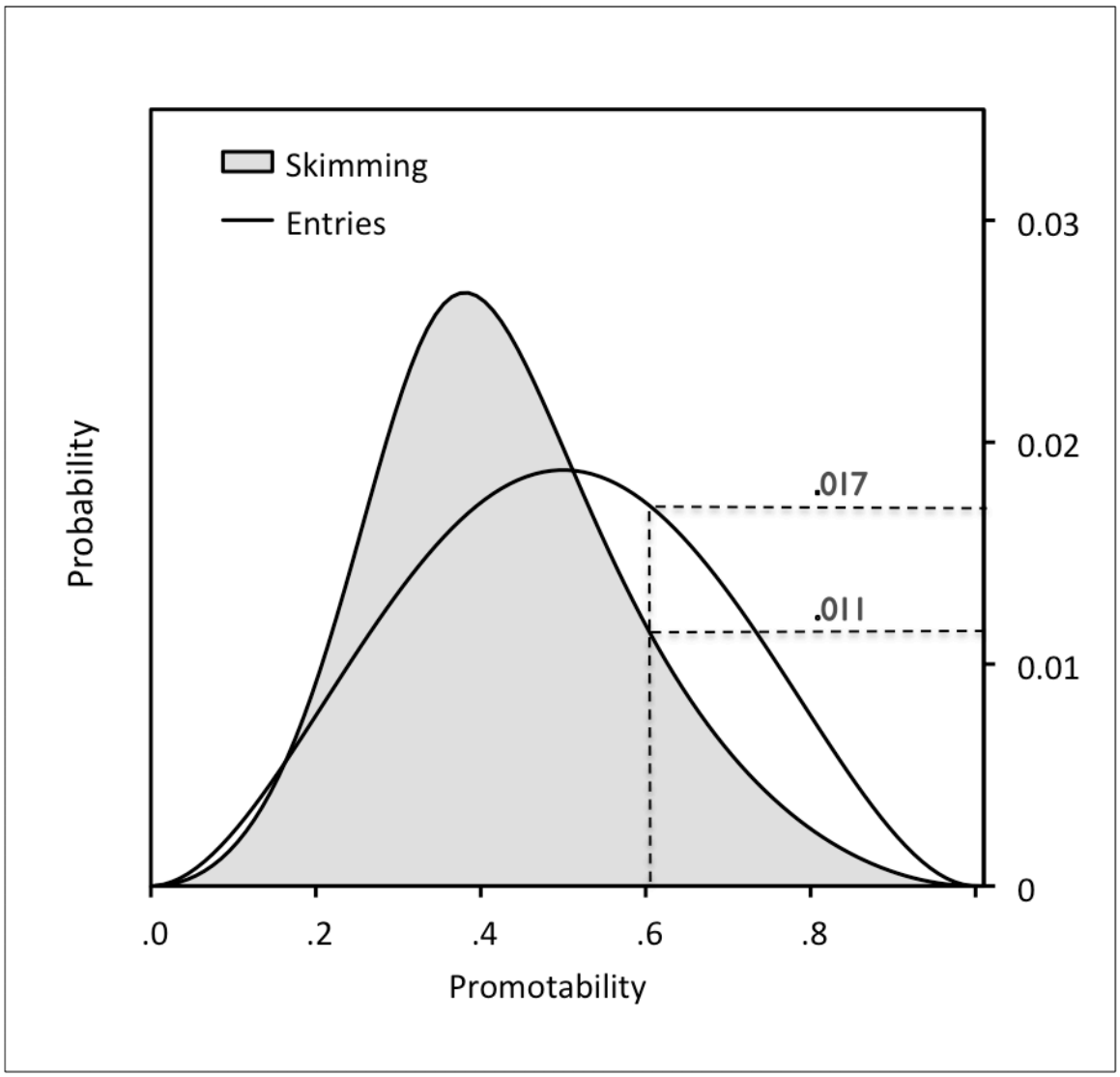

Figure 10. The effect of Skimming on the distribution of Promotability, with annotations for the exemplar *x*=0.605. The probability distribution of Promotability for Incumbents (shaded) is superimposed upon that for Entries to illustrate how Skimming depletes the upper tail of the distribution for Entries (as in Figure 3). Probabilities are annotated for the exemplar *x*=0.605, with the values of Promotability for an Entry, 0.017, and for an Incumbent, 0.011.

