

Lazear, Edward P.

Working Paper

The Peter Principle: A Theory of Decline

IZA Discussion Papers, No. 759

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Lazear, Edward P. (2003) : The Peter Principle: A Theory of Decline, IZA Discussion Papers, No. 759, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/21410>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 759

The Peter Principle: A Theory of Decline

Edward P. Lazear

April 2003

The Peter Principle: A Theory of Decline

Edward P. Lazear
*Stanford University
and IZA Bonn*

Discussion Paper No. 759
April 2003

IZA

P.O. Box 7240
D-53072 Bonn
Germany

Tel.: +49-228-3894-0
Fax: +49-228-3894-210
Email: iza@iza.org

This Discussion Paper is issued within the framework of IZA's research area *Mobility and Flexibility of Labor*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor, (2) internationalization of labor markets, (3) welfare state and labor market, (4) labor markets in transition countries, (5) the future of labor, (6) evaluation of labor market policies and projects and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website (www.iza.org) or directly from the author.

This paper is forthcoming in a Special Issue of the Journal of Political Economy in Memory of Sherwin Rosen. Sherwin Rosen was my most important teacher, my valued colleague and dear friend. Sherwin served on my thesis committee and taught me much of what I know. Throughout the thirty years that we were friends, Sherwin was a constant source of inspiration, wisdom, and kindness. A deep thinker who opened up a number of areas of research, Sherwin was interested in hierarchies and promotion, so this paper is very much in keeping with his research agenda and derives from it.

ABSTRACT

The Peter Principle: A Theory of Decline*

Some have observed that individuals perform worse after being promoted. The Peter Principle, which states that people are promoted to their level of incompetence, suggests that something is fundamentally misaligned in the promotion process. This view is unnecessary and inconsistent with the data. Below, it is argued that ability appears lower after promotion purely as a statistical matter. Being promoted is evidence that a standard has been met. Regression to the mean implies that future ability will be lower, on average. Firms optimally account for the regression bias in making promotion decisions, but the effect is never eliminated. Rather than evidence of a mistake, the Peter Principle is a necessary consequence of any promotion rule. Furthermore, firms that take it into account appropriately adopt an optimal strategy. Usually, firms inflate the promotion criterion to offset the Peter Principle effect, and the more important is the transitory component relative to total variation in ability, the larger the amount that the standard is inflated. The same logic applies to other situations. For example, it explains why movie sequels are worse than the original film on which they are based and why second visits to restaurants are less rewarding than the first.

JEL Classification: J00, J6

Keywords: Peter principle, regression to the mean, stochastic

Edward P. Lazear
Graduate School of Business
Stanford University
CA 94305-6010
USA
Email: lazear@stanford.edu

* This research was supported in part by the National Science Foundation. Useful comments were provided by many. Po-Han Fong, Thomas P. Gherig, Kevin J. Murphy, Torsten Persson, Paul Pfleiderer, Kathryn Shaw, Eskil Wadensjo, and Michael Waldman were particularly helpful. An earlier version of this paper was printed as an NBER working paper.

The Peter Principle states that workers are promoted to their level of incompetence.¹ One interpretation is that firms systematically make mistakes in their promotion decisions. Another, favored here, is that the decline in ability that is seen after promotion is the natural outcome of a statistical process that displays regression to the mean. Workers are promoted on the basis of having met some standard. To the extent that ability is the sum of both permanent and transitory components, those who meet the standard have expected transitory components that are positive. The expectation of the post-promotion transitory component is zero, implying a reduction in expected ability. Firms that understand the statistical process take this phenomenon into account by adjusting the promotion standard, but the result remains: Expected ability for those promoted is lower after promotion than before.²

There is substantial evidence of the Peter Principle.³ In addition to papers from the

¹Peter and Hull (1969).

²It is also true that those who are denied promotion do better after they are turned down than they did before the decision was made, for the same reason.

³Two recent papers [Fairburn and Malcolmson (2000) and Faria (2000)] on this topic use a very different approach from this paper and from each other. The Peter Principle is a by-product of using promotion to solve a moral hazard problem in Fairburn and Malcolmson. Rather than motivate through money, which induces influence activity, firms choose promotion because then managers must live with the consequences of their decisions. Too many workers are promoted under certain circumstances, resulting in a Peter Principle effect. In Faria (2000), workers have two skills. Those who are good at one are necessarily less good at another when on the frontier. Faria argues that this is what is meant by the Peter Principle.

marketing and organizational behavior literature,⁴ there are a number of findings in empirical labor economics that support the claim. In an early paper that used subjective performance evaluation, Abraham and Medoff (1980) reported that workers' subjective evaluation scores fell, the longer they were on the job. In Lazear (1992), it was found that the coefficient of job tenure in a wage regression was actually negative. The longer a worker was in a particular job, given his tenure in the firm, the lower his wage. The reason presumably is that the better workers are promoted out of the job so those with a given number of years of firm experience who have fewer years in a job are less likely to have gotten stuck in that job. Baker, Gibbs and Holmstrom (1994) replicate this finding in their data and Gibbs and Hendricks (2001) find that raises and bonuses fall with tenure.

The tone of the literature outside of economics is that there is something wrong with promotion dynamics, and this anomaly shows up as the Peter Principle. (Indeed, the book written by Peter and Hull is entitled, *The Peter Principle: Why Things Always Go Wrong*.) The approach taken here is different. The Peter Principle results from optimal adjustment to decision-making under uncertainty. It is argued that even when firms use exactly the right promotion rule, the Peter Principle effect will be observed. The fact that reversals of promotion decisions are rare is more compatible with the view that they were correct in the first place than that something went wrong.

More often, the Peter Principle is interpreted in a multi-factor context. Individuals who are good in one job are not necessarily good in the job into which they are promoted. As a result, individuals appear incompetent in the job in which they settle. To obtain this result, it is merely necessary to make a slight modification in the regression-to-the-mean structure. Here, general

⁴See, for example, Anderson, Dubinsky, and Mehta (1999).

ability is combined with a job-specific ability to produce output. Regression to the mean results because positive readings on the job-specific component prior to promotion are uncorrelated with the job-specific component after promotion.⁵

The fact that promoted individuals are less able than their apparent pre-promotion ability induces firms to adjust in two respects. First, firms select their promotion rule with the understanding that the pre-promotion ability is a biased estimate of true ability for those who exceed some standard. Second, as the variance in the transitory component of ability rises relative to the variance in the permanent component, the adjustment factor becomes greater. In the typical case, the standard that one must exceed to obtain a promotion increases with the relative importance of the transitory component because the regression effect is larger when the variance of the transitory component is large.

The model presented below yields the following results:

1. Promoted individuals' performance falls, on average, relative to their pre-promotion performance.
2. Firms that take the decline into account adjust their promotion rule accordingly, but this does not negate the observation that ability declines after promotion.
3. The importance of the Peter Principle depends on the amount of variation in the transitory component relative to the permanent. The Peter Principle is most pronounced when the transitory component is large.

⁵The structure is a variant of the Jovanovic (1979a,b) model that was modified and used in a context closer to this structure in Lazear, (1986).

4. The length of the pre-promotion period depends on the ratio of transitory variation to permanent variation. As the transitory component becomes more important, firms lengthen the pre-promotion period.
5. Movie sequels are systematically worse than the original on which they are based.
6. Follow-up visits to good restaurants provide poorer meals than the first sampling.
7. Absent learning effects, second term elected officials are less effective than they were during the first term.

