

WORKING PAPER

2002

Department of Economics
Tufts University
Medford, MA 02155
(617) 627 – 3560
<http://ase.tufts.edu/econ>

Survival of the Unfittest: How Dodos Become Managers

Dan Bernhardt
Department of Economics
University of Illinois
Champaign, IL 61820
danber@uiuc.edu

Eric Hughson
Leeds School of Business
University of Colorado
Boulder, CO 80309
eric.hughson@colorado.edu

Edward Kutsoati
Department of Economics
Tufts University
Medford, MA 02155
edward.kutsoati@tufts.edu

August 15, 2002

Abstract

This paper investigates the consequences of noisy evaluation of worker skills for skill investment and hiring. Individuals skew investment toward skills that most managers can best evaluate. In turn, this reinforces a tendency for managers to hire job applicants whose attributes they can best evaluate. In a dynamic context, where workers eventually become managers, managerial evaluation skills can become more and more skewed over time. Over the long haul, this has the effect of reducing both average employee productivity and the average quality of the job applicant pool. We distinguish two factors that affect a manager's ability to evaluate a particular skill: the manager's innate ability at that skill (nature), and his investment in the skill (nurture). We characterize how the dynamic evolution of the distributions of worker skill investments and worker productivity is affected by the relative importance of "nature" versus "nurture" for managerial skill evaluation. Surprisingly, investment distortions are *magnified* if some firms either (a) compensate managers with equity, rather than a share of profit, in order to induce managers to internalize the future evaluation consequences of their hiring decisions, or (b) strategically select the managerial expertise of their personnel department; unless both (i) identifying whether a worker is skilled matters more than distinguishing among skilled workers, and (ii) the initial investment distortions are sufficiently small. Finally, we show that short-run affirmative action policies can be effective only if hiring more multi-skilled workers matters more than distinguishing among workers according to a single skill.

1 Introduction

Suppose that employers are better able to evaluate some attributes of a job applicant's skills than others. One reason may be because certain attributes are easier to evaluate than others. Alternatively, those making the hiring decisions, having themselves been hired because they possess certain attributes, can better evaluate the same abilities in others. For instance, an evaluator who has good interpersonal skills, but is relatively weak analytically, may be more capable of evaluating an applicant's interpersonal skills than his analytical skills. An evaluator whose goal is to maximize expected firm profit may therefore hire those who are most skilled in the areas that he can evaluate. Consequently, if potential job applicants know what evaluators are looking for, they will attempt to develop those skills that can be best evaluated by employers, even if it is common knowledge that those skills contribute *less* to on-the-job productivity than others. In a dynamic context, where workers eventually become managers, the distribution of managerial evaluation skills can become more and more skewed over time, and average employee productivity can fall.

In this paper, we investigate the consequences of noisy evaluation of worker skills for skill investment and hiring in a simple overlapping generations framework. The current generation of managers hires the next generation of workers. There are two types of job applicant. Each type has a comparative advantage in developing one of two skills. Evaluation is noisy because each manager can evaluate only one of the skills. Taking this noisy evaluation procedure into account, job applicants then optimally make their skill investments. Those hired then become the next crop of managers, and make the next period's hiring decisions. The likelihood that a manager can evaluate a given skill depends on two factors: the manager's innate ability at that skill ("nature"), and his time investment in the skill ("nurture"). We characterize how the dynamic evolution of the distributions of worker skill investments and worker productivity is affected by the relative importance of nature versus nurture for managerial skill evaluation.

We first analyze a potential employee's static skill investment problem given a particular distribution of managerial evaluation skills. We show that agents skew skill investment toward the skill that most managers can evaluate, regardless of the potential employee's innate ability. This is not generally socially efficient, even when most managers can evaluate the more productive skill. In such an instance, all job applicants over-invest in acquiring the more productive skill. This reduces output both because of the decreasing marginal returns to investment in a given skill and because those with innate talents in acquiring the less productive skill inefficiently shift their investment away toward the other skill. As a result, both the expected productivity of the job applicant pool *and* expected employee productivity are lower than the social optimum.

We then investigate the equilibrium dynamics, when the current generation of workers becomes the next generation of managers. Investment distortions by workers introduce another dynamic source of inefficiency: If too many workers invest in the same skill, then subsequently, too few managers will be able to evaluate the other skill, leading to an increasingly overly-specialized distribution of managers. In turn, the overly-specialized distribution of managers increases the distortion in worker skill investments, so that over time, the situation worsens.

We characterize the limiting distributions of managerial skills and show that generally, these distributions feature over-specialization. The equilibrium steady state distributions of managerial (and worker) skills depend on whether managerial expertise (the skill that managers can evaluate) is determined primarily by the manager’s innate ability or primarily by the manager’s skill investment.

If managerial expertise depends principally on innate ability (“nature”), then there is a unique limiting distribution of managerial types. The proportion of managers of each type is determined by the relative productivity of the skills and the proportion of the population that possess innate ability in each. In contrast, if managerial expertise is learned, depending largely on the time invested in each skill (“nurture”), then there are multiple limiting distributions of managerial skills. Indeed, if “nurture” is the only determinant of managerial skills, then the only stable limiting distributions feature managers who are either all one type or all the other, depending on which type of manager initially predominates in hiring decisions. If nurture predominates, but innate ability also has an impact on the determination of managerial expertise, then the two skills co-exist in any steady state; and the two stable steady states feature mostly one type or mostly the other.

If nurture predominates, and the initial distribution of managers is relatively evenly balanced, then, over time, average employee productivity declines until the steady state is reached. If the initial distribution of managers is heavily skewed toward either skill, average employee productivity initially improves, but levels off at one of the steady states. In each case, corporate culture runs amok, as managers endogenously hire too many of their own type in equilibrium. The mix of managerial types that maximizes the expected productivity of those hired is more evenly balanced than the limiting distribution of evaluation skills. Hence, the mix of managerial types that obtains in any equilibrium differs from the mix that maximizes employee quality or output: generically, equilibria are inefficient. The distortion in skill investments due to the noise in the selection procedure is only magnified if the less productive skill dominates in the equilibrium.

In our base framework, firms take the managerial expertise of their personnel departments as exogenous, determined by past hiring decisions, and provide their managers the incentive to hire so as to maximize period profits. But, as we demonstrate, this leads to long-run distortions in the investment decisions of workers. One can imagine, therefore, that firms might be more strategic in

their design of compensation schemes or personnel departments.

Such considerations lead us to explore how the qualitative results are affected if some firms (i) strategically design their personnel departments, selecting the expertise of their managers to maximize firm value; or (ii) compensate managers with equity, rather than claims to the profits generated by the workers that they hire. When managers are compensated with equity, they internalize both the production of their hires, *and* their subsequent hiring decisions.

Surprisingly, in the long run, both forms of strategic behavior, while privately optimal, may be socially sub-optimal, actually *increasing* investment distortions, and reducing long-run output in the economy. To understand why, consider first a single firm that selects the expertise of its managers strategically, and suppose, for simplicity, that both skills contribute equally to output. Then, generically, this firm maximizes firm value by specializing in one type of manager. As long as there is sufficient dispersion in realized skills so that identifying the *level* of a worker's skill contributes more to firm profits than identifying *whether* a worker develops a skill, then a value-maximizing firm selects managers that are best at evaluating the skill that most workers acquire.