4.6 Expected Promotability given Current Tenure

We computed expected Promotability given current Tenure using Bayes Rule. For each of the 100 intervals, we computed the distribution of Tenure for each of the years *y*=0, 1, 2, ... 40. Then for each year, *y*, we weighted the expected values of *X* by the 100 probabilities for the distribution of Tenure. Finally, we divided the results by two, since we expect to intercept a completed Tenure at its midpoint.

5. DISCUSSION

5.1 All Organizations are Different

Personnel processes vary widely across organizations; just about every organization is unique in some respect. The model described herein is not expected to apply exactly to many organizations, or, for that matter, to any organization. The normal distribution is well known by practitioners; however, there are few, if any, distributions in practice that are precisely "normal." Box and Draper (1987) make the often-quoted statements: "Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful." (Box and Draper, 1987, p. 74.) So the question at hand, regarding the present paper, is not whether the model applies to a practitioner's data exactly, but can it be parameterized to be close enough to true model to be useful?

Our aim is to capture the essence and complexity of personnel processes; communicate these to the reader; familiarize the reader with certain important phenomena that are both prevalent, and often misunderstood, in workplace data, such as SSB; and to provide a fundamental statistical model as a "building block" that the practitioner can adapt to his or her needs. This importance of building blocks is espoused by I/O Psychologists. They assert how critical it is to have accurate estimates of the relationships among theoretical variables, because relationships at the latent-variable level are the building blocks for new and improved theory (Viswesvaran and Ones 1995; Hunter and Schmidt 2004; Schmidt, Shaffer and In-Sue 2008).

5.2 Parsimony

The statistical model presented herein is intentionally parsimonious. Almost all organizations differ with respect to their staffing strategies and histories, so most users will want to elaborate the fundamental model we present here. In practice, Performance may well increase over time with the acquisition of job knowledge. On the contrary, Performance might decrease as the incumbent's knowledge becomes obsolete. Rather than making specific assumptions about such relationships, we have assumed that Performance remains constant over time; we leave it to the user to "add in" adjustments that tailor the model to the user's organization, rather than assume the existence of factors that the user must first "subtract out" before proceeding. In summary, we adopted a policy of parsimony, i.e., minimizing complexity, when defining the model, so the user can adapt the model to the organization of interest by elaborating the fundamental statistical model. Furthermore, we urge the user to adopt the "Principle of Parsimony" (Ramsey and Schafer 2002), which is to employ the simplest model that adequately explains the data.

Schmidt and Hunter (2004) present a conventional path model of Performance that loads Tenure positively, under the assumption that job knowledge increases over time, and increased knowledge leads to increased Performance. Ratings are based, at least in part, upon Performance. However, they have additional error components, due to the variability of Performance, as well as variability associated with the Supervisor's perception of the incumbent's Performance.

Our hope is to raise the consciousness of the reader to factors that pose serious threats to validity, such as SSB, and inspire others to attempt to collect better personnel data, and to attain a better understanding of personnel decision processes. The cost of a single poor decision can be high. Accordingly, we urge practitioners to be mindful of the threats to validity presented herein and in (Shadish, Cook, and Campbell, 2002), and to improve and objectify the personnel data-collection and decision-making processes.

When developing the principles in this paper, we assumed that personnel transition processes have reached a steady state, so that historical records can provide reliable estimates of promotion and demotion rates, and transition processes have had sufficient time to operate Incumbents. If rates have varied substantially over time, the rate history is brief, the Group is expanding or contracting, or the Group is small, then the practitioner could attempt to adjust for such factors, or use it with caution.