A. Model

Let there be two periods. Each worker has a time-invariant component of ability, denoted $A \sim f(A)$, and a time-varying component of ability, denoted ε_1 for period 1 and ε_2 for period 2. Let the time-varying components be i.i.d. with density $g(\varepsilon)$. The firm can observe $A + \varepsilon_t$ in each period, but cannot disentangle the time-varying component of ability from the permanent component. There are a variety of interpretations that are consistent with this specification. One can think of the ε_t as being a true transitory aspect of ability or just measurement error. Later, the interpretation of different jobs will be considered.

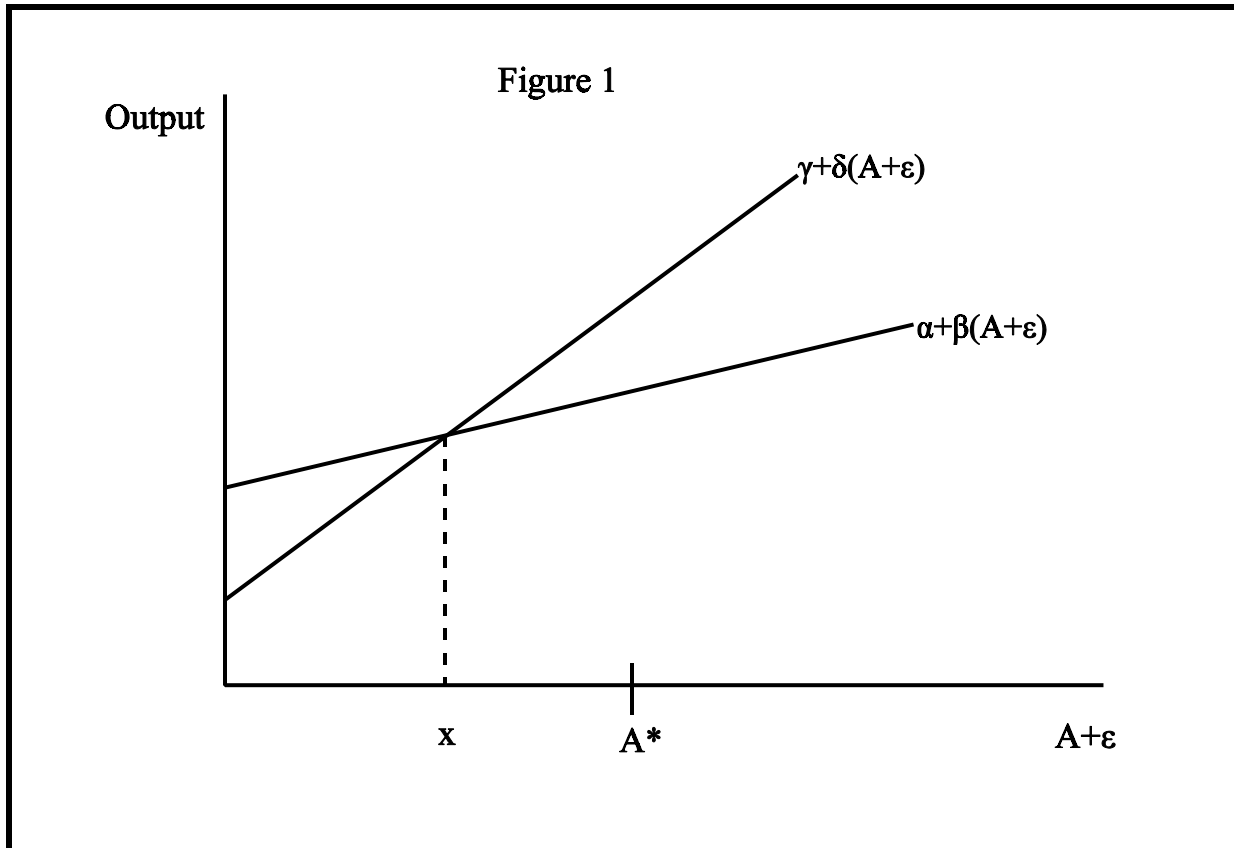
There are two jobs (two are sufficient), which we denote difficult and easy. An individual's productivity in the easy job is given by

$$\alpha + \beta(A + \varepsilon_t)$$

and in the difficult job is given by

$$\gamma + \delta(A + \varepsilon_t)$$

where $\alpha > \gamma$ and $\delta > \beta$. Thus, it pays to assign a worker to the difficult job if and only if



$$A+\epsilon_t > x$$

where

$$x = (\alpha - \gamma) / (\delta - \beta).$$

The situation and the crossing point that correspond to x are shown in figure 1.⁶ The setup seeks to capture the idea that the most able have a comparative advantage in the difficult job.

Assume that individual ability $f(A)$ is such that in the absence of information, it pays to

⁶This production structure is similar to that used in a comprehensive analysis by Gibbons and Waldman (1999), who also allow for transitory and permanent components with regression. The focus of their paper is earnings and promotion. Neither optimal decision-making by firms given the transitory component, nor strategic effort in response to promotion rules are central to their discussion.

assign everyone to the easy job in period 1.⁷ Intuitively, this assumption amounts to saying that most people are not well-suited to the difficult and that in the absence of countervailing information, individuals are assigned to the easy job. With one exception, noted in the appendix, it is sufficient to assume symmetric ignorance, so that workers are no better informed about their abilities than firms.

After the first period, firms obtain an estimate of A , namely $\hat{A} = A + \varepsilon_1$. Since ε_1 is the period one transitory component (either measurement error or transitory ability), it is A and not \hat{A} on which a promotion decision should be made. But A is not observed, so firms are forced to base their decision on \hat{A} .

1. Workers perform worse after being promoted

Firms must select some criterion level, A^* , such that if $\hat{A} > A^*$ the worker is promoted to the difficult job. If \hat{A} is less than A^* , the worker remains in his current job. It is now shown that workers who are promoted have levels of ability in period 1 that are higher on average than their ability in period 2.

First, note that the expectation of ε_1 given that an individual is promoted, is

⁷This amounts to assuming that $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\alpha + \beta(A + \varepsilon)) dGdF > \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\gamma + \delta(A + \varepsilon)) dGdF$.

$$\begin{aligned}
E(\varepsilon_1 | A + \varepsilon_1 > A^*) &= \int_{-\infty}^{\infty} \int_{A^*-A}^{\infty} \frac{1}{1 - G(A^* - A)} \varepsilon g(\varepsilon) f(A) d\varepsilon dA \\
&= \int_{-\infty}^{\infty} E(\varepsilon | \varepsilon > A^* - A) f(A) dA
\end{aligned}$$

which is positive since $f(A)$ is positive and the conditional expectation of ε given ε greater than any number is positive (because the unconditional expectation of ε is zero).

Thus, the conditional expectation of ε_1 is positive among those who are promoted.

Now, in period 2, the expectation of the transitory component is

$$E(\varepsilon_2 | A + \varepsilon_1 > A^*) = 0$$

because ε_2 is independent of A and of ε_1 . As a result, for any promoted individual with ability A ,

$$A + E(\varepsilon_1 | A + \varepsilon_1 > A^*) > A + E(\varepsilon_2 | A + \varepsilon_1 > A^*).$$

Thus, expected ability falls for promoted individuals from period 1 to period 2.

Individuals who are promoted are promoted in part because they are likely to have high permanent ability,⁸ but also because the transitory component of their ability is high. One of the

⁸The notation ε_1 and A could be swapped in the above discussion to show that $E(A | A + \varepsilon_1 > A^*) > E(A)$.

reasons that academics tend to write better papers before they receive tenure is that they would not have received tenure had they not written the better-than-average papers. The point is obvious, but is made graphic by the following example. Suppose that a firm promotes all individuals who can obtain three heads on three consecutive coin tosses. Only one in eight will be promoted. But when the firm asks their promoted individuals to repeat the feat, only one in eight will measure up. Seven out of eight will do worse than they did before being promoted. The reason is that all of the “performance” on the coin toss is transitory since tosses are independent.