Only if the dispersion in realized worker skill levels is sufficiently small, is it possibly optimal to specialize in evaluating the skill that fewer workers acquire. Even then, the strategic firm will select managers who are best at evaluating the predominant skill if the measure of workers acquiring the predominant skill sufficiently exceeds the measure acquiring the other skill. That is, a profit-maximizing firm goes against the cultural tide only if the tide is not "too strong." In short, the strategic firm will specialize in the evaluation of the predominant skill unless: (i) the initial distribution of managerial expertise is not too skewed (so that investments are not too skewed); and (ii) there is relatively little dispersion in the talents of workers who succeed in acquiring skills, so that hiring more multi-skilled matters more than distinguishing among skilled workers.

Now suppose more firms strategically design their personnel departments. The above discussion implies that their presence improves long-run equilibrium outcomes only if they go against the tide. Importantly, in such a case, the presence of such strategic firms can make an efficient, expertise-balanced steady state locally stable even if managerial expertise is determined primarily by nurture.

The flip side to this observation is that if the strategic firms specialize in the evaluation of the predominant skill, then their presence only reinforces the investment distortions in the long-run steady states, and speeds the economy's convergence to these inefficient states. By specializing in the evaluation of the predominant skill, these strategic firms cause workers to further skew their investments, which, in turn, further skews the limiting distribution of managers. If all firms design their personnel departments, then as long as firms are not indifferent to their choice of managerial expertise, the resulting equilibria have all one type of manager or all the other. It is precisely such

equilibria that *minimize* expected employee output.

Consider now the consequences of compensating managers with equity rather than with a share of next period's profits. If managers are only compensated with next period's profits, then their hiring decisions will ignore the implications for future hiring decisions. If, instead, managers are compensated with equity, then they internalize the consequences for future hiring. If the current distribution of worker investments is skewed in one direction, then next period's distribution of managerial evaluation skills will be similarly skewed, causing the investments of workers in the next period to be skewed. Anticipating this, as long as there is sufficient dispersion in realized skill levels, managers who receive equity compensation will want to hire workers who are better able to evaluate the same skill in the next period, setting *slacker* productivity standards for such workers. Not only does this further skew the long-run distribution of managers, and hence further skew worker investments, but these firms also hire less productive employees than they might.

Our results also have implications for the efficacy of temporary affirmative action policies. For example, in an economy where most men are better able to develop one skill, while most women are better able to develop the other skill, a historical bias toward hiring men can be entrenched, as yesterday's male hires are today's evaluators. Our analysis reveals that if the mix of managers reflects the distribution of employees, then temporary affirmative action policies have no long run impact. If nurture predominates, then the "good" balanced steady state remains unstable, so that continued intervention would be required. If nature predominates, then while temporary affirmative action can speed or hinder convergence to the unique steady state, it cannot alter the ultimate outcome.

If, however, enough firms strategically design their personnel departments, then temporary affirmative action policies can have permanent impacts. A policy that mandates the increased hiring of women would increase the measure of managers who can evaluate the skills that most women develop. In turn, this more-balanced distribution of managerial expertise reduces investment distortions, *possibly* causing firms to specialize in the evaluation of the non-predominant skill, *i.e.* the skill that most women develop. In turn, this would lead to an even more balanced distribution of managerial expertise and investments so that the affirmative action policy becomes self-enforcing.

Our paper is related to work by Bradford Cornell and Ivo Welch (1996), and Susan Athey, Christopher Avery and Peter Zemsky (2000). Using a model similar to ours, Cornell and Welch (1996) show that noisy evaluation can lead to hiring discrimination even when employers do not intend to discriminate. Stephen Coate and Glenn Loury (1993), Andrea Moro and Peter Norman (1999) and Norman (1999) also explore how multiple equilibria can arise in static contexts when there is noisy evaluation and workers invest in skills. The basic idea is that output depends on a

worker's costly, unobserved human capital investment and job placement (skilled or unskilled): only trained workers are productive in the skilled position. Firms see a worker's sex and a noisy signal of whether the investment was made. While sex does not directly affect the cost of investment, if firms believe that, for example, women are less likely to invest than men, then firms are more likely to place women in positions where human capital is irrelevant, making investment by women less attractive, so that these beliefs are self-fulfilling. The antecedent to all of these papers is contribution by Edmund Phelps (1972). Phelps supposes that firms have noisier signals of some agents' (women, say) abilities than others, and hire to maximize profits, and derives the static implications. In some sense, the subsequent papers and ours serve to endogenize the noise in the signals and the consequences for investment behavior. In addition, our paper derives the dynamic and output implications.

Athey, Avery, and Zemsky (2000) provide a model of promotional policies that seeks to explain the evolution of diversity within a firm. In their paper, entry-level employees receive mentoring only from managers who are of their "type." This mentoring augments the ability of lower-level employees, increasing their chances of being promoted. Consequently, although the entering workers in a firm may be diverse, the higher echelons are not. In a similar vein, Canice Prendergast (1993) shows how subjective performance evaluation in a workplace can give rise to "yes men," workers acting to conform to the opinions of their supervisors. Our analysis extends this research agenda in that (i) we investigate how the distribution of managerial types affects skill acquisition decisions of potential employees, and (ii) derive the resulting consequences for the evolution of the distribution of managerial types over time.

Evidence of the effects of noisy evaluation processes can be found in David Granick's (1972) study of managerial selection across four countries. Granick found that, (i) the probability that a young manager is promoted is determined in large part by whether his skills coincide with those of the superior; and (ii) an employee may receive a high performance ratings under one superior only to see his ratings in the same post turn strongly negative after a change in the immediate boss. In a study of large and medium-sized U.S. firms, Michael Unseem (1989) observes that in some companies the proportion of managers with one or another skill was exceptionally high or low. For example, in some companies the "MBA culture" dominates. In the words of an inside observer, "They want their own type. The frame of reference here is the MBA" (Unseem p.125). The same study notes that if a CEO held an MBA degree, the chief operating officer is 2.06 times as likely to have an MBA as well. Hence, in a more general model of firm hierarchy, our theory can explain why engineers do not get promoted beyond a certain level without obtaining MBAs.

Finally, our paper complements the literature on uniform social behavior (*e.g.*, Thomas Schelling

(1978), Abijit Banerjee (1992), and Sushil Bikhchandani, David Hirshleifer and Ivo Welch (1992)). However, these models typically generate local conformity through positive payoff (co-ordination) externalities in an adopted technology (*e.g.*, Brian Arthur (1989, 1994), or a reputational effect that induces herding (*e.g.*, David Scharfstein and Jeremy Stein (1990), Jeffrey Zweibel (1995)). In contrast here, the incentive to get hired, coupled with noisy evaluation of job candidates, leads to the distortions in skill investments and resulting inefficiencies.

The paper is organized as follows: Sections 2 to 4 present the model, analysis and findings. Section 5 quantifies our findings numerically. Section 6 modifies the basic model to allow for (i) strategic design of firms' personnel departments, and (ii) compensating managers with equity, rather than claims to next period's profits. Section 7 concludes. All proofs are in an appendix.

2 The Model

Consider an overlapping generations model in which agents live for three periods. Each period, a continuum of measure one of agents are born. They are endowed with one unit of time that they can devote to developing two types of skills, *Analytical* skills and *Business Sense* skills, which we respectively label A and B for short. Training is costless. There is a continuum of measure one of long-lived atomistic firms. Each period, each firm can hire a measure ν of trained agents. In his second period of life, a trained agent becomes a worker if hired. Otherwise, the agent pursues a reservation alternative for the rest of his life. A trained worker with skill levels s_A, s_B produces

$$Y = a_A s_A + a_B s_B, \tag{1}$$

where $a_j > 0$ reflects the contribution of skill j to output. Without loss of generality, we assume that $a_A \geq a_B$ so that skill A is more valuable. We normalize the price of the output good to 1. All workers receive a common wage W . The wage W is high enough that potential employees value employment in this industry to their reservation alternative, but small enough that the expected profit from hiring the marginal worker at wage W is positive.

