INTERNATIONAL ECONOMIC REVIEW
Vol. 41, No. 2, May 2000

A MODEL OF DECLINING STANDARDS*

BY JOEL SOBEL[1]

Department of Economics, University of California, San Diego, U.S.A.

This paper presents a model in which relative standing determines standards. There are three kinds of agents in the model: candidates who wish to pass a test, members of the elite who have passed the test, and the judge who decides who passes. In order to pass, a candidate's performance must be at least as good as the performance of a representative member of the elite. Without perturbations in the underlying data, the model predicts that standards will not change. Perturbations in the preferences used to judge candidates lead to a reduction in standards.

1. INTRODUCTION

In many situations, individuals need to pass a test in order to obtain a credential, signal their competence, or get a promotion. This article investigates the dynamic implications of one rule for determining what is necessary to obtain a credential.

I present a model in which relative standing determines the level of achievement needed to pass a test. A candidate passes a test provided that his or her performance compares favorably with that of people who have passed the test recently. The main result is that standards are likely to decline under such a system.

The setting is simple. There are three kinds of agents: (1) candidates who wish to pass a test, (2) members of the elite who have already passed the test, and (3) the judge who decides who passes. In order to pass, a candidate's performance must be at least as good as the performance of a representative member of the elite. Candidates know what it takes to pass the test. Their effort determines their performance. They choose the least expensive effort level that will enable them to pass. Members of the elite retire and are replaced by candidates who have passed the test recently. Without perturbations in the underlying data (in this case, the preferences of the judge), the model predicts that standards will not change. The cost needed to pass the test remains constant. Nevertheless, perturbations in the preferences used to judge candidates systematically lead to a reduction in standards. The logic of the result is simple. A change in preferences provides new ways to pass the test. A candidate can pass by imitating the behavior of the previous cohort of

*Manuscript submitted July 1997; revised October 1998.

[1] Various committee assignments helped to form my thoughts on this subject. Thanks to my fellow masochists for making these jobs less onerous. Thanks also to Julian Betts, Vince Crawford, Garey Ramey, Joel Watson, referees, and several seminar audiences for comments and encouragement. Thanks also to Cynthia Bansak and Michael Davis for research assistance and to the National Science Foundation and CASBS for financial support.

candidates. When the judge's preferences change, candidates may be able to provide an equivalent or superior performance at a lower cost. If future candidates are again judged by the original preferences, then the cost of acquiring the skills needed to pass drops. In this sense, the higher-level equilibria are unstable.

It is common to use relative performance to compensate workers or to determine who should get promoted. Relative measures can be used to create incentives that encourage workers to select high levels of effort. When the output of one worker provides information about private actions or information of another worker, using the information can benefit an employer both because it enables the employer to draw more precise inferences about the worker's effort and because it creates opportunities for sharing risk. The literature on tournaments (Green and Stokey, 1983; Holmstrom, 1982; Lazear and Rosen, 1981; Nalebuff and Stiglitz, 1983a, 1983b) makes these points clearly. Shleifer (1985) demonstrates the value of basing compensation on the performance of other firms in a regulatory context.

In this article there is no competition between candidates. In contrast to tournament models, relative performance is not used to determine which of several possible workers will receive a prize. Instead, candidates must compare well against a standard created by the performance of past candidates. If a candidate does as well as successful candidates from the past, then he or she knows that he or she will pass.

Relative comparisons are an essential aspect of many certification processes. The result identifies a factor that could lead to grade inflation, declines in the quality of high school graduates,[2] and declines in the minimum needed to pass a graduate school qualifying examination.

The model of this article describes aspects of the promotion process in research universities. Promotion to tenure is the outcome of a departmental and administrative review. The rules governing the process are complicated, but settlements based on recent legal decisions suggest that a sufficient condition for promotion is that external reviewers compare a candidate favorably with recently promoted members of the department.[3] Explicit competition is ruled out because there is not an open search for a new tenured position. Instead, internal candidates are reviewed separately. Furthermore, the criteria for promotion are modified periodically. Recently, administrators have suggested that teaching quality should be given more weight in the review process. Thus the promotion process shares some of the features of the theoretical model.