As a general matter, the larger is the transitory component relative to the permanent component, the more important is the Peter Principle effect. If there were no transitory component, there would be no regression to the mean. Thus, the importance of “luck” is positively associated with the force of the Peter Principle.

2. The Promotion Rule

Firms know in advance that there will be some expected fall in productivity among those promoted and adjust their promotion standard accordingly. Below, the general optimization problem for the firm is presented. Then, the way in which the rule operates is demonstrated by an example.

The firm’s problem is to maximize profits (or worker utility) by selecting the job for each candidate with the highest expected value. Recall that individuals who have period 2 ability greater than x , defined above, would be assigned to job 2 were second period ability known. The firm does not see A , but only \hat{A} and must choose some criterion, A^* , such that it promotes workers whose observed ability in period 1 is greater than A^* . This is equivalent to promoting individuals when $A > A^* - \varepsilon_1$. Thus, the firm wants to choose A^* so as to maximize

$$(1) \quad \underset{A^*}{\text{Max}} \int_{-\infty}^{\infty} \int_{A^*-\varepsilon_1}^{\infty} \int_{-\infty}^{\infty} (\gamma + \delta(A + \varepsilon_2)) f(A) g(\varepsilon_1) g(\varepsilon_2) d\varepsilon_2 dA d\varepsilon_1$$

$$+ \int_{-\infty}^{\infty} \int_{-\infty}^{A^*-\varepsilon_1} \int_{-\infty}^{\infty} (\alpha + \beta(A + \varepsilon_2)) f(A) g(\varepsilon_1) g(\varepsilon_2) d\varepsilon_2 dA d\varepsilon_1$$

Because the expectation of ε_2 is zero, (1) can be written as

$$(2) \quad \underset{A^*}{\text{Max}} \int_{-\infty}^{\infty} \int_{A^*-\varepsilon_1}^{\infty} (\gamma + \delta A) dFdG + \int_{-\infty}^{\infty} \int_{-\infty}^{A^*-\varepsilon_1} (\alpha + \beta A) dFdG$$

The choice of A^* depends on the distribution. However, two examples reveal that A^* does not equal x as a general rule. Instead, in typical cases, firms adjust A^* upward. Knowing that worker ability in period 2 will differ from worker ability in period 1, firms usually set the bar higher than they would were ability observed in period 1 carried over directly to period 2.

Actual solutions that provide intuition are available for given distributions. Consider, for example, the case where A , ε_1 , and ε_2 are all distributed normally, with mean zero and variance equal to 1. Let $\alpha=1$, $\beta=.5$, $\gamma=0$, $\delta=1$. Then x , the ability level at which jobs produce equal value, is 2 since

$$\alpha + \beta(A+\varepsilon) = \gamma + \delta(A+\varepsilon)$$

for $A+\varepsilon = 2$. However, A^* is 4.01. The firm sets its promotion standard more than two standard deviations higher than the crossing point in figure 1 because it understands that the worker's ability

in period 2 is likely to be lower than it was in period 1 for the promoted group. As a result, the firm insists on a very high level of observed ability in period 1 in order to warrant promotion. Statements like, “tenure requires that the faculty member be the best in his field, having produced outstanding research” is a manifestation of the upward adjustment.

Consider the same example with a twist. Let the distribution of A remain the same, namely, normal with mean 0 and standard deviation of 1, but let the standard deviation of ε_1 fall to 0.1. Then, A^* drops from 4.01 to 2.08. Although the firm still adjusts its promotion criterion upward from x , the adjustment is much smaller because the importance of the transitory component has been diminished. There is regression to the mean, but the regression that takes place is small relative to the amount in the prior example. When the standard deviation of ε is zero, the promotion standard is 2, which is exactly x as expected. Then, the distribution of ε_1 is degenerate, so that all observed ability in period 1 is permanent ability. The problem in (2) becomes

$$\text{Max}_{A^*} \int_{A^*}^{\infty} (\gamma + \delta A) f(A) dA + \int_{-\infty}^{A^*} (\alpha + \beta A) f(A) dA$$

which has first order condition

$$\frac{\partial}{\partial A^*} = -(\gamma + \delta A^*) f(A^*) + (\alpha + \beta A^*) f(A^*) = 0 \quad .$$

The solution is

$$(\gamma + \delta A^*) = (\alpha + \beta A^*)$$

which is the crossing point, i.e., x , in figure 1. When there is no transitory component, the firm simply promotes those whose permanent ability places them better in the difficult job than in the easy job.

It is possible to derive the relation between A^* and x in more general terms.⁹ The first-order condition to (2) is

$$(\gamma + \delta A^* - \alpha - \beta A^*) \int_{-\infty}^{\infty} f(A^* - \varepsilon_1) g(\varepsilon_1) d\varepsilon_1 = (\delta - \beta) \int_{-\infty}^{\infty} \varepsilon_1 f(A^* - \varepsilon_1) g(\varepsilon_1) d\varepsilon_1$$

The integral on the l.h.s. is always positive, so the whether A^* exceeds x or not depends on the sign of the integral on the r.h.s. Assume that both $f()$ and $g()$ are symmetric densities and let $g()$ be symmetric around zero and $f()$ be symmetric around \bar{A} . Write the integral on the right side as

$$\int_{-\infty}^0 \varepsilon_1 f(A^* - \varepsilon_1) g(\varepsilon_1) d\varepsilon_1 + \int_0^{\infty} \varepsilon_1 f(A^* - \varepsilon_1) g(\varepsilon_1) d\varepsilon_1$$

Use a change of variable in the first integral of $u = -\varepsilon_1$ and in the second, allow $u = \varepsilon_1$. Because of symmetry, $g(u) = g(-u)$ so one can write the two integrals as one:

⁹I am indebted to Wing Suen for this derivation.

$$\int_0^{\infty} u[f(A^*-u) - f(A^*+u)]g(u)du$$

Suppose that the firm wants to promote fewer than 50% of the people, i.e., that $x > \bar{A}$. (Recall that x is the value such that $\alpha + \beta x = \gamma + \delta x$. If $x = \bar{A}$, then because of symmetry of the density functions, half of the population would have $A > x$ and half would have $A < x$.) Under such circumstances, $f(x-u) > f(x+u)$ (because f is unimodal around \bar{A}). Thus, the r.h.s. of the first-order condition is positive for $A^* = x$, which implies that

$$\alpha + \beta A^* > \gamma + \delta A^*$$

or that $A^* > x$. Thus, the firm adjusts the cutoff level upward when fewer than 50% of the workers are better suited to the difficult job than the easy job.

The same reasoning applies in reverse. If more than 50% are to be promoted, then $x < \bar{A}$ which means that $f(x-u) < f(x+u)$. As a result, the r.h.s. of the first-order condition is negative at x , which means that A^* must be less than x to satisfy optimality. Thus, when more than 50% better suited to the difficult job than the easy job, the firm reduces the promotion cutoff below that which would be optimal were there no error in period one. (Actually, under such circumstances, workers would initially be assigned to the difficult job and the standard would be one such that workers who fell below it would be demoted after the first period.)

The intuition is this. Although there is always regression to the mean, adjusting the promotion level upward reduces the probability that the firm will make a bad promotion decision.