In their final period of life, workers become managers with expertise in *one and only one* skill and evaluate new job applicants who will then fill the measure ν of job openings. The mechanism for determining managerial expertise is detailed later. We first assume that managers are the residual claimants to the firm's period revenues, paying the new workers their wages out of the realized revenues generated by their work. That is, each manager receives an equal share of period profits from the hiring decisions that they make, $\Pi = E\{(Y - W)\nu \mid \text{evaluations}\}$. Consequently, a firm's managers collectively seek to hire to maximize next period's profits. In section 6, we consider the possibility that the previous period's managers sell the firm to the current period's management;

and that the current management will then sell these claims to the firm's profit stream to the generation of workers that it hires. In this case, a firm's management will make its hiring decisions to maximize share value/discounted firm profits.

The nature of compensation is sufficiently important to discuss further. The assumption that a worker's wage does not depend on his realized or observed skill levels simplifies the analysis because it ensures that workers will make their skill investments so as to maximize the probability that they are hired. This common wage level would follow if workers either did not observe their skill level or if it were non-verifiable to a third party, making it non-contractible—the firm would always assert that the worker had the skill level that minimized wage compensation. Workers care about total lifetime compensation, and the precise division between wage and profit is unimportant as long as the last worker hired generates a profit for the firm, and when a worker becomes a manager, the profit share dominates his reservation alternative.

2.1 Skill Development.

Agents have an innate propensity for developing one of the two skills. A fraction β , $0 < \beta < 1$, have natural ability in skill A (type- A agents); the remainder, $1 - \beta$, have natural ability in skill B (type- B agents). This innate propensity is captured by θ_j^i , which denotes a type- i worker's innate ability to develop skill j . For simplicity, we assume θ_j^i takes two values:

$$\theta_j^i = \begin{cases} H, & \text{if } j = i; \\ L, & \text{if } j \neq i, \end{cases}$$

where $0 < L < H < 1$.

A type- i agent who invests t_j^i attempting to acquire skill j develops the skill with probability

$$P(t_j^i, \theta_j^i) = \theta_j^i (t_j^i)^\psi,$$

where $0 < \psi < 1$. If the agent develops the skill, then his realized skill level, $s_j > 0$ is drawn from a distribution $F(\cdot)$, with associated positive density, $f(\cdot)$, on its bounded support, $[\underline{z}, \bar{z}]$. To ease the analysis, we assume (as in Cornell and Welch (1996)) that the realized skill level s_j does not depend on either his innate skill level, θ_j^i , or his time investment, t_j^i . An agent who does not acquire skill j has skill $s_j = 0$. This abstraction, which does not affect our qualitative findings, simplifies the analysis for two reasons. First, although expected productivity in a skill rises with the time spent developing the skill, realized skill levels reveal no information about the applicant's type. Therefore, managers do not use realized skill levels to make inferences about investments. Second, it simplifies a young agent's objective, because the probability of being hired takes a simple form.

2.2 Managerial Evaluation and Hiring.

Each manager can evaluate one and only one skill of an applicant. Type- A managers observe the level of skill A only, whereas type- B managers observe the level of skill B only. The current proportion of type- A managers in the economy is given by μ . For the most part, the distribution of managerial evaluation skills across firms is unimportant, as agents cannot selectively target firms, although it may help understanding to assume that the proportion of type A managers does not vary across firms. Each manager evaluates an equal number of job applicants, so that the measure of job applicants evaluated by type A managers in the economy is μ . For the static problem, it is unimportant how managerial ability is determined, and we take μ as given. Later, when we analyze the evolution of μ over time, we characterize how the skill a manager can evaluate depends on his innate ability and the skill investment choice.

The nature of managerial expertise is sufficiently important to elaborate upon. To ease the analysis, we make several simplifying assumptions: (i) each manager *perfectly* observes one skill and does *not* observe the other; (ii) each potential hire is evaluated by only one manager; and (iii) workers cannot selectively target firms (e.g. an agent with innate analytical ability cannot apply only to firms with personnel departments that can evaluate analytical skills).

Our qualitative findings are not sensitive to perturbations of the assumptions that preserve the qualitative feature that expertise in personnel departments is representative of the distribution of expertise in the economy. For example, investment distortions are reduced, but not eliminated (i) if some managers can observe both skills, but either some managers who hire observe only one skill, or managers receive noisier signals about one skill; (ii) if multiple managers can evaluate an agent, but either the probability of being evaluated only by a given managerial type rises with the measure of managers with expertise in that skill; or signals are noisy, and the worker expects to be evaluated more frequently by managers with the predominant managerial expertise; (iii) if selective targeting is not perfect, so that, *ceteris paribus*, the more firms there are with managers with expertise in one skill, the more likely a worker is to target such a manager.

Given their evaluations of applicants' realized skill levels, managers hire to maximize expected profits, Π , of the firm. Profit maximization implies that managers hire those applicants with greatest expected productivity given the observed skill levels.

We assume that the number of job openings in a firm, ν , is small enough that an agent is not hired if the manager who evaluates him finds him unskilled. This assumption ensures that a manager never wants to hire a worker who fails to acquire the skill that the manager can evaluate (*i.e.*, acquiring skills is a good thing, $E(Y|s_j = s > 0) > E(Y|s_j = 0)$). This assumption implies

that workers will devote all of their training time to acquiring skills, so that

$$t_A^i + t_B^i = 1. \quad (2)$$

Given this, it simplifies presentation to denote the time that a type- i agent invests in skill i (skill with comparative advantage) by t_i and $(1 - t_i)$ into the other skill. Therefore, t_A is type- A 's investment in skill A , and t_B is type- B 's investment in skill B .

Later, we provide a technical assumption that details conditions on primitives that ensure that no manager will, in fact, hire a worker who fails to acquire the skill that the manager can evaluate. Hence, an agent is hired only if he has a skill that the evaluating manager can observe, *and* if the skill level is sufficiently high to meet the cut-off for employment. Because realized skill levels do not reveal investment decisions, a type- i manager's inference about an agent's skill level in skill j depends on his beliefs $\{\tilde{t}_A, \tilde{t}_B\}$ about agents' time investments. After observing s_i , a type- i manager forms expectations about an agent's s_j , using beliefs \tilde{t}_A and \tilde{t}_B . From a type- i manager's perspective, the expected productivity, Y , of a job applicant who has obtained skill i is given by

$$\begin{aligned} E(Y | s_i > 0) &= a_i s_i + a_j E(s_j | s_i > 0) = a_i s_i + \Pr(\text{acquire skill } j | s_i > 0) a_j \bar{s} \\ &= a_i s_i + \frac{\Pr(\text{acquire both skills})}{\Pr(\text{acquire skill } i)} a_j \bar{s}, \end{aligned} \quad (3)$$

where $\bar{s} = E(s)$. That is, the manager's expectation of the applicant's productivity equals the realized productivity in the skill the manager can observe plus the expected productivity in the skill that the manager cannot observe.