Limited data on promotion trends in the University of California system are available. Fifty-eight percent of the assistant professors hired at the University of California at Berkeley between 1970–1971 and 1979–1980 were promoted to tenure; that figure rose to 72 percent for those hired between 1985–1986 and 1990–1991. At

[2]The report of the National Commission on Excellence in Education (1983) raised concerns about the quality of high school education. The report calls attention to a feature that plays an important role in my models when it observes (p. 20) that "'minimum competency' examinations...fall short of what is needed, as the 'minimum' tends to become the 'maximum,' thus lowering educational standards for all." Over the past decade, there is some evidence that performance of secondary students on standardized tests has been improving, however.

[3]WAGE (1993) describes recent cases in the Law School and Mathematics Department at the University of California at Berkeley.

the University of California at San Diego, 63 percent of assistant professors hired between 1969–1970 and 1978–1979 eventually received tenure; this figure rose to 72 percent for the period between 1979–1980 and 1987–1988 (see Scull et al., 1995). These percentages may accurately reflect a trend in promotion rates, or they could reflect an increase in the quality of the pool of candidates. The forces identified in this article also may play a role in these promotion trends.

Two other articles study standard setting in similar settings. Barberà et al. (1998) analyze a model of voting for new members of a club in which voting behavior is strategic. This article examines a finite-horizon model in which candidates' characteristics are fixed, judges are the current members of the elite, and candidates need the support of only one judge to join the elite. Judges decide if and when to support a candidate. Sobel (1999) analyzes a related model in which there is more than one judge. When the preferences of judges are heterogeneous, declining standards are possible even without perturbations. This article demonstrates the possibility of rising standards when promotions require the support of a supermajority of the judges.

Costrell (1994) presents a model of educational standards that suggests that these standards may be set lower than is socially optimal. Betts (1998) reconsiders the model. He presents alternative assumptions that make the direction of bias in standard setting ambiguous. The emphasis of these articles is different from this one. They focus on the welfare implications of static models. This article identifies a particular bias in a dynamic model.

Leland (1979) introduces a model that explains why professional groups may wish to set standards. Leland assumes that professionals have better information about the quality of the service they offer to customers. The professionals can charge a price corresponding to the value of the average quality supplied by all professionals. If the professionals that supply the highest quality have higher outside opportunities, Leland demonstrates that the market might fail. The best-quality sellers may accept their outside opportunity rather than be treated like average-quality suppliers. Imposing minimum quality standards may enable the market to survive, although not necessarily at socially optimal levels. Leland concentrates on the effect of quality standards given a distribution of abilities. He does not discuss how setting standards may influence the abilities of professionals.

The next section contains the formal model. Section 3 concludes the paper.

2. PERTURBATIONS MAY DECREASE STANDARDS; THEY DO NOT INCREASE STANDARDS

There are three types of agents in the model: candidates, elites, and a judge. The candidates want to acquire a credential or get promoted. Refer to their objective as passing a test. Candidates choose an effort level that determines their performance on the test. Elites are agents who have passed the test. The performance of the candidates is compared with that of the elite. The judge determines which candidates should pass.

Performance is described by an element x of $\mathbf{R}^{\mathbf{n}}_{+}$. There is a cost function $C(x)$; $C(x)$ is the effort that a candidate must expend to achieve performance x. The judge has preferences described by a utility function $U(\cdot)$. Assume that these functions are continuous and strictly increasing. A candidate passes in period t provided that he or she meets or exceeds a standard set by the group promoted in the preceding period.

Assume that in each period some candidates pass. These candidates must be at least as good as the rth member of the existing population of elites. Following a promotion, an appropriate number of elites, selected at random, retire. This assumption keeps the population of elites constant.