Job requirements frequently change over time, which complicates the modeling of personnel processes. The needed job knowledge might be evolving in response to rapidly changing market demands. Recent entries might have training that is more relevant than some Incumbents with greater Tenure. Incumbents may have updated their skills. These are some of the factors that the analyst must attempt to take into account by elaborating the model.

In most organizations there will be other idiosyncrasies with regard to personnel decisions that the analyst should attempt to take into account. For example, there might be a waiting period of at least one year before an employee can be promoted; or promotion after one year's service might be automatic, barring unacceptable Performance.

5.3 Elaborating the Model

We use the bivariate beta to model the joint distribution between Promotability and Ratings. As we have made clear, we do not expect this model to apply directly to any specific organization. In our experience, some empirical distributions in organizations are approximately bell-shaped, like the beta (3,3). However, many are not, and many are asymmetric. Accordingly, we suggest that the practitioner consider the following procedure:

1. Transform the data for both Promotability and Ratings to the unit interval. We recommend a transformation that tends to symmetrize these distributions.

2. Fit beta distributions to the transformed data. One method for doing this is the method of moments (Johnson and Kotz 1970b, p47). Most modern statistical software packages have provisions for such estimation.

3. Analyze the data, as in the Methods section, and make inferences.

4. Finally, invert the transformations (from step 1) and express study inferences in terms of the original units of the data.

5.4 Exponential Lifetimes

Lifetimes (e.g., Tenures) can often be usefully represented with the exponential distribution (Johnson and Kotz 1970a, p208). The exponential distribution is easy to manipulate in mathematically, and can be a reasonable approximation for the true (complex) distribution of Tenure in many Job Groups. Although the statistical model presented here incorporates the exponential distribution, the fundamental approach is not limited to this distribution for Tenure. Rather the analyst can replace it with another distribution with positive support that better models the distribution of Tenure in the practitioner's Job Group.

5.5 Reason for Leaving

In our opinion, a serious limitation of our model is not endemic to the model itself, but a pervasive pattern we have noted in the coding and maintenance of human resource (HR) data. Generally, we have observed the absence of complete, accurate, and detailed records of personnel decisions. A prime example is "reason for leaving," which is important for the application of the Methods proposed in this paper. In many HR database systems, and probably all large organizations with a computerized HR database, a code is entered into the database that represents the reason for termination. In the authors' experiences, these codes are often missing, lack specificity, and are inaccurate, sometimes intentionally so. Suppose, for example, an employee is given the choice to resign or be fired, and chooses to resign. A "constructive discharge" such as this is often coded as a voluntary resignation rather than a "demotion." In such a case, the poor Performance associated with a forced resignation is not captured by the data and remains unknown to the analyst.

Consider the opposite situation, where an employee resigns to take a better job. Such a departure may be due to superior Performance, and should be classified as a promotion according the theory we present. But once again, it is not unusual that the reason for termination is coded as a simple resignation and the superior Performance of the employee is not captured by the HR database. The miscoding of reasons for leaving is common in the workplace largely because of the difficulty in doing so, i.e. the subjective judgments that must be made are hard. Our recommendation to human resource managers is to invest the time and effort it takes to create and enter detailed and defensible reasons for termination, and enter codes for these reasons into the HR database. This will require a well-documented, and labor-intensive effort to assign reasons for termination that can then be associated with Performance. Such information should have great value to the analysis we propose herein, and will provide a better understanding of the personnel dynamics in the organization. Otherwise, if the data are unreliable, then so are the inferences that can be drawn from such data.

5.6 Incorporating Transfers

We define transfers as those job changes that are unrelated to job performance. We recommend that transfers be removed from the analysis data set to be analyzed. In this context, transfers include conventional lateral transfers as well as departures believed to be unrelated to Performance. The latter departures include resignations because of relocation, staying at home to take care of family members, illness, and even death. When the reason for departure is unknown, or cannot be associated with Performance, we recommend that these cases be omitted from the analysis and that the practitioner draw inferences only from the data considered to be reliable.