Hence, the probability that a type- A manager hires a job applicant given that he observes $s_A > 0$ is equal to the probability that the applicant's expected productivity exceeds some endogenous hiring standard, $K(\nu, \mu)$,

$$\begin{aligned} Pr(A \text{ hires}) &= Pr\left(\left(a_A s_A + \frac{m_{AB}}{m_A} a_B \bar{s}\right) \geq K(\nu, \mu)\right) \\ &= Pr\left(s_A \geq \frac{K(\nu, \mu) - \frac{m_{AB}}{m_A} a_B \bar{s}}{a_A}\right) = Pr(s_A \geq K_A(\nu, \mu)), \end{aligned} \quad (4)$$

where

$$m_A = \beta H \tilde{t}_A^\psi + (1 - \beta) L (1 - \tilde{t}_B)^\psi$$

is the measure of agents believed to have acquired skill A , and

$$m_{AB} = HL[\beta \tilde{t}_A^\psi (1 - \tilde{t}_A)^\psi + (1 - \beta) \tilde{t}_B^\psi (1 - \tilde{t}_B)^\psi]$$

is the measure believed to have acquired both skills. $K(\nu, \mu)$ is the (endogenous) minimum equilibrium standard that a worker's observed productivity must exceed in order to be hired and

$K_A(\nu, \mu)$ is a type- A manager's hiring standard as a function of the level of skill A . Note that the standard for hiring, $K(\nu, \mu)$, depends on both the measure of job openings, ν , and the proportion of type- A managers, μ . It depends on μ because the current distribution of managerial skills affects optimal skill investments, and hence the distribution of realized skill levels. Similarly, the probability that a type- B manager hires a job applicant is given by

$$Pr(B \text{ hires}) = Pr\left(s_B \geq \frac{K(\nu, \mu) - \frac{m_{AB}}{m_B} a_A \bar{s}}{a_B}\right) = Pr(s_B \geq K_B(\nu, \mu)), \quad (5)$$

where

$$m_B = \beta L(1 - \tilde{t}_A)^\psi + (1 - \beta) H \tilde{t}_B^\psi$$

is the measure of applicants believed to have acquired skill B and $K_B(\nu, \mu)$ is the quality cut-off below which a type- B manager will not hire.

In equilibrium, the quality of the marginal worker, $K(\nu, \mu)$, is determined by setting the measure of those hired equal to the measure of job openings,

$$\mu m_A Pr(\text{manager A hires}) + (1 - \mu) m_B Pr(\text{manager B hires}) = \nu, \quad (6)$$

which is equivalent to

$$\beta Pr(\text{type-A applicant is hired}) + (1 - \beta) Pr(\text{type-B applicant is hired}) = \nu. \quad (7)$$

An Agent's Problem

Given the distribution of managerial types, μ , and the associated equilibrium hiring standards, $K(\nu, \mu)$, a type- A agent chooses t_A to maximize the probability of being hired:

$$\max_{t_A} Pr(A \text{ is hired}) = \max_{t_A} \mu Pr(A \text{ hires}) H t_A^\psi + (1 - \mu) Pr(B \text{ hires}) L(1 - t_A)^\psi, \quad (8a)$$

and a type- B agent chooses t_B to maximize the expected probability of being hired:

$$\max_{t_B} Pr(B \text{ is hired}) = \max_{t_B} \mu Pr(A \text{ hires}) L(1 - t_B)^\psi + (1 - \mu) Pr(B \text{ hires}) H t_B^\psi. \quad (8b)$$

The associated first order conditions are:

$$\mu Pr(A \text{ hires}) H(1 - t_A)^{1-\psi} = (1 - \mu) Pr(B \text{ hires}) L t_A^{1-\psi} \quad (9a)$$

and

$$(1 - \mu) Pr(B \text{ hires}) H(1 - t_B)^{1-\psi} = \mu Pr(A \text{ hires}) L t_B^{1-\psi}. \quad (9b)$$

Solving for t_A and t_B yields

$$t_A = \frac{(\mu H Pr(A \text{ hires}))^{\frac{1}{1-\psi}}}{(\mu H Pr(A \text{ hires}))^{\frac{1}{1-\psi}} + ((1 - \mu) L Pr(B \text{ hires}))^{\frac{1}{1-\psi}}} \quad (10a)$$

and

$$t_B = \frac{((1 - \mu)HPr(B \text{ hires}))^{\frac{1}{1-\psi}}}{(\mu LPr(A \text{ hires}))^{\frac{1}{1-\psi}} + ((1 - \mu)HPr(B \text{ hires}))^{\frac{1}{1-\psi}}}. \quad (10b)$$

Period Equilibrium

For any managerial skill distribution, μ , a (Perfect Bayesian) period equilibrium is a set of investments, $\{t_A^*(\mu), t_B^*(\mu)\}$; managerial beliefs, $\{\tilde{t}_A, \tilde{t}_B\}$; and hiring standards, $\{K_A^*(\nu, \mu), K_B^*(\nu, \mu)\}$ such that

- (i) Given conjectures $\hat{K}_A^*(\nu, \mu)$, $\hat{K}_B^*(\nu, \mu)$, $t_A^*(\mu)$ and $t_B^*(\mu)$ maximize (8a) and (8b) respectively.
- (ii) Given $t_A^*(\mu)$ and $t_B^*(\mu)$, $K_A^*(\nu, \mu) = \hat{K}_A^*(\nu, \mu)$ and $K_B^*(\nu, \mu) = \hat{K}_B^*(\nu, \mu)$ maximize expected managerial profits, Π .
- (iii) $t_A^*(\mu)$, $t_B^*(\mu)$, $K_A^*(\nu, \mu)$ and $K_B^*(\nu, \mu)$ solve equation (6).
- (iv) Managerial beliefs are consistent: $\{t_A^*(\mu), t_B^*(\mu)\} = \{\tilde{t}_A, \tilde{t}_B\}$.
- (v) Agents' beliefs are consistent: $\{K_A^*(\nu, \mu), K_B^*(\nu, \mu)\} = \{\hat{K}_A^*(\nu, \mu), \hat{K}_B^*(\nu, \mu)\}$.

Our first result characterizes the time each type of agent devotes to acquiring each skill, and how these investments depend on the proportion of managers, μ , that can evaluate type A skills. More importantly, Lemma 1 characterizes how the hiring standards by each managerial type depends on μ , and how the probability that a type A agent is hired depends on μ .

Lemma 1 *Period Equilibrium Characterization.*

- (a) *If no managers can evaluate a particular skill, agents optimally spend all their time developing the other skill: If $\mu = 0$, then $t_A = 0$ and $t_B = 1$. If $\mu = 1$, then $t_A = 1$, and $t_B = 0$.*
- (b) *Agents invest relatively more time in the skill in which they have a comparative advantage: $t_A \geq 1 - t_B$, with strict inequality if $\mu \in (0, 1)$.*
- (c) *The time invested by all agents in skill A (skill B) rises (falls) with μ .*
- (d) *The probability that a type- A agent (type- B agent) is hired rises (falls) with the proportion of managers that can evaluate skill A : $\frac{\mu Pr(A \text{ hires})}{(1-\mu)Pr(B \text{ hires})}$ is increasing in μ , so that the chance of being hired with skill A increases with μ . However, $\frac{Pr(A \text{ hires})}{Pr(B \text{ hires})}$ falls in μ .*
- (e) *$Pr(A \text{ is hired}) \in [\frac{\nu L}{\beta L + (1-\beta)H}, \frac{\nu H}{\beta H + (1-\beta)L}]$ and $Pr(B \text{ is hired}) \in [\frac{\nu L}{\beta H + (1-\beta)L}, \frac{\nu H}{\beta L + (1-\beta)H}]$.*

The key results in Lemma 1 are that (i) as the probability of being evaluated by a type- A manager increases, workers spend more time developing skill A at the expense of skill B ; (ii) the more type- A managers there are, the more workers who are hired by type- A managers; and (iii) the more type- A managers there are, the more likely a worker with innate ability at developing A skills is hired.