However, at the same time, the adjustment increases the probability that it will fail to promote a qualified worker i.e., it reduces the false positive while increasing the false negative error. Conversely, lowering the promotion cutoff reduces the probability that someone who erroneously was observed to be a poor worker is not promoted, but increases the probability that the firm promotes too many bad workers. Thus, there is a tradeoff. When fewer than 50% are to be promoted, the expected cost of making a false positive error exceeds that of making a false negative error so that the criterion must be adjusted upward. To the extent that most hierarchies are narrower at the top than at the bottom, $A^* > x$ is probably the more typical case, So $A^* > x$. Standards are adjusted upward.

3. Another Interpretation of the Peter Principle

Rosen (1986) presents a model of sequential promotions where individuals are sorted by ability at each stage such that the entering class at each round of a tournament are of equal (ex ante) ability. Rosen uses the model to determine the optimal compensation at each level to motivate workers. Sorting is also an issue because the pool of workers at successive rounds have higher ability than earlier rounds. The Rosen model is in some ways more general than the one described here because it allows for effort as well as ability differences. The focus is quite different, however, because neither the optimal promotion rule nor the worker's output over time is an important part of the analysis. It is likely that many of the Peter Principle results that come out of this paper could have been derived in that important paper on sequential promotions.¹⁰

¹⁰Gibbons and Waldman (1999b) and Prendergast (1999) survey the large literature on careers and promotions.

Still, the Rosen model does not fit one of the most common interpretations of the Peter Principle, which is that workers are promoted to their level of incompetence because a worker who is good in one job is not necessarily good in a job one level up. Fine professors do not necessarily make good deans (although not all would interpret moving to the dean's job as a promotion).¹¹ A slight modification of the definitions above and some of the formulas permit this interpretation.

To see this, allow ε_1 to be defined as the job-specific component of ability associated with the easy job and ε_2 as the job-specific component of ability associated with the difficult job. Individuals are assigned to the easy job in period 1 for the reason given before: Most are better at the easy job and in the absence of information, the easy job is the right assignment. After evaluation, \hat{A} is observed and the worker is promoted or not. If he is not promoted, then his ability post-promotion is $A+\varepsilon_1$. If he is promoted then his ability after promotion is $A+\varepsilon_2$.

Under this interpretation, workers who are not promoted have output that remains constant over time and equal to

$$\alpha + \beta(A+\varepsilon_1) \quad .$$

Those who are promoted have output equal to

$$\gamma + \delta(A+\varepsilon_2) \quad .$$

The argument of the first section holds:

¹¹Anderson, Dubinsky, and Mehta (1999) claim that the data reveal a Peter Principle for sales managers because the skills needed by salespeople are generally distinctly different from those needed by sales managers.

$$\begin{aligned}
E(\varepsilon_1 | A + \varepsilon_1 > A^*) &= \int_{-\infty}^{\infty} \int_{A^*-A}^{\infty} \frac{1}{1 - G(A^* - A)} \varepsilon g(\varepsilon) f(A) d\varepsilon dA \\
&= \int_{-\infty}^{\infty} E(\varepsilon | \varepsilon > A^* - A) f(A) dA
\end{aligned}$$

which is positive since $f(A)$ is positive and the conditional expectation of ε given ε greater than any number is positive (because the unconditional expectation of ε is zero). But the expectation of ε_2 is zero for promoted workers because ε_1 and ε_2 are uncorrelated. As a result, expected ability is higher in pre-promotion than for promoted workers post-promotion.

This does not necessarily imply that output is lower after promotion because workers are in different jobs. On the contrary, if A^* is chosen optimally, it must be the case that expected output for the promoted workers is higher in the difficult job than in the easy job. If it were not, it would be better to raise A^* until expected output were higher. Rather the point is that after promotion, the average promoted worker is not as able in the difficult job as he was in the easy job, i.e.,

$$E(A + \varepsilon_2 | \text{promoted}) < E(A + \varepsilon_1 | \text{promoted}) .$$

Also true is that within any job, those left behind and not promoted have lower average ability than those of their cohorts entering into that job. If there were a series of promotion rounds, then at every level, those who were not promoted would have a job-specific component that is

negative. This can be seen simply by examining the first round, which can be thought of as a “promotion” from being out of the firm to being hired as a worker. (Individuals must exceed some standard in order to be hired.) Since it has already been shown that

$$E(\varepsilon_1 | A + \varepsilon_1 > A^*) > 0$$

and since $E(\varepsilon_1) = 0$, it must be true that

$$E(\varepsilon_1 | A + \varepsilon_1 \leq A^*) < 0 .$$

They appear “incompetent” because within any given job, the actual ability of those who are not promoted out of the job is lower than the unconditional expectation of ability for that job. Those who are left behind and become the long-termers are worse than those who come into the job. They are incompetent relative to the entry pool because the competent workers are promoted out of the job. In a tournament with enough steps, each competent worker would continue to be promoted until he too was incompetent, i.e., until $E(\varepsilon_t) < 0$ for those whose highest job attained is job t . This is the Peter Principle: Workers are promoted to their level of incompetence. Those who are “competent” are promoted again.¹²

4. Mistake or Optimal Adjustment?

The original book was entitled, *The Peter Principle: Why Things Always Go Wrong*. The implications of the view that promotion decisions are biased are different from those of this model.

¹² One difference between this interpretation of the Peter Principle and the one used in the rest of this paper is that output of those not promoted does not rise under the job-specific interpretation. Since ε_1 is a job-specific effect and not a transitory component, those who are not promoted have ability $A + \varepsilon_1$ in both periods.

The view in this model is that if A^* is chosen optimally, most of the time firms will not want to undo their decision. Even though ability is below that predicted by a naive use of the first period estimate, the promoted worker's ability is still above x , at least on average. If it were not, then the choice of A^* would have been sub-optimal. Ability falls, but does not fall below x in most cases so the firm does not want to reverse its decision. The behavioral view that "things go wrong" is different. If the firm really made a mistake, then it would want to demote or fire workers in most cases, which leads to the inevitable question, "why are demotions so rare?" This analysis provides an answer. The promotion rule is chosen optimally so that ability is not as high as it was before promotion, but it is still high enough to justify the promotion.

B. Strategic Behavior by Workers

So far, worker effort has been assumed to be given. In this section, we relax the assumption that effort is given in order to determine how workers may game the system to alter their promotion possibilities. Whether workers over-produce during the probationary period depends crucially on the compensation scheme.

1. Efficient Effort with Worker Job Choice

The first result is that if workers are paid piece rates and allowed to choose their own jobs, all is efficient, even if workers know their ability and firms do not. In order to examine incentives, it is necessary to define more terms: Let μ_e be the amount of effort that an individual chooses if he is in the easy job and μ_d be the amount of effort he chooses if he is in difficult job. Let $C(\mu)$ be the cost incurred for any given level of effort μ . Then, if A is known, but ε is not, individuals for whom

$$(3) \quad \int_{-\infty}^{\infty} (\gamma + \delta(A + \varepsilon + \mu_d))g(\varepsilon)d\varepsilon - C(\mu_d) > \int_{-\infty}^{\infty} (\alpha + \beta(A + \varepsilon + \mu_e))g(\varepsilon)d\varepsilon - C(\mu_e)$$

choose the difficult job. Those for whom the condition in (3) does not hold choose the easy job.

Effort levels in (3) are merely the optimal levels, given the job chosen.