5.7 How to Use of a Snapshot Sample

We have taken aim at snapshot samples and the perils of SSB. A natural question to ask is: how can one make good use of a snapshot sample? If the purpose of the analysis is *descriptive*, then a Snapshot might be the type of sample needed. For example, suppose an employer wishes to compare job satisfaction to that in comparable organizations, or to external benchmarks for the industry. In such cases, a Snapshot might be the appropriate sample for such comparative purposes.

On the other hand, suppose the purpose of the study is *explanatory*, and that the goal is to identify factors that predict job satisfaction (or dissatisfaction). For an analytical study such as this, it is important that the sample be representative of all Entries for the period of interest. It may be difficult if not impossible to survey employees who have departed the Group, and whose opinions might be critical. A validity study conducted on a sample in which only Current employees can be sampled is referred to as a concurrent validity study. I/O psychologists are keenly aware of the threat to validity posed by concurrent validity studies and have developed adjustment factors to take potential biases into account. The present authors are reminded of a corporate meeting during which the CEO noted how low the attendance was and asked if anyone in the audience knew why. A venerable survey statistician in the audience gave a reply, which brought down the house: "Why don't we hear from the people who aren't here?"

5.8 Age and Ability

The courts make the assumption that, in the absence of age discrimination, there should not be a correlation between age and ability, or between age and variables designed to measure ability. This is reasonably consistent with the bulk of psychological research for the general population (Sterns and McDaniel 1994, Warr 1994, Beier and Ackerman 2005). We make the same assumption in this study. However, as is demonstrated in this paper, this assumption is not necessarily true for a snapshot sample. When applied to incumbents in a Group, the courts' assumption might well be wrong. SSB can easily bias the observed correlation between age and Performance related variables such as Ratings and layoffs. Skimming can create the false impression of age discrimination. Conversely, Dredging can mask the effects of true age discrimination.

Age Stereotyping. SSB is a phenomenon that has the potential of contributing to the formation of erroneous age stereotypes. Consider the following scenario: a Skimming Strategy is in full force and induces an association between age and ability. Employees in the Group are likely to perceive this association. Suppose the boundaries of the Group have effectively become the boundaries of their world of work. In such a case, these employees might extrapolate inferences based upon what they observe in their Group to the workplace as a whole, and erroneously infer that age is associated with a decline in ability. Now consider an alternative scenario: what would happen if Incumbents had broader exposure to Incumbents in Groups ranked higher in the job hierarchy? What these workers might well observe is that many of the older workers have been promoted, now reside in high places, and are running the company.

The Legal Challenge. The legal challenge is to get the attention of the courts and attorneys with regard to the complex relationships among Ability, Performance, and Ratings. One hurdle is conveying these principles to legal professionals, such as lawyers and judges. Another is communicating the results to laypersons, such as members of a jury. In the absence of a careful assessment of the possibility of SSB, we suggest that the finder of fact be circumspect about the strength of the statistical evidence.

6. CONCLUSION

We proposed theoretical models for analyzing Snapshots of job Groups in the workplace. This has long been the purview of social scientists, notably Industrial/Organizational Psychologists, who have advanced the theory of job performance in organizations and meta-analytic methods. We defined three stereotypical Staffing Strategies – Skimming, Dredging, and Trimming – which are based upon the relative magnitudes of promotion and demotion rates, and we illustrated their calculated effects. We emphasized the distinction between a Snapshot sample of Incumbents and a representative sample of Entries. The threat to validity posed by confusing the two is so common and so perilous that we give it a name: Snapshot Selection Bias. We discussed the consequences of this bias for analytic studies of personnel data in general and age discrimination in particular.

REFERENCES

Beier, M. E. and Ackerman, P. A. (2005), "Age, Ability, and the Role of Prior Knowledge on the Acquisition of New Domain Knowledge: Promising Results in a Real-World Learning Environment," Psychology and Aging, 20, 341–355.

Box, G. E. P. and Draper, N. R. (1987), *Empirical Model-Building and Response Surfaces.* New York: John Wiley and Sons.