The more time that agents devote to acquiring skill A , the less time they have to give to acquiring skill B . Hence, as μ increases, the probability that an agent, observed by a type- A manager to have acquired skill A , also acquired skill B falls, and the probability that an agent, observed by a type- B manager to have acquired skill B , also acquired skill A rises. The consequence is that as μ increases, type- A managers set relatively higher hiring standards, and type- B managers set relatively lower standards.

Taken together, the results of Lemma 1 indicate that regardless of innate ability, economic agents tend to skew skill investment toward those skills on which they are more likely to be evaluated.

2.3 Transition to Managerial Expertise

To analyze the evolution of the distribution of managerial skills, and hence evaluation skills, we must determine how a worker's future managerial type is related to both worker type and innate ability. Recall that each manager can evaluate only one skill. We assume that the probability that a type- i worker becomes a type- i manager reflects both his innate ability as well as the time the worker devoted to acquiring skill i . Specifically, this probability is given by

$$T(i|i) = (1 - \gamma) + \gamma t_i,$$

so that the probability that he becomes a type- j manager is

$$T(j|i) = \gamma(1 - t_i).$$

Here, $\gamma \in [0, 1]$ captures the extent to which both “*nature*” (innate ability) and “*nurture*” (effort) determine managerial expertise. The higher is γ , the more important is “*nurture*” in determining managerial type.

Let the subscript t denote the time period. Then the distribution of managerial skills (parameterized by μ) evolves as follows: the fraction of type- A managers at time $t + 1$ as a function of the fraction at time t , $\mu_{t+1}(\mu_t)$ is given by

$$\mu_{t+1}(\mu_t) = \frac{1}{\nu} [\beta((1 - \gamma) + \gamma t_A)Pr_t(A \text{ is hired}) + (1 - \beta)\gamma(1 - t_B)Pr_t(B \text{ is hired})] \quad (11)$$

The first term on the right-hand side of equation (11) is the fraction of type- A workers that develop managerial expertise in evaluating skill A , and the second term is the fraction of type- B workers that develop expertise in evaluating skill A .

Assumption: The measure of job openings in the economy satisfies $\nu < \bar{\nu}$, where $\bar{\nu}$ solves

$$K\left(\bar{\nu}, \frac{1}{2}\right) = \frac{2(HL)^{\frac{1}{1-\psi}}}{(H^{\frac{1}{1-\psi}} + L^{\frac{1}{1-\psi}})^{1+\psi}}(a_A + a_B)\bar{s}.$$

This technical assumption is sufficient to ensure that a minimum standard is required for hiring.

3 Steady State

Existence of a Steady State

We now show that given any initial distribution of managerial types, independently of the relative importance of nature and nurture in determining managerial type, that the distribution of managerial types always reaches a steady state. That is, the equilibrium distribution of managerial types remains unchanged from one period to the next: $\mu_{t+1}(\mu_t) = \mu_t$. Since μ_{t+1} maps the unit interval $[0, 1]$ onto itself, existence of a steady state follows from standard fixed point arguments if μ_{t+1} is continuous in μ .

Lemma 2 $t_A(\mu)$, $t_B(\mu)$, $K_A(\nu, \mu)$ and $K_B(\nu, \mu)$ are (jointly) continuous in μ . Furthermore, μ_{t+1} is continuous in μ_t .

Proposition 1 *There exists at least one steady state in μ ($0 < \mu < 1$).*

Characterization of Steady States

We now characterize the steady-state distributions of managerial skills. An important question is whether or not the initial distribution of managerial skills affects the steady-state distributions. We show that if and only if managerial expertise depends sufficiently on skill selection (nurture), then the steady state distribution depends critically on the initial distribution of managerial skills.

Case (i): Expertise Determined solely by “Nature”

When managerial type is determined solely by nature, workers with an innate ability in a particular skill can, as managers, observe only that skill. Then, the steady state proportion of managers, μ^* , depends on the distribution of innate abilities in the population, the abilities of each type to learn both skills, and the productivity of each ability, but is independent of the initial distribution of managerial types, μ_0 .

Proposition 2 *If “nature” is the sole determinant of managerial expertise, i.e. if $\gamma = 0$, then there is a unique and stable steady state at some $\mu = \mu^*$.*

If the current proportion of type- A managers, μ_t , is less than μ^* , then agents shift investments toward developing skill B because there is a high probability of being evaluated by a type- B manager. However, the fraction of type- A agents hired exceeds the current proportion of type- A managers. Because type- A workers become type- A managers, it follows that the proportion of type- A managers next period, $\mu_{t+1} = \frac{\beta}{\nu} Pr_t(A \text{ is hired})$, rises. Similarly, if $\mu_t > \mu^*$, then $\mu_{t+1} < \mu_t$.

It is straightforward to show that μ^* is an increasing function of both $(a_A - a_B)$ and β . To see this, note that if $\gamma = 0$, then fixing a_A, a_B , given any μ_t , $\mu_{t+1}(\mu_t)$ is an increasing function of β . So, too, fixing β , for any μ_t , $\mu_{t+1}(\mu_t)$ is an increasing function of $(a_A - a_B)$.

If agents are equally likely to be of either type and the skills are equally valuable, then $\mu^* = \frac{1}{2}$. Thus, if $a_A > a_B$, and $\beta = \frac{1}{2}$, then $\mu^* > \frac{1}{2}$ —when managerial expertise is determined solely by nature, all else equal, in the steady state, there are more managers with expertise in the more efficient skill. If, instead, the skills are equally valuable, then the managerial type that predominates in the steady state corresponds to the innate ability that predominates in the population: $\mu^* < \frac{1}{2}$ if $\beta < \frac{1}{2}$ and $\mu^* > \frac{1}{2}$ if $\beta > \frac{1}{2}$. When $a_A > a_B$, if $\beta > \frac{1}{2}$, then μ^* consists of an even higher proportion of type- A managers. However, if $a_A > a_B$, and $\beta < \frac{1}{2}$, then the type of manager that predominates in the steady state depends on whether or not the *efficiency* effect (how much a_A exceeds a_B) dominates the *size* effect (how small is β).

Case (ii): Expertise Determined solely by “Nurture”

When managerial type is determined solely by “nurture”, the probability that a worker becomes a type- i manager is given by the proportion of time he spent developing skill i . We have already shown that in this instance, $\mu = \{0, 1\}$ are steady states (Lemma 1(a)). In fact, these are the only stable steady states:

Proposition 3 *If “nurture” is the sole determinant of managerial expertise ($\gamma = 1$), then there are three steady states: there are stable steady states at $\mu = \{0, 1\}$ and an unstable steady state at some $\mu = \mu^s$ ($0 < \mu^s < 1$).*

If the current distribution of managerial evaluation skills is such that $\mu_t < \mu^s$, then both agent types shift their time investments toward skill B . As a result, there are more type- B managers in the next generation, and the proportion of type- A managers falls over time: $\mu_{t+1} < \mu_t < \mu^s$. (See also Lemma 3 below). In the long run, all managers are of type B and hence, by Lemma 1(b), no agent invests in skill A .