Let $x_t(i)$ for $i = 1, \dots, R$ be the characteristics of the R members of the elite at time t. $x_0(i)$ is the performance of an original member of the elite. For convenience, order these elements so that $U[x_t(i)] \geq U[x_t(i+1)]$ for all $i = 1, \dots, R-1$. Candidates at t solve the problem

$$(1) \qquad \min C(x) \quad \text{subject to} \quad U(x) \geq U\left[x_{t-1}(r)\right].$$

That is, a candidate passes at time t if the judge believes that the candidate is as good as the rth best member of the elite at time $t-1$.

$U(\cdot)$ and $C(\cdot)$ are continuous and increasing. Hence it follows that Problem (1) has a solution and that the constraint in Problem (1) binds at each solution. Let x^*_t be a solution to Problem (1), P be the payoff associated with passing the test, and $W(x)$ be the best available outside opportunity to an agent with characteristic x. It follows that a candidate will select effort x_t^* provided that

$$(2) \qquad P - C(x_t^*) \geq \max_x\{W(x) - C(x)\}.$$

Assume for convenience that Inequality (2) always holds. This assumption simplifies the exposition. The main result does not change if different candidates have different values of P and $W(\cdot)$ and that only some of the agents attempt to pass the test.

The dynamic process converges with probability one to a configuration in which every member of the elite population performs at a standard equal to that of the rth best member of the original population.

PROPOSITION 1. *If a candidate must perform at a standard equal to that of the* rth *best member of the current of the elite in order to pass, at any time t after which all members of the original population have retired,* $U[x_t(i)] = U[x_0(r)]$ *for all i.*

This proposition states that after the original members of the elite have retired, subsequent elite populations are homogeneous and perform at a level that is equivalent (from the point of view of the judge) to the characteristic of the rth best member of the original population.

PROOF. Assume that there is precisely one candidate in each period. If more than one candidate passes in a period, then convergence to a homogeneous population is faster. Call the skill of the rth best member $x(r)$, and call any x such that $U(x) = U[x(r)]$ *typical.* I claim that eventually every member of the elite is typical. The reason is that under the maintained assumptions, every new member of the elite is typical. The first replacement must be typical (because standards are never exceeded). Hence, prior to the first retirement, there are $R+1$ members of the elite, and at least the rth and $r+1$th are typical. It follows that after retirement, the rth best member of the elite will be typical. Furthermore, if neither the newcomer nor the original rth best elite retire, then in the next period there will be at least two people with typical performance levels. Repeating this process always leads to entry at the typical performance level. The number of elite at this level cannot go down. It will go up whenever someone with an atypical performance level retires. When the original population of elites has retired, every member of the elite will be typical. ■

In the period of adjustment, the performance of the elite could be increasing. For example, if candidates need to be as good as the best current member of the elite, then until every member of the elite has an equivalent performance level, a newcomer would always be at least as good as the person he or she replaces.

In light of the Proposition 1, assume that at any time t, $U[x_i(t)] \equiv U_t$. In this case, the criterion for passing becomes: A candidate passes in period t if his or her characteristic x satisfies $U(x) \geq U_{t-1}$. That is, in order to pass, a candidate must be at least as good as a representative member of the previous elite. In period t an individual will select x to solve

$$(3) \qquad \min C(x) \quad \text{subject to} \quad U(x) \geq U_{t-1}.$$

It follows from Proposition 1 that any initial specification of elite performance uniquely determines the limiting performance of the elite. The simple observation demonstrates that the model of standards has multiple equilibria. Standards remain constant over time (after every member of the original elite retires); the cost of meeting those standards does not change.

In this model, perturbations systematically lower standards. To make this point as simply as possible, assume that candidates have identical cost functions (this assumption is not necessary for Proposition 1), that the solution to Problem (3) is unique, and that members of the elite retire after one period. These assumptions guarantee that each member of the elite picks the same performance level. I discuss how to relax this assumption at the end of the section. Assume that in period t the judge evaluates individuals using different preferences; call the new utility function $V(\cdot)$. If candidates are aware of this change, then they will select their effort to solve

$$(4) \qquad \min C(x) \quad \text{subject to} \quad V(x) \geq V(x_{t-1}).$$

Notice that x_{t-1} is feasible for Problem (4). Hence the solution to this problem must cost no more than $C(x_{t-1})$. Moreover, if the indifference curves through x_{t-1} of $V(\cdot)$ differ from those of $U(\cdot)$, then the solution to Problem (4) will lead to a cost strictly less than $C(x_{t-1})$. When all functions are differentiable and solutions to Expression (4) lie in the interior of $\mathbf{R}^{\mathbf{n}}_{+}$, a sufficient condition for the indifference curves to differ is that denoting the derivative of a function F by DF, $DU(x_{t-1}) \neq DV(x_{t-1})$.