Because the expectation of ε is zero, (3) can be rewritten as

$$(4) \quad \alpha + \beta(A + \mu_e) - C(\mu_e) < \gamma + \delta(A + \mu_d) - C(\mu_d).$$

If the condition in (4) holds, a worker prefers the difficult job. If it does not, the easy job is selected.¹³ This is the same as the efficiency condition so workers choose jobs and effort efficiently under these conditions. There is no distortion in effort choice. The worker internalizes everything. This is a simple problem of occupational choice with effort. By contrast, were the firm to choose the job for the worker, then effort would be distorted, although the surprising result is that the distortion is as likely to take the form of under-work as it is of over-work. This is shown in the appendix.

¹³In a competitive market with rising supply price for workers (because they are distinguished by ability), firms earn zero profit. The marginal worker is the one for whom ability A_0 is low enough that

$$\alpha + \beta(A_0 + \mu) - C(\mu) = 0.$$

2. Tournaments

The usual intuition that most have about promotions inducing atypically high effort in period 1 comes from a tournament-like payment structure.¹⁴ When period 2 wages depend on the job rather than the output in the job, all workers put forth more effort than they would in the absence of period 2 promotion concerns.

The intuition holds whether the tournament is against another player or against a standard. For the purposes here, there is little difference between competing against another player and competing against a standard. In the Lazear-Rosen tournament structure, any level of effort can be implemented for any standard simply by choosing the wage spread appropriately. In the case of a standard, wages in period 2 are fixed in advance and depend only on promotion. Even if workers receive no wage prior to promotion, workers who are ignorant of A put forth effort in order to maximize

$$\text{Max}_{\mu_1} \int \{ W_d \Pr(A + \mu_1 + \varepsilon_1 > A^*) + W_e [1 - \Pr(A + \mu_1 + \varepsilon_1 > A^*)] - C(\mu_1) \} g(A) dA$$

where W_d is the difficult job's wage and W_e is the easy job's wage.

The first-order condition is

¹⁴Here again, Rosen is instrumental. The first paper on the subject is Lazear and Rosen (1981).

$$\int \{(W_d - W_e) \frac{\partial \Pr(A + \mu_1 + \varepsilon_1 > A^*)}{\partial \mu_1} - C'(\mu_1)\} dG = 0$$

or

$$(W_d - W_e) g(A^* - \mu_1 - A) = C'(\mu_1) \quad .$$

The firm can obtain any level of effort, μ_1 , simply by setting the spread between the difficult job wage and easy job wage appropriately. Then it is only necessary to set the expected wage sufficiently high to attract the marginal worker.

What is clear, however, is that effort in period 1 exceeds that in period 2. The tournament structure induces individuals to work at some positive level in period 1, but to reduce effort in period 2. In this stylized model, since there is no contingent reward in period 2, effort falls to zero. But the general point is that the tournament against a standard creates incentives to perform better in the pre-promotion period than in the post-promotion period.

Firms understand that their compensation schemes induce strategic behavior by workers and set A^* accordingly. Although this may mitigate the effects of the behavior, it in no way changes the results of this section. Since all derivations hold for any given A^* , they hold for the A^* chosen to take these effects into account.

As is the case in the tournament against a standard, workers put forth more effort before the promotion decision than after the promotion decision in a tournament against another player. This follows directly from Lazear and Rosen (1981), where effort during the contest period exceeds effort after the contest period. Worker effort during the contest period is monotonically increasing in the spread between the winner's wage and the loser's wage. After the contest has been decided, effort falls off.

In both the tournament story and the regression-to-the-mean story, worker output declines after the promotion decision. In the tournament context, it is because effort declines. In the regression-to-the-mean version, it is because of the statistical proposition that ensures that winners do worse after promotion. There is a difference, however. In tournaments, even losers reduce effort after the promotion has been decided, so expected output for all workers falls over time. In the statistical version, winners' output fall and losers' output rise above their pre-promotion levels on average.

To amplify this point, just as those who are promoted have higher-than-average pre-promotion transitory error, ε_1 , so do those who fail to be promoted have lower-than-expected transitory components. Other things equal, this implies that those who do not get a promotion should do better after being turned down than they did before. Thus, faculty who are denied tenure and move to other schools should do better on average at those other schools than they did when they were assistant professors at the first institution.

Observing this effect may be difficult for a number of reasons. For example, a worker's

output might depend on the individuals with whom he works. In an up-or-out context,¹⁵ those who fail to be promoted may find that the complementary factors in the new job are not as productivity-enhancing as those in the first job. Furthermore, motivation is an issue. To the extent that an individual believes that he is in the running for promotion, tournament effects are present, inducing effort. After the promotion has been denied, the incentives vanish, reducing effort and output.

C. Other Examples of the Principle

The regression-to-the-mean phenomenon that is observed as the Peter Principle in the labor market has other manifestations. For example, it is often observed that sequels are rarely as good as the original movie on which the sequel is based.¹⁶ If each movie is thought of as having a theme-constant component, A , and a transitory component ε_t , e.g., actors, specific story, or direction, associated with each particular film, then the same analysis holds. In order for a sequel to be made, the value of the original film, $A + \varepsilon_t$, must be estimated to be greater than A^* . But given that the value exceeds the threshold level, A^* , the expectation of the value of the sequel will be less than the original simply because

$$E(\varepsilon_1 \mid \text{sequel is made}) > 0,$$

¹⁵See Kahn and Huberman, (1988).

¹⁶Data from Teichner and Luehrman (1992) establish clearly that revenues are lower and costs higher on sequels than the are on the original on which they are based. See, for example, their exhibits 4 and 5.

but

$$E(\varepsilon_2 \mid \text{sequel is made}) = 0.$$

As a result, an original film must be sufficiently good to generate a sequel because studios adjust upward their cutoff levels, knowing that the second film is likely to be inferior to the first.

It is straightforward to test this proposition. Among other things, it implies that measures of film quality such as academy awards or ticket sales should be higher on the original film than on the sequels.¹⁷

The fact that sequels do worse than originals is not evidence that the studio made a mistake. Once again, if A^* is chosen optimally, then the average sequel is profitable, even if not as profitable as the original. A decline in profitability does not imply that something went wrong in decision making.

Similarly, the first meal in a good restaurant is often the best, followed by less satisfying repeat visits. Just as above, think of A as the restaurant component of the first meal (recipes, management) and ε_1 as the transitory component associated with the meal itself (that night's chef, specific ingredients, dinner companion). A second visit to the restaurant is made only if

$$A + \varepsilon_1 > A^* .$$

Once again, the expected value of the second meal lies below that of the first, conditional on deciding to make a second visit to the restaurant. The larger is the transitory component, the larger

¹⁷A countervailing effect is the notoriety that is created by the first film, which makes it easier to sell tickets on the sequel than on the original. Even if consumers understand that the sequel is worse than the original, more tickets might be sold on the sequel if, say, the actors and director are not well-known before the first film is made.

is the discrepancy between first and second meal and the higher is the standard set to merit a second visit. As before, the fact that the second visit is not as good as the first does not imply that a mistake was made. The second meal on average is good enough to justify a repeat visit to the restaurant, even if it is not as good as the first. The point can also be used to explain why favorite restaurants go out of fashion. A restaurant becomes a favorite in part because of the permanent component (e.g., good recipes and an insightful owner) and in part because of potentially transitory components (e.g., the current chef and the service of the staff). A favored restaurant can be thought of as one that has gotten a draw of $A + \varepsilon_1 > A^*$.

It is favored precisely because the value of the output exceeds some standard. Over time, ε_1 is replaced by transitory effects, the expectation of which is zero. The quality falls and the restaurant goes out of fashion.