Case (iii): “Nature” versus “Nurture”

The number of equilibrium steady states depends critically on whether nature predominates (γ

is small), or whether nurture predominates (γ is large). When “nature” predominates, there is a unique equilibrium steady state. However, if γ is large enough, we will show that there exist two stable steady states. In one, agents skew investment toward skill A and in the other, they skew investment toward skill B .

Lemma 3 *There exists $0 < \mu^T < 1$, such that for $\mu \underset{(<)}{>} \mu^T$,*

$$\beta(1 - t_A)Pr(A \text{ is hired}) \underset{(>)}{<} (1 - \beta)(1 - t_B)Pr(B \text{ is hired}).$$

Call μ^T the cross-over distribution of managerial skills. Lemma 3 implies that if $\gamma > 0$, then starting from a managerial distribution $\mu \underset{(<)}{>} \mu^T$, the measure of type- B (type- A) workers who ‘switch’ to become type- A (type- B) managers exceeds the measure of type- A (type- B) workers who become type- B (type- A) managers. Hence, there will be more type- A ’s in the next generation of the managerial pool, relative to the previous generation, if $\mu > \mu^T$. Formally,

Corollary 1 *Given μ_t , $\mu_{t+1}(\mu_t)$ is decreasing in γ if $\mu_t < \mu^T$, and increasing in γ if $\mu_t > \mu^T$.*

Corollary 1 is useful for characterizing steady states as γ varies between 0 and 1. If there exists a steady state $\mu^* \neq \mu^T$, then $|\mu^* - \mu^T|$ increases as γ approaches 1.

Proposition 4 *There exists some $\hat{\gamma}$, such that there exists a unique and stable steady state at some μ^* if $\gamma < \hat{\gamma}$. For any $\gamma > \hat{\gamma}$, the economy has three steady states: stable steady states at some $\mu = \{\mu^B(\gamma), \mu^G(\gamma)\}$ and an unstable steady state at some $\mu^s(\gamma) \in (\mu^B(\gamma), \mu^G(\gamma))$. Also, $\mu^B(\gamma)$ is decreasing in γ and $\mu^G(\gamma)$ is increasing in γ .*

Proposition 4 shows that if sufficient weight is attached to nurture in the determination of managerial skills, then the steady state distribution of managerial skills may feature a higher proportion of either type- A or type- B managers. Which type predominates depends on the initial managerial distribution: for $\gamma > \hat{\gamma}$, if the initial proportion, μ_0 , of type- A managers is less than $\mu^s(\gamma)$, then workers invest more in skill B , and in the limit, $\mu_\infty = \mu^B(\gamma) < \mu^s(\gamma)$. Put more starkly, the “unfittest” not only survive, but in the long run, they dominate, and the bad steady state persists. If instead, $\mu_0 > \mu^s(\gamma)$, then $\mu_\infty = \mu^G(\gamma) > \mu^s(\gamma)$.

As we shall see, both limiting distributions are inefficient relative to the optimal distribution of managerial skills because excessive effort is spent developing skills that can be evaluated by managers. Corporate culture runs amok as too many of one type of worker are hired. Average worker productivity declines, even when the more efficient skill predominates in equilibrium. When the less efficient skill predominates, average worker quality is further reduced.

4 Efficiency

We now show that the limiting distributions of our economy are generically inefficient. Our criterion for comparing efficiency across steady states is expected worker productivity, or, equivalently, average worker quality. This criterion also corresponds to expected industry output and profits. Employee quality is given by

$$EQ(\mu) = \Pr(\text{hired by A})E(Y^i|\text{hired by A}) + \Pr(\text{hired by B})E(Y^i|\text{hired by B}), \quad (12)$$

where

$$E(Y^i|\text{hired by A}) = a_A E \left[s_A \mid s_A \geq K^A(\nu, \mu) \right] + \frac{m_{AB}}{m_A} a_B \bar{s} \quad (13)$$

and

$$E(Y^i|\text{hired by B}) = a_B E \left[s_B \mid s_B \geq K^B(\nu, \mu) \right] + \frac{m_{AB}}{m_B} a_A \bar{s}. \quad (14)$$

The average quality of the applicant pool is

$$\begin{aligned} AQ(\mu) &= \Pr(\text{type-A})E(Y^i|i \text{ is type-A}) + \Pr(\text{type-B})E(Y^i|i \text{ is type-B}) \\ &= \beta(a_A H t_A^\psi + a_B L(1 - t_A)^\psi) \bar{s} + (1 - \beta)(a_A L(1 - t_B)^\psi + a_B H t_B^\psi) \bar{s}. \end{aligned} \quad (15)$$

Note that a higher average applicant pool quality need not imply a higher average employee quality, as employee quality also reflects the ability of managers to identify worker skills.

To maximize worker productivity, the distribution of managers must provide agents with the right skill investment incentives. Each distribution of evaluation skills, μ , induces best responses $t_A(\mu)$, $t_B(\mu)$, and hiring cutoffs $K_A(\mu, \nu)$ and $K_B(\mu, \nu)$. Let μ^o be the mix of managers that maximizes expected employee productivity: $\mu^o = \operatorname{argmax}_\mu EQ(\mu)$. The optimal mix, μ^o , depends on all of the primitives describing the economy, except for γ .

To demonstrate that equilibrium steady states are generically inefficient, we show that in every economy, generically, there is at most a single value of γ that leads to a steady state in which expected employee productivity is maximized.

Proposition 5 *For any $(a_A - a_B)$, there exists at most one value of β that can support the optimal managerial mix, μ^o , for more than one value of γ .*

It is worth noting that if the economy is symmetric, *i.e.*, $a_A = a_B$ and $\beta = 0.5$, then $\mu_\infty = 0.5$ is one equilibrium (but not necessarily stable) steady state for all values of γ , and μ_∞ is also the efficient distribution.

These results help us characterize the consequences of increased “nurture” for long-run efficiency. Generically, increases in γ may first lead to improvements in the average quality of workers in

the limiting distribution, but eventually, increases in γ must lead to a more unbalanced mix of managerial evaluation skills than that which maximizes industry output.

5 Example

We now illustrate our results numerically. Suppose that $P(t, \theta) = \theta\sqrt{t}$, $\theta \in \{0.3, 0.6\}$; half the population has an innate ability to acquire skill A , and half has an innate ability to acquire skill B , $\beta = 0.5$; and skill A is more productive than skill B , with $a_A = 1$ and $a_B = 0.9$. There is a measure $\nu = 0.1$ of job openings, and realized skills s are distributed uniformly on $[0, 1]$.

Table 1: Equilibrium μ_∞ as a Function of Nature vs. Nurture

γ	μ_0	μ_∞	t_A	t_B	Expected Employee Quality	Expected Applicant Quality
0		.548	.885	.674	.917	.471
0.75		.889	.994	.088	.910	.378
0.9	< .276	.043	.028	.998	.818	.368
0.9	= .276	.276	.534	.933	.886	.455
0.9	> .276	.962	.999	.011	.897	.329
1	< .317	0	0	1	.800	.203
1	= .317	.317	.613	.910	.892	.462
1	> .317	1	1	0	.889	.225
0.455		.701	.957	.418	.925	.450

NOTES: Equilibrium steady states as a function of γ (nurture) and initial proportion of type- A managers (μ_0). μ_∞ is the limiting proportion of type- A managers and t_A and t_B represent the limiting amounts of time spent by types A and B in acquiring the skills in which they have a comparative advantage. The last row shows the managerial mix ($\mu_\infty = 0.701$) that maximizes expected employee productivity. This is only an equilibrium when $\gamma = 0.455$. In this example, $P(t, \theta) = \theta\sqrt{t}$, $\theta \in \{0.3, 0.6\}$; skill A is more productive than skill B , $a_A = 1 > a_B = 0.9$, $\beta = 0.5$, $\nu = 0.1$, and realized skills, s , are distributed uniformly on $[0, 1]$.