The implication of this result is apparent. When preferences shift back to $U(\cdot)$ in a subsequent period, the cost of meeting the new standard cannot be greater than it was prior to the perturbation, and if x_{t-1} does not solve Problem (4), it will be lower. A one-time perturbation in the preferences of the judge leads to a permanent reduction in standards. These comments establish Proposition 2.

PROPOSITION 2. *If the elite are homogeneous and a candidate must perform at a level x_t that satisfies $V(x_t) \geq V(x_{t-1})$ in period t and a level x_{t+1} that satisfies $U(x_{t+1}) \geq U(x_t)$ in period $t+1$, then $C(x_{t-1}) \geq C(x_t) \geq C(x_{t+1})$. Moreover, the first inequality is strict if $U(\cdot)$, $V(\cdot)$, and $C(\cdot)$ are differentiable at x_{t-1} and $DY(x_{t-1}) \neq DV(x_{t-1})$. [The second inequality is strict if $DU(x_t) \neq DV(x_t)$.]*

The cost of meeting the standard decreases (weakly) whenever preferences change. That is, standards decrease even if the perturbation is permanent. Interpreting the change as a temporary perturbation identifies a clear sense in which equilibria that require a more costly level of skill are less stable than lower-level equilibria.

Figure 1 illustrates the proposition. The original elite population supplies the characteristic x_0; this vector lies on the isocost curve C_0. If the preferences of the judges change (so that one indifference curve is given by V), then it is optimal to supply x_1; this performance lies on the lower isocost curve C_1.[4] Finally, if the preferences of the judge switch back to U, then candidates can pass by supplying x_2. The performance x_2 lies on a still lower isocost curve C_2. Moreover, the judge views x_2 as strictly worse than the original performance level x_0 because the indifference curve labeled U^* is strictly below the original indifference curve labeled U.

How should one understand Proposition 2? A decrease in the cost of meeting the standards means that then it is easier to meet those standards. A change in preferences forces the judge to apply different weights to different components of the performance vector. Proposition 2 demonstrates that this change results in falling standards. Hence policy changes that temporarily increase the weight placed on reading relative to mathematics in the high school curriculum or teaching quality relative to research in tenure reviews could have the effect of lowering quality of high school graduates or tenured faculty. The cost of meeting standards could go down both when preferences change from $U(\cdot)$ to $V(\cdot)$ and when they change back again.

To provide an intuition for the result, imagine that preferences $U(\cdot)$ place little emphasis on the teaching of assistant professors, while $V(\cdot)$ weights teaching more heavily. When preferences change, the new generation is judged in a way that rewards teaching, but against a standard determined by the skills of people who were not rewarded for teaching quality when they made their own investment choice. The newcomers can meet the new standards by teaching somewhat better than their elders, while doing considerably less research. If, sometime in the future, the relative importance of research is increased, then the new generation of tenure candidates will be compared with a standard set by the good teachers. It is likely that these

[4]The judge's indifference curves are linear in Figure 1, but linearity is not necessary for the results.