Cook T. D., Campbell, D. T. and Peracchio, L. (1990), "Quasi experimentation," In Dunnette, M. D. and Hough, L. M. (Eds.), *Handbook of Industrial and Organizational Psychology* (pp. 491-576), Second Edition, Volume 1, Palo Alto: Consulting Psychologists Press.

Digman J.M. (1990). "Personality structure: Emergence of the five-factor model," *Annual Review of Psychology*, 41, 417–440.

Feller, W. (1970), *An Introduction to Probability Theory and Its Applications, Volume II* (2nd ed.), New York: John Wiley and Sons, Inc.

Hunter, J. E. (1980). Test validation for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB). Washington, DC: U.S. Department of Labor.

Hunter, J. E., & Hunter, R. F. (1984). "Validity and utility of alternate predictors of job performance," *Psychological Bulletin*, 96, 72–98.

Hunter, J. E. and Schmidt F. L. (2004), *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2nd ed.), Thousand Oaks, CA: Sage.

Johnson, N. L. and Kotz, S. (1970a), *Distributions in Statistics: Continuous Univariate Distributions–1*, New York: Wiley.

Johnson, N. L. and Kotz, S. (1970b), *Distributions in Statistics: Continuous Univariate Distributions –2*, New York: Wiley.

Michael, J. R. (1999), "The Flow of Ability Through a Job Group," unpublished Ph.D. dissertation, The George Washington University.

Michael, J. R. and Schucany, W. R. (2002), "The mixture approach for simulating bivariate distributions with specified correlations," *The American Statistician*, 56, 48–54.

Mobley, W. H. (1991), "Intermediate Linkages in the Relationship between Job Satisfaction and Employee Turnover," In *Motivation and Work Behavior*, Steers, R. M. and Porter, L. W. (Ed.), New York: McGraw-Hill, Inc.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). "Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance," *Journal of Applied Psychology,* 78*,* 679–703.

Peter, L. J., and Hull, R. (1969), *The Peter Principle*, New York: Bantam Books.

Ramsey, F. L. and Schafer, D. W. (2002). *The Statistical Sleuth: A Course in Methods of Data Analysis*. Duxbury: Pacific Grove, CA.

Rushton, J. Philippe and Jensen, Arthur R. (2005), "Thirty years of research on race differences in cognitive ability," Psychology, Public Policy, and Law, Vol. 11, No. 2, 235–294.

Schmidt, F. L., and Hunter, J. E. (1998), "The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings," *Psychological Bulletin*, 124, 262–274.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1981). "Task differences and validity of aptitude tests in selection: A red herring." *Journal of Applied Psychology*, 66, 166–185.

Schmidt, F. L., Law, K., Hunter, J. E., Rothstein, H. R., Pearlman, K., & McDaniel, M. (1993), "Refinements in validity generalization methods: Implications for the situational specificity hypothesis," Journal of Applied Psychology, 78, 3–13.

Schmidt, F. L, Le, Huy. and Oh, I. (2009), "Correcting for the Distorted Effects of Study Artifacts in Meta-Analysis," In H. H. Cooper, L. V., Hedges, and Valentine, J. C. (Eds.), Chapter 17 in, *The Handbook of Research Synthesis and Meta-Analysis* (2nd ed.), New York, NY: Russell Sage Foundation Publications

Schmidt, F, L, Shaffer, J.A. and In-Sue, O.H. (2008), "Increased accuracy for range restriction corrections: implications for the role of personality and general mental ability In job and training performance," *Personnel Psychology*, 61, 827–868.

Shadish, W. R., Cook, T. D. and Campbell D. T. (2002), *Experimental and Quasi-Experimental Design, for Generalized and Causal Inference*, Dallas: Houghton Mifflin.

Snedecor, G.W. and Cochran, W.G. (1967), *Statistical Methods*, Ames Iowa: The Iowa State University Press

Spearman, C. (1904), "General intelligence, objectively determined and measured," *American Journal of Psychology*, 15, 201–293.