Table 1 calculates equilibrium steady states as a function of γ (“nurture”). μ_0 is the initial fraction of type- A and μ_∞ is the limiting distribution. The bottom line of Table 1 gives the managerial mix, $\mu = 0.701$, that maximizes expected employee quality. This optimal mix is supported as a steady state only by $\gamma = 0.455$.¹ Lesser values of γ have too few type- A managers in the steady

¹Note that $\gamma = 0.455$ does not maximize the expected quality of the applicant pool.

state, while greater values of γ have too many type- A managers (if the steady state is unique); or either too many or grossly too few if there are multiple steady states. For $\gamma \leq 0.75$, there is a unique steady state that is reached independently of the initial distribution of managers, μ_0 .

For $\gamma > 0.75$, there are three steady states: an unstable steady state (with a managerial distribution that depends on γ), and the two stable steady states. When “nurture” predominates so there are multiple steady states, then the stable steady states feature over-specialization, which reduces quality of both the employee- and applicant-pool, *even* when the efficient skill dominates. That is, even if skill A dominates in the steady state, both agent types under-invest in the acquisition of skill B . Note how both the average quality of the job applicant pool and average employee quality are affected by the distribution of managerial types. The more skewed is the distribution of managerial types, the greater is the effort distortion of potential employees. Loosely speaking, the greater is effort distortion, the fewer potential employees actually acquire skills, so that less-skilled workers are hired. The average quality of the workforce is lower when skill investment is more distorted, and lower still when the distortion is toward the less-efficient skill B .

Figure 1 illustrates the evolution of the proportion of type- A managers and average managerial quality. Notice that the curves $\mu_{t+1}(\gamma = 0)$ and $\mu_{t+1}(\gamma = 1)$ intersect each other at $\mu = \mu^T$. Recall from Lemma 3 that at μ^T , the measure of type- A agents that become type- B managers is exactly equal to the measure of type- B hires that become type- A managers in the next period. All steady states, both stable and unstable ones, occur where the $\mu_{t+1}(\gamma)$ curves intersect the 45° line. The curve $\mu_{t+1}(\gamma = 1)$ intersects the 45° line at $\mu = 0$, 0.317 , and $\mu = 1$. If managerial expertise is completely determined by nurture and $\mu_0 < 0.317$, then eventually all managers will be of type- B even though skill B is less efficient. This is the ‘worst’ steady state in terms of average productivity: expected employee quality is 0.80 and the quality of the average job applicant is a mere 0.203 .

Figure 2 shows the dynamics of investments into skill A for $\gamma = 1$ when the initial distribution is greater than the unstable steady state (*i.e.*, $\mu_0 = 0.317 + \epsilon$, for small $\epsilon > 0$). The greater is the percentage of type- A managers, the greater is the investment by both types of worker into development of skill A . Figure 2 also displays expected worker quality as a function of μ . As indicated in the bottom line of the table, maximum quality occurs at $\mu = 0.701$. This is not a steady state value unless $\gamma = 0.455$. Although it is possible that the initial distribution of managerial expertise is such that the evolution of managerial expertise approaches the optimal mix, in which case average worker quality initially improves, the process continues. Generically, corporate culture runs amok in the long run, as both managerial expertise and worker skill investments become overly-specialized.

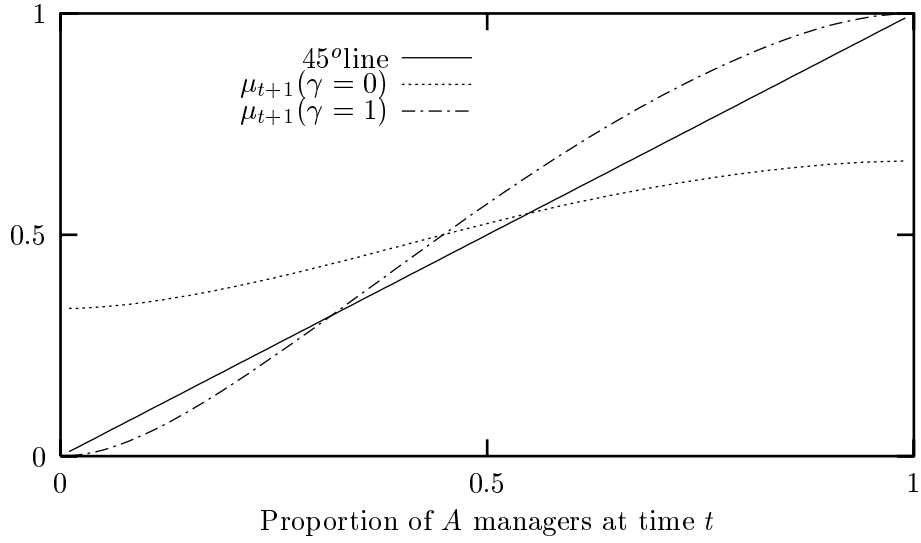


Figure 1: The dynamic evolution of managerial skills

NOTES: μ_{t+1} is plotted against μ_t for $\gamma = 0$ and $\gamma = 1$. A steady state is represented by intersection with the 45° line. It can be seen that, for $\gamma = 1$, if $\mu < (>)0.317$, then the limiting distribution consists of only type-*B*'s (type-*A*'s).

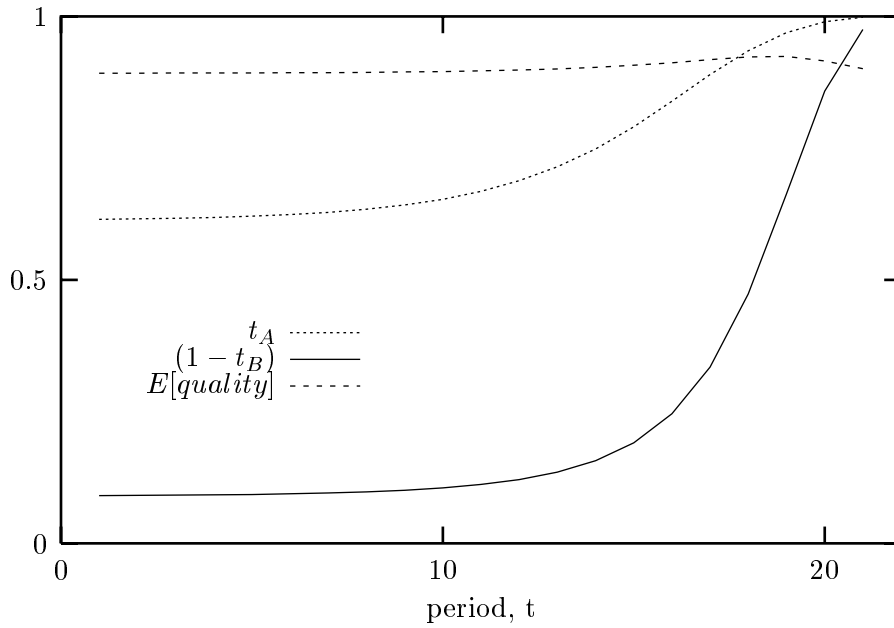


Figure 2: Investments in Skill A

NOTES: Dynamics of investments into skill *A* and employee quality, when $\gamma = 1$ and the initial distribution, μ , is greater than the unstable steady state (*i.e.*, $\mu_0 = 0.317 + \epsilon$). Employee quality increases initially, then fall as $\mu \rightarrow \mu_\infty$.