The same logic can be applied to the “*Sports Illustrated* Effect.” It is claimed that it is a curse to be on the cover of *Sports Illustrated* because athletes’ performance falls thereafter. Again this reflects regression to the mean, but it does not imply that *Sports Illustrated* chose the wrong athlete for the cover.

In the political arena, re-elections of elected officials occur when constituents view performance as having exceeded some standard. The post-election performance should be worse absent learning and seniority effects. If the learning effect is not too strong, second term presidents will do less well than they did during their first term. But this does not imply that voters made a mistake. Even knowing that there is regression, voting for the incumbent is rationale if he exceeds a sufficiently high standard.

Finally, there is a close relation of this analysis to that of the winner's curse.¹⁸ If a bidder submits a bid based on a naive estimate, the winner will lose money, because on average the winner has obtained a higher-than-average error. As a result, he shades his bid so in equilibrium he adopts the rent maximizing strategy. The same is true here. By adjusting the cutoff criterion, A^* , the decision maker maximizes rent

D. Length of Probationary Period and Relative Importance of the Transitory Component

The longer a firm waits to make a promotion decision, the better the information. One would expect that transitory components that bias a decision could be reduced or eliminated if the firm waited long enough to make a promotion decision. The cost of waiting, however, is that workers are in the wrong job for more of their lifetimes. For example, suppose that it were possible to get a perfect reading on A by waiting until the date of retirement. The information would have no value because the worker would have spent his entire working career in the easy job, even if he were better suited to the difficult job. The tradeoff is modeled. The conclusion is that as the variance of ε_1 rises, it becomes more valuable to wait on a promotion decision.

To see this, let us add one period to the previous model (without effort). Now, ε_1 , ε_2 , and ε_3

¹⁸See Wilson (1969). Winner's curse usually relates to a reading relative to others' readings rather than the time dimension of taking multiple readings, sometimes in different settings. The transitory versus permanent component is central to the discussion of this paper, but not the theme of most of the winner's-curse literature.

Actually, loser's curse is as important to the assignment problem as is winner's curse. In the job context, the goal is to assign a worker to the right job. Workers who do not satisfy the promotion criterion are, on average, undervalued just as those who are promoted are overvalued. The optimal selection of the cutoff point trades off the two kinds of errors.

refer to the transitory component in periods 1, 2 and 3 and assume that they are distributed i.i.d. and to reduce notation, that $E(A)=0$. Suppose that by waiting two periods, an employer can obtain a perfect reading of A . Under those circumstances, the optimum is simply to promote those for whom $A>x$. The cost is that when the firm delays its promotion decision to the end of period 2, all workers are in the easy job during period 2 even though it might be better to place some in the difficult job in period 2. Expected output over the lifetime is then

$$\begin{aligned}
 \text{Expected Output if Wait} &= 2(\alpha + \beta E(A + \varepsilon_1)) + \int_x^\infty (\gamma + \delta A) dF + \int_{-\infty}^x (\alpha + \beta A) dF \\
 (5) \qquad \qquad \qquad &= 2\alpha + \int_x^\infty (\gamma + \delta A) dF + \int_{-\infty}^x (\alpha + \beta A) dF
 \end{aligned}$$

The alternative is to make a decision after one period, using imperfect information and recognizing that sorting will be imperfect. To make things simple, assume that a firm that makes a promotion decision at the end of period 1 cannot reevaluate at the end of period 2. The gain is that workers are sorted early so that very able people can be put in the difficult job more quickly. The cost is that more errors are made in assigning workers to jobs. Then, expected output over the three periods is

$$\begin{aligned}
 \text{(6)} \quad \text{Expected Output Early} &= (\alpha + \beta E(A + \varepsilon_1)) + 2 \int_{-\infty}^{\infty} \int_{A^* - \varepsilon}^{\infty} (\gamma + \delta A) dF dG + 2 \int_{-\infty}^{\infty} \int_{-\infty}^{A^* - \varepsilon} (\alpha + \beta A) dF dG \\
 &= \alpha + 2 \int_{-\infty}^{\infty} \int_{A^* - \varepsilon}^{\infty} (\gamma + \delta A) dF dG + 2 \int_{-\infty}^{\infty} \int_{-\infty}^{A^* - \varepsilon} (\alpha + \beta A) dF dG
 \end{aligned}$$

In an extreme case, it is clear that it pays to decide early. If the distribution of ε is degenerate so that there is no error, then (14) becomes

$$\text{(7)} \quad \text{Expected Output Early} = \alpha + 2 \int_x^{\infty} (\gamma + \delta A) dF dG + 2 \int_{-\infty}^x (\alpha + \beta A) dF$$

The r.h.s. of the expression in (7) must exceed the r.h.s. of (5) because

$$\gamma + \delta A > \alpha + \beta A \quad \text{for } A > x$$

since that is how x is defined. Thus, when the variance in ε shrinks to zero, it always pays to promote early.

The example used earlier shows that it sometimes pays to defer the promotion decision until the end of period 2. As before, let $\alpha=1$, $\beta=.5$, $\gamma=0$ and $\delta=1$, where the distributions of A and ε_1 are normal with variance equal to one. As shown earlier, the optimal cut point is $A^*=4.01$. Then, the r.h.s. of (5) equals 3.004. The r.h.s. of (6) is 3.000. Thus, deferring the promotion decision until the

second period pays when the variance in ε is 1. Other numerical examples show that the advantage of deferring the promotion decision becomes larger for higher variances in ε .

The general point is that when the distribution of ε is sufficiently tight, it pays to make the promotion decision early. When it is sufficiently diffuse, it pays to make the promotion decision later. Later promotion decisions are more accurate, but result in workers' spending a longer proportion of their worklife in the wrong job.

Conclusion

Workers who are promoted have been observed to have exceeded some standard. Part of the observation is based on lasting ability, but part is based on transitory components that may reflect measurement difficulties, short-term luck, or skills that are job specific. As a result, there is regression to the mean, creating a "Peter Principle." Workers who are promoted do not appear to be as able as they were before the promotion.

Firms take this phenomenon into account in setting up their promotion rule. Under general conditions, when fewer than 50% of the workers are better suited to the high level job, the firm adjusts the promotion standard upward to compensate for the regression to the mean. The amount of the adjustment depends on the tightness of the error distribution. When the pre-promotion error has high dispersion, promotion standards are inflated by more than they are when the error dispersion is low.

The statistical argument has been contrasted with incentive arguments. Whether workers over-produce because they are gaming the system depends on the payment structure. If, for example, output were observable so that workers could be paid on the basis of output both before

and after the promotion decision, then it would be optimal to allow workers to make their own job choice. Under these circumstances, there is no distortion in effort; all is efficient. When a tournament structure is chosen because of inability to observe output, workers produce more before promotion than they do after promotion. Although tournaments result in declining output after promotion, the implications of tournaments for losers and winners are different from those of the statistical argument. In particular, in tournaments, output of both losers and winners falls after promotion. The statistical argument implies that losers' output rises and winners' output falls after promotion.

The Peter Principle can be interpreted to mean that workers are not as able as perceived before promotion or that they were better in their prior job relative to their peers than they are in their current one. In a multi-level firm, the typical worker who remains at a given level is "incompetent" in that he is not as good as the average worker coming into the job, nor is he as good as he was in his previous assignment relative to the comparison set.

One way to offset the Peter Principle is to wait for a longer time before making a promotion decision. The advantage is that the job assignment is better than it would have been had the decision been made earlier. The disadvantage is that able workers remain in the wrong job for a longer period of time.

The logic of the Peter Principle applies in other contexts as well. The regression-to-the-mean phenomenon implies that movie sequels are lower quality than the original films on which they are based and that excellent restaurant meals are followed by ones that are closer to mediocre.