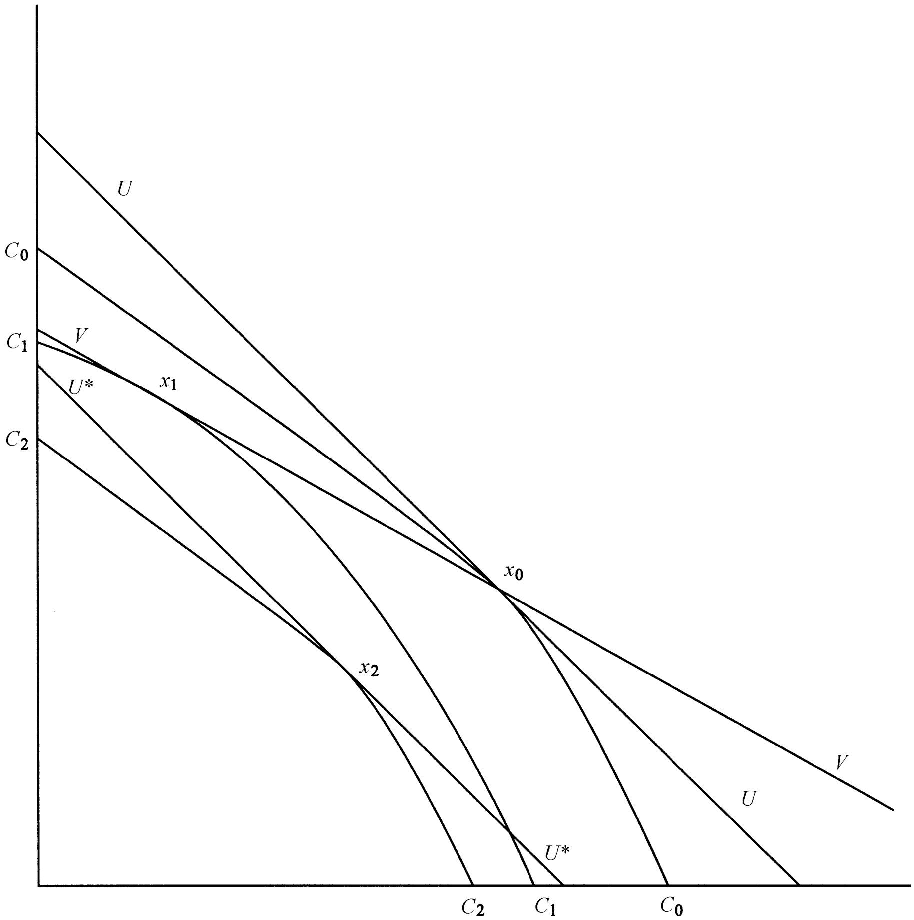

FIGURE 1

standards can be met by characteristics that would not have been satisfactory relative to the original population.

One might question how sensitive the basic result is to changes in assumptions. The formulation of the problem is stark. It involves only two functions: the utility function $U(\cdot)$ and the cost function $C(\cdot)$. Only perturbations in $U(\cdot)$ have been considered. What happens if one perturbs $C(\cdot)$ instead? Changing costs is likely to change the number of agents who meet the standard (i.e., those agents for whom the cost of meeting the standard is less than the net value of passing the test). In the most important sense, a temporary change in $C(\cdot)$ does not change standards at all. Provided that at least one candidate can meet the standard, the utility level that must be offered to the judges does not change when the cost function changes.

Hence, when the cost function returns to normal, the system will return to the original equilibrium.

Second, one could investigate the implications of assuming that different candidates had different cost functions, there were multiple solutions to Problem (3), or the elite population changed slowly. There would still be multiple equilibria. One would expect different performance levels to be represented in the elite. If the preferences of the judge do not change and the cost functions are increasing, then $U(x_t)$ does not depend on the particular characteristic x_t selected. When preferences change, however, the optimization problem (Problem 4) with the single constraint $V(x) \geq V(x_t)$ depends on the choice of x_t. Proposition 3 continues to hold if you replace Problem (4) with

$$(5) \quad \min C(x) \quad \text{subject to} \quad V(x) \geq V(y) \quad \text{for some } y \text{ in the previous population,}$$

which requires only that a candidate be as good as *some* member of the current elite to pass. Call this rule the *worst-case standard.* If the criterion for passing is more stringent than the constraint in Problem (5), there is no guarantee that perturbations will lead to falling standards.