6 Strategic firm behavior

Heretofore, we assumed that firms make hiring decisions to maximize profits taking as given the distribution of managerial expertise in their personnel department. Specifically, a particular firm's managerial expertise at date t , μ_t , reflects past hiring decisions, which, in turn, depend on the realized distribution of managerial expertise generated by hiring in the previous period. Proposition 5 says us that generically, this eventually leads to inefficient outcomes: the equilibrium limiting distribution of managerial expertise induce sub-optimal investment decisions by employees.

But suppose now that rather than taking the distribution of managerial expertise as exogenous, some firms design their personnel departments strategically, selecting the mix of managerial expertise that maximizes firm profits. If a firm chooses managerial expertise period by period, then maximizing period profit also maximizes firm value. The questions that we now address are: (i) what distribution of managerial types would an *optimizing* firm choose in order to maximize expected profits? and (ii) what are the consequences for long-run expected employee productivity and investment decisions in the economy if a significant fraction of firms chose their personnel department design?

6.1 Optimal personnel department

We suppose that firms use only one manager to evaluate a job applicant, and that firms have no information about individual applicants prior to investigating, so that they cannot target managerial expertise to particular applicants. Although these are clearly simplifying assumptions, our qualitative results extend to reasonable permutations of these assumptions. Generically, the expected profit from observing one skill exceeds that from observing the other. It follows that a strategic firm will choose a personnel department that specializes in observing one skill or the other—the firm will choose to observe the skill that maximizes expected employee productivity. To see this, note that because each firm is small, an individual firm's personnel department decisions do not affect workers' investments. Further, there are constant returns to scale in managerial expertise, because each manager evaluates job applicants independently. Hence, if it is optimal for one manager to observe skill A , then it is optimal for all managers to investigate that skill.

The question then becomes which skill would firms prefer to have their managers observe? In particular, under what conditions does a firm have an incentive to specialize in the skill that workers devote most of their time toward developing?

To answer this question, we first consider a single strategic firm and suppose that $m_j > \nu$ so that there are more job applicants with skill j than slots (if not, the value of investigating skill j is

further reduced). Then, a firm that observes skill j will hire the top ν/m_j percent of applicants who develop skill j . Hence, if the firm uses managers with expertise in evaluating skill A , the average output per employee is

$$E[Y^i | \mu = 1] = \frac{m_A}{\nu} a_A \int_{F^{-1}\left(1 - \frac{\nu}{m_A}\right)}^{\bar{z}} s f(s) ds + \frac{m_{AB}}{m_A} a_B \bar{s}. \quad (16)$$

If, instead, the firm uses managers with expertise in evaluating skill B , average worker output is

$$E[Y^i | \mu = 0] = \frac{m_B}{\nu} a_B \int_{F^{-1}\left(1 - \frac{\nu}{m_B}\right)}^{\bar{z}} s f(s) ds + \frac{m_{AB}}{m_B} a_A \bar{s}. \quad (17)$$

Define Δ to be the difference in expected productivity from evaluating skill A rather than skill B :

$$\Delta \equiv \left(\frac{m_A}{\nu} a_A \int_{F^{-1}\left(1 - \frac{\nu}{m_A}\right)}^{\bar{z}} s f(s) ds - \frac{m_B}{\nu} a_B \int_{F^{-1}\left(1 - \frac{\nu}{m_B}\right)}^{\bar{z}} s f(s) ds \right) - (m_A a_A - m_B a_B) \frac{m_{AB}}{m_A m_B} \bar{s}. \quad (18)$$

The firm will therefore select managers with expertise in evaluating skill A if and only if $\Delta \geq 0$.

Consider first a symmetric environment in which each skill contributes equally to output, $a_A = a_B = a$ and there are equal measures of potential hires with innate abilities to acquire each skill, $\beta = 0.5$. Then, if there are equal proportions of managers with each skill expertise in the economy, $\mu = 0.5$, then $m_A = m_B$ and it does not matter which skill is investigated. However, if there are more managers with expertise in skill A , $\mu > 0.5$, then more workers acquire skill A than skill B .² As a result, the first term in Δ (*i.e.* the difference in integrals) is positive, but the second term is negative. The tradeoff between investigating skill A and skill B is then the tradeoff between the two terms. Qualitatively,

1. The first term captures the fact that if a manager inspects skill A , then since $m_A > m_B$, the manager sees more candidates who acquire a skill. The expected quality of the top $\frac{\nu}{m_A}$ percent of the distribution exceeds the expected quality of the top $\frac{\nu}{m_B}$ percent.
2. Because $m_A > m_B$, someone who is observed to have skill B is more likely to have acquired both skills than someone who is observed to acquire skill A . Hence, for any observed skill level, someone observed to acquire skill B is expected to be more productive than his skill A counterpart.

Inspection of equation (18) reveals that the sign of Δ depends crucially on the dispersion in the distribution of realized skill levels relative to the mean skill acquired, \bar{s} . Increasing the dispersion

²So, too, more workers acquire skill A than B if $\mu = 0.5$, but $\beta > 0.5$.

in realized employee skills (*e.g.* by a mean preserving spread of the upper tail of $F(\cdot)$), then the first term rises relative to the second. As a result, if there is sufficient dispersion, $\Delta > 0$, so that the firm should investigate the skill A that most employees develop. Intuitively, when employees vary significantly in ability, identifying which have high ability becomes more important. Consequently, the value of seeing more applicants with ability rises, so the value of staffing the personnel department with managers who can evaluate the skill that most applicants acquire rises.

If instead, the distribution of realized skills lacks sufficient dispersion, then $\Delta < 0$. In particular, letting $\underline{z} \rightarrow \bar{z} > 0$, the first on the RHS of (18) goes to zero, so that $\Delta < 0$, implying that the optimal personnel department design only features managers with expertise in the skill B that fewer workers acquire. Intuitively, if there is little dispersion in skills, identifying multi-skilled applicants contributes more to output than distinguishing among applicants whose skills can be evaluated. Hence, it is optimal to swim against the corporate tide and staff the personnel department with management that can evaluate the non-predominant skill, the one that most firms do not evaluate.

To glean further insight into the optimal personnel department design, we explore the impact of a marginal increase in the measure of managers with skill A expertise on Δ :

$$\begin{aligned} \frac{\partial \Delta}{\partial \mu} &= \frac{1}{\nu} [E_A(\mu) - K_A^*(\mu)] \frac{\partial m_A}{\partial \mu} - \frac{1}{\nu} [E_B(\mu) - K_B^*(\mu)] \frac{\partial m_B}{\partial \mu} \\ &\quad - a\bar{s} \left[\frac{m_{AB}}{m_A m_B} \left(\frac{\partial m_A}{\partial \mu} - \frac{\partial m_B}{\partial \mu} \right) + (m_A - m_B) \frac{\partial \left(\frac{m_{AB}}{m_A m_B} \right)}{\partial \mu} \