References

- Anderson, Ralph E., Alan J. Dubinsky, and Rajiv Mehta. "Sales Managers: Marketing's Best Example of the Peter Principle?" *Business Horizons*, 4, 2, 1, 19 (1999), 19-26.
- Baker, George, Michael Gibbs and Bengt Holmstrom. "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics*, November, 1994: 881-919.
- Fairburn, James A. and James M. Malcomson. "Performance, Promotion, and the Peter Principle," *Review of Economic Studies*, forthcoming, 2000.
- Faria, Joao Ricardo. "An Economic Analysis of the Peter and Dilbert Principles," Unpublished manuscript, University of Technology, Sydney, Australia, (2000).
- Gibbons, R. and Waldman, M., "Careers in Organisations: Theories and Evidence" in *Handbook of Labor Economics*, vol. 3B, ch 36. Amsterdam: North Holland, 1999. (1999b)
- Gibbons, Robert and Michael Waldman. "A Theory of Wage and Promotion Dynamics," *Quarterly Journal of Economics*, 114: 4 (November 1999) 1321-58 .
- Gibbs, Michael and Wallace Hendricks. "Are Formal Salary Systems a Veil?" University of Chicago, August, 2001.
- Jovanovic, Boyan. "Job Matching and the Theory of Turnover," *Journal of Political Economy* 87 (October 1979): 972-90. (1979a)
- Jovanovic, Boyan. "Firm-Specific Capital and Turnover," *Journal of Political Economy* 87 (December 1979): 1246-60. (1979b)
- Kahn, Charles, and Huberman, Gur. "Two-Sided Uncertainty and 'Up-or-Out' Contracts," *Journal of Labor Economics* 6 (October 1988): 423-44.
- Lazear, Edward P. "The Job as a Concept," in *Performance Measurement and Incentive Compensation*, ed. William J. Bruns, Jr. Cambridge: Harvard Business School Press, 1992.
- Lazear, Edward P. "Raids and Offer-Matching," in Ehrenberg, Ronald, ed. *Research in Labor Economics*, Vol. 8, 1986, part A: 141-65.
- Lazear, Edward P., and Rosen, Sherwin. "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy* 89 (October 1981): 841-64.

Medoff, James, and Abraham, Katharine. "Experience, Performance, and Earnings," *Quarterly Journal of Economics* **95** (December 1980): 703-36.

Peter, L.J. and R. Hull. *The Peter Principle: Why Things Always Go Wrong*. New York: Morrow (1969).

Prendergast, Canice. "The Provision of Incentives in Firms," *Journal of Economic Literature*, March 1999, 37, 7-63

Rosen, Sherwin. "Prizes and Incentives in Elimination Tournaments," *American Economic Review* **76** (September 1986): 701-15.

Teichner, William A. and Timothy A. Luehrman. "Arundel Partners: The Sequel Project," Harvard Business School Case 9-292-140. Boston: Harvard Business School Publishing, 1992.

Wilson, R., "Competitive Bidding with Disparate Information," *Management Science*, Vol. 15, No. 7 (March 1969), pp. 446-448. Reprinted in Steven A. Lippman and David K. Levine (eds.), *The Economics of Information*, Edward Elgar Publishing, London, 1994; and Paul Klemperer (ed.), *The Economic Theory of Auctions*, Edward Elgar Publishing, London, 1999.

Appendix

Effort is Distorted When Firms Assign Jobs, but Worker Know Ability

Assume asymmetric information where workers know their abilities, A , but firms do not. Define μ_1 as effort in period 1, μ_{2e} as effort in period 2 if the worker is not promoted, and μ_{2d} as effort in period 2 if the worker is promoted. Note that effort in period 1 is determined before the promotion decision is made, so period 1 effort is independent of promotion. The cost of effort is given by $C(\mu)$. For simplicity, $C(\mu)$ is assumed to be independent of ability and the same across periods.

The worker is paid a piece rate so in period 2, a worker who has not been promoted chooses effort, μ_{2e} , so as to solve

$$\underset{\mu_{2e}}{\text{Max}} \quad \alpha + \beta E(A + \mu_{2e} + \varepsilon_2) - C(\mu_{2e})$$

or

$$\underset{\mu_{2e}}{\text{Max}} \quad \alpha + \beta(A + \mu_{2e}) - C(\mu_{2e}).$$

The first order condition to is

$$(A1) \quad C'(\mu_{2e}) = \beta.$$

An analogous problem can be solved for those who are promoted. Their problem is

$$\underset{\mu_{2d}}{Max} \quad \gamma + \delta(A_2 + \mu_{2d}) - C(\mu_{2d})$$

which has first-order condition

$$(A2) \quad C'(\mu_{2d}) = \delta.$$

Eq. (A1) and (A2) define μ_{2e} and μ_{2d} . Promoted workers put forth more effort in period 2 because the marginal return to effort is higher in the difficult job than in the easy job, i.e., $\delta > \beta$. Given this, the worker solves a two-period problem in period 1, knowing that he will choose μ_{2d} and μ_{2e} , depending on whether or not he is promoted.

The worker who knows his own ability has a first period problem given by¹⁹

$$\underset{\mu_1}{Max} \quad \alpha + \beta E(\mu_1 + A + \varepsilon_1) - C(\mu_1) + \text{Prob}(A + \mu_1 + \varepsilon_1 > A^*) E\{\gamma + \delta(A + \mu_{2d} + \varepsilon_2) - C(\mu_{2d})\} \\ + \text{Prob}(A + \varepsilon_1 + \mu_1 \leq A^*) E\{\alpha + \beta(A + \mu_{2e} + \varepsilon_2) - C(\mu_{2e})\}$$

or

$$(A3) \quad \underset{\mu_1}{Max} \quad \alpha + \beta(\mu_1 + A) - C(\mu_1) + [1 - G(A^* - \mu_1 - A)] [\gamma + \delta(A + \mu_{2d}) - C(\mu_{2d})]$$

¹⁹The discount rate is assumed to be zero.

$$+ G (A^* - \mu_1 - A) [\alpha + \beta (A + \mu_{2e}) - C (\mu_{2e})] .$$

The first-order condition is

$$(A4) \quad \beta - C(\mu_1) + g(A^* - \mu_1 - A) \{[\gamma + \delta(A + \mu_{2d}) - C(\mu_{2d})] - [\alpha + \beta(A + \mu_{2e}) - C(\mu_{2e})]\} = 0 .$$

Efficient effort is supplied when workers set $C(\mu_1) = \beta$. According to the first-order condition in (8), this occurs only when the last term on the l.h.s. is equal to zero. In general, it will not be zero. In fact, the last term is positive, implying over-investment, when

$$[\gamma + \delta(A + \mu_{2d}) - C(\mu_{2d})] > [\alpha + \beta(A + \mu_{2e}) - C(\mu_{2e})] .$$

Sufficiently high-ability workers prefer job 1 because they earn more in job 1. As a result, they overwork in period 1 to enhance the probability that they will be promoted. Because the firm cannot distinguish effort from ability, workers who want to be promoted have an incentive to work too hard in order to fool the firm into believing that their ability levels are higher than they actually are.

Less intuitive, the converse is also true. Low-ability workers, i.e., those for whom A is sufficiently low so that

$$[\gamma + \delta(A + \mu_{2d}) - C(\mu_{2d})] < [\alpha + \beta(A + \mu_{2e}) - C(\mu_{2e})] ,$$

under-work.²⁰ These workers underachieve because they do not want to take the chance of being promoted. From their point of view, a promotion is bad because they are likely to earn less in the difficult job than in the easy job.