There are two other extensions to Proposition 2 using the worst-case standard. First, suppose that candidates do not always select x to minimize the cost of meeting the standard. There may be private benefits to performing at a particular level that are not captured by entering the elite. For example, candidates may do more than is necessary to pass the test to be eligible for further credentials in the future, to win broader acclaim, or for self-esteem. Under these circumstances, the elite population will not be homogeneous. If the worst member of the elite determines the standard for future candidates, however, then all that is necessary is that there is at least one candidate for which the constraint binds in each period. If at least one candidate who is marginal in this sense passes in each period, then perturbations will reduce standards as in Proposition 2.

Similarly, the worst-case standard would lead to declining standards if there is a stochastic relationship between a candidate's effort choice and his or her performance provided that in every period at least one marginal candidate is promoted.

3. Conclusions

This article demonstrates that a particular rule for determining who should pass a test leads to erosion of standards. The result does not imply that decreasing standards are inevitable. Rather, it suggests a possibly surprising feature of grading or promotion procedures that contrasts with standard results on relative compensation schemes. The result may be useful in examining the performance of incentive schemes based on relative performance. There are two lessons for individuals seeking to design rules that maintain standards. First, when passing is based on relative standing, one should select the comparison group with care. Standards decline when the judge compares one group of candidates with one utility function against a group of elite that passed relative to another utility function. If the judge has preferences U, he or she should compare candidates with members of the elite who themselves were evaluated using U. Second, marginal passes make bad precedents. The use of a worst-case standard, which passes anyone who compares favorably with any successful candidate from the past, tends to erode standards.

The model does not permit welfare analysis. Whether decreasing standards are harmful either to society or to individuals requires a more elaborate model that completely specifies the preferences of the agents over the membership of the elite and the outside opportunities of failed candidates. In such a setting one could hope to use these preferences to construct optimal standards and then to study the implications of imposing these standards on the membership of the elite.

By concentrating on the implications of particular rules, I have obtained results that may be descriptive of some situations. I believe that the results are particularly relevant to situations in which competitive forces are not present and recent cases are used to determine outcomes.

REFERENCES

BARBERÀ, S., M. MASCHLER, AND J. SHALEV, "Voting or Voters: A Model of Electoral Evolution," CORE discussion paper, 1998.

BETTS, J., "The Impact of Educational Standards on the Level and Distribution of Earnings," *American Economic Review* 88 (1998), 266–75.

COSTRELL, R., "A Simple Model of Educational Standards," *American Economic Review* 84 (1994), 956–71.

GREEN, J., AND N. STOKEY, "A Comparison of Tournaments and Contracts," *Journal of Political Economy* 91 (1983), 349–64.

HOLMSTROM, B., "Moral Hazard in Teams," *Bell Journal of Economics* 13 (1982), 324–40.

LAZEAR, E., AND S. ROSEN, "Rank-Order Tournaments as Optimal Labor Contracts," *Journal of Political Economy* 89 (1981), 841–64.

LELAND, H. E., "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards," *Journal of Political Economy* 87 (1979), 1328–46.

NALEBUFF, B., AND J. STIGLITZ, "Prizes and Incentives: Toward a General Theory of Compensation and Competition," *Bell Journal of Economics* 14 (1983a), 166–78.

——— AND ———, "Information, Competition, and Markets," *American Economic Review*, *Papers and Proceedings* 73 (1983b), 278–83.

NATIONAL COMMISSION ON EXCELLENCE IN EDUCATION, *A Nation at Risk*: *The Imperative for Educational Reform* (Washington: U.S. Government Printing Office, 1983).

SHLEIFER, A., "A Theory of Yardstick Competition," *Rand Journal of Economics* 16 (1985), 319–27.

SCULL, A. T., P. CHURCHLAND, R. SALMON, "Report of the Committee to Review Trends in Promotion to Tenure," UCSD Academic Senate, La Jolla, 1995.

SOBEL, J., "On the Dynamics of Standards," UCSD discussion paper 1999.

WAGE (We Advocate Gender Equality) Newsletter, Berkeley, October 1993.