Sterns, H. L. and McDaniel, M. A. (1994), "Job performance and the older worker," In Rix, Sara E. (Ed.), *Older Workers: How Do They Measure Up? An Overview of Age Differences in Employee Costs and Performance* (pp. 27-51), Washington, DC: AARP.

The American Statistical Association (1999) *Ethical Guidelines for Statistical Practice*, The American Statistical Association, http://www.amstat.org/about/ethicalguidelines.cfm

The American Statistical Association (2011), Surveys, http://www.amstat.org/careers/surveys.cfm, The American Statistical Association/Washington, DC.

Vardi, Y. (1982), "Nonparametric Estimation in the Presence of Length Bias," *The Annals of Statistics*, 10, 616-620.

Vardi, Y. (1988), " Statistical Models for Intercepted Data ," *Journal of the American Statistical Association*, 83, No. 401, 616-620.

Viswesvaran C, Ones D., S. (1995). “Theory testing: Combining psychometric meta-analysis and structural equations modeling,” *Personnel Psychology*, *48*, 865–885.

Warr, P. (1994). “Age and employment,” In H. C. Triandis, M. D. Dunnette,& L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 4, pp. 485–550). Palo Alto, CA: Consulting Psychologists Press.

APPENDIX: DERIVATION OF VECTOR VARIABLES FOR THE EXEMPLAR x=0.605

We present here the derivations of the Vector Variables listed in Table 3 and illustrate them with the exemplar x=0.605.

$$
\begin{aligned}
x &= 0.605 \\
f_x &= \Pr(0.600<X<0.610) \\
&= F(0.610) - F(0.600) \\
&= 0.700 - 0.683 = 0.017 \\
g_x &= f_x = 0.017 \\
\lambda_{dx} &= \Pr(Y<C_d) \\
&= \Pr(Y<0.017) \\
&= G_y(C_d = 0.017;\ \alpha=\beta=3,\ x=0.605) \\
&= \sum_{k=0}^{9} h(k;\ n,x)\ E(v;\ \alpha+k,\ \beta+n-k) \\
&= \sum_{k=0}^{9} h(k;\ n,x)\ E(0.017;\ 3+k,\ 12-k) \\
&= h(0;\ 9,x)\ E(0.017;\ 3,\ 12) + h(1;\ 9,x)\ E(0.017;\ 4,\ 11) + \\
&\quad h(2;\ 9,x)\ E(0.017;\ 5,\ 10) + h(3;\ 9,x)\ E(0.017;\ 6,\ 9) + \\
&\quad h(4;\ 9,x)\ E(0.017;\ 7,\ 8) + h(5;\ 9,x)\ E(0.017;\ 8,\ 7) + \\
&\quad h(6;\ 9,x)\ E(0.017;\ 9,\ 6) + h(7;\ 9,x)\ E(0.017;\ 10,\ 5) + \\
&\quad h(8;\ 9,x)\ E(0.017;\ 11,\ 4) + h(9;\ 9,x)\ E(0.017;\ 12,\ 3) \\
&= (0.000)\ (0.448) + (0.001)\ (0.215) + (0.020)\ (0.080) + \\
&\quad (0.071)\ (0.023) + (0.162)\ (0.023) + (0.249)\ (0.005) + \\
&\quad (0.254)\ (0.001) + (0.167)\ (0.000) + (0.064)\ (0.000) + \\
&\quad (0.011)\ (0.000) \\
&= 0.005. \\
\lambda_{px} &= 0.295 \text{ (derived similarly)} \\
\lambda_{ex} &= \lambda_{dx} + \lambda_{px} \\
&= 0.300 \\
t_x &= 1/\lambda_{ex} \\
&= 3.331 \\
z_x &= (t_x/5)\ f_x \\
&= (0.666)\ (0.017) \\
&= 0.011
\end{aligned}
$$

In the last equation, the derivation of z_x takes LBS into account by selecting (retaining, over time) Entries in proportion to their expected lengths (Tenures). Thus, at x=0.605, the expected Promotability of an Incumbent (0.011) is two-thirds that of an Entry (0.017).