The workers who are most likely to distort their effort in period 1 are those for whom

²⁰Since μ_2 is independent of A , there is always an A sufficiently low to make this condition hold.

$$g(A^* - \mu_1 - A)$$

is high (see eq. (8)) and for whom

$$| [\gamma + \delta(A + \mu_{2d}) - C(\mu_{2d})] - [\alpha + \beta(A + \mu_{2e}) - C(\mu_{2e})] |$$

is high. Under standard assumptions about the distribution of ε , in particular that

$$\lim_{\varepsilon \rightarrow -\infty, \infty} g(\varepsilon) = 0,$$

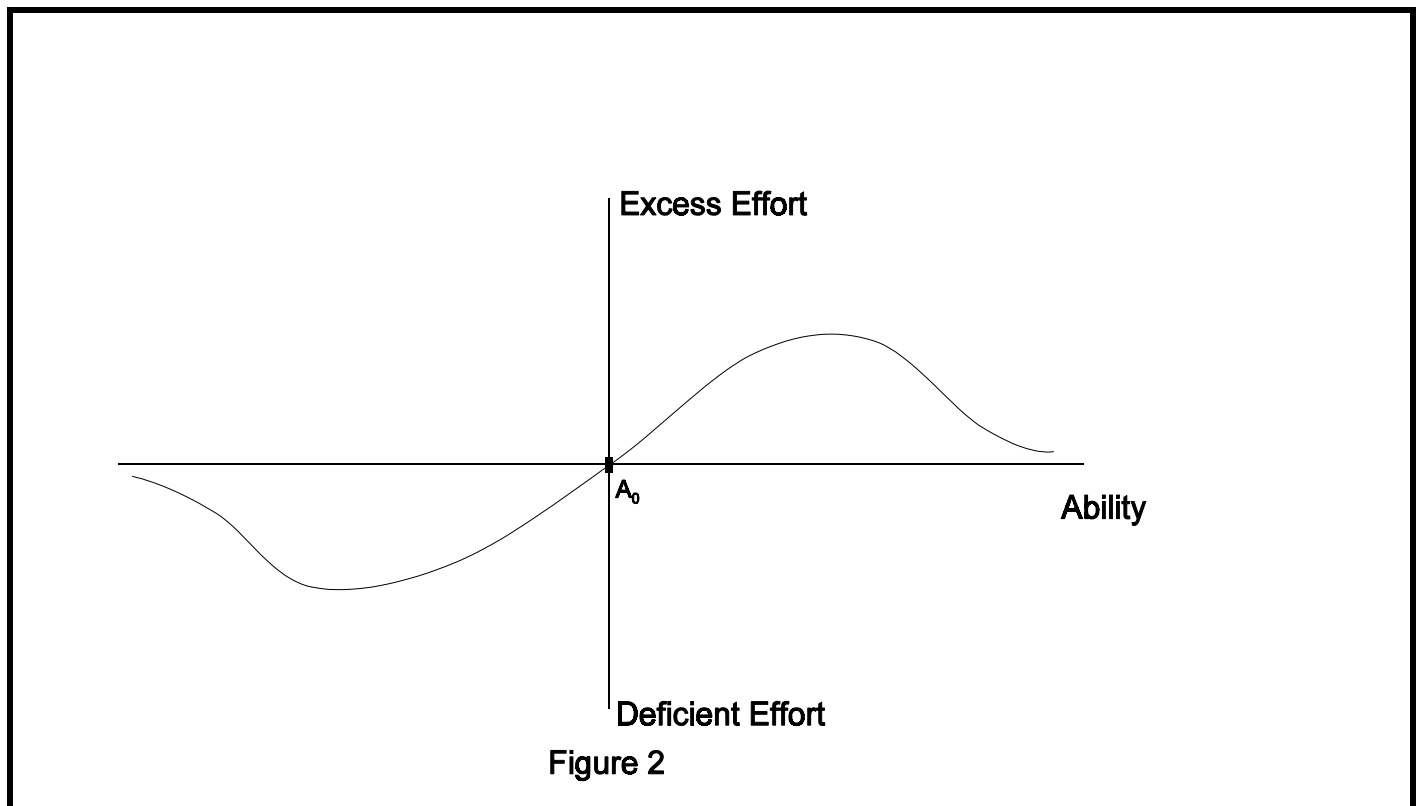
very high- and very low-ability workers choose the efficient level of effort in period 1. They have little to fear in terms of incorrect promotion decisions. The extremely able are almost certain to be promoted, so that extra effort has very little effect on the probability of promotion. Conversely, the totally inept are almost certain to avoid promotion, so that reducing effort has almost no effect on

lowering the probability of promotion.

Also true is that those whose underlying ability is very near the efficient job switch point (x in figure 1) do not distort effort much. Even if they are misclassified, they have little to lose. Define A_0 such that

$$[\gamma + \delta (A_0 + \mu_{2d}) - C(\mu_{2d})] = [\alpha + \beta(A_0 + \mu_{2e}) - C(\mu_{2e})]$$

Then, excess effort is zero at both extremes and also at A_0 , which is likely to be close to x . The pattern of distortion is shown in figure 2. Those at A_0 do not distort at all. Those at the ability



extremes to do not distort. Those with ability less than the switch point under-work and those with ability more than the switch point over-work.

IZA Discussion Papers

No.	Author(s)	Title	Area	Date
743	G. Heineck J. Schwarze	Substance Use and Earnings: The Case of Smokers in Germany	5	03/03
744	M. R. West L. Woessmann	Which School Systems Sort Weaker Students into Smaller Classes? International Evidence	2	03/03
745	L. Woessmann	Educational Production in East Asia: The Impact of Family Background and Schooling Policies on Student Performance	2	03/03
746	A. Ammermueller H. Heijke L. Woessmann	Schooling Quality in Eastern Europe: Educational Production During Transition	4	03/03
747	J. Messina	Sectoral Structure and Entry Regulations	3	03/03
748	M. Pflüger	Economic Integration, Wage Policies and Social Policies	2	04/03
749	V. Jakobsen N. Smith	The Educational Attainment of the Children of the Danish 'Guest Worker' Immigrants	1	04/03
750	R. Hujer M. Caliendo S. Thomsen	New Evidence on the Effects of Job Creation Schemes in Germany - A Matching Approach with Threefold Heterogeneity	6	04/03
751	L. Borghans B. ter Weel	Are Computer Skills the New Basic Skills? The Returns to Computer, Writing and Math Skills in Britain	5	04/03
752	N. Malchow-Møller J. R. Skaksen	Skill-Biased Technological Change in Denmark: A Disaggregate Perspective	5	04/03
753	S. Burgess H. Turon	Unemployment Equilibrium and On-the-Job Search	1	04/03
754	H. Turon	Separability of Duration Dependence and Unobserved Heterogeneity	1	04/03
755	T. Dohmen	In Support of the Supporters? Do Social Forces Shape Decisions of the Impartial?	7	04/03
756	N. Datta Gupta R. Oaxaca N. Smith	Swimming Upstream, Floating Downstream: Comparing Women's Relative Wage Positions in the U.S. and Denmark	5	04/03
757	J. Hunt	Teen Births Keep American Crime High	5	04/03
758	J. Lise S. Seitz J. Smith	Equilibrium Policy Experiments and the Evaluation of Social Programs	6	04/03
759	E. P. Lazear	The Peter Principle: A Theory of Decline	1	04/03

An updated list of IZA Discussion Papers is available on the center's homepage www.iza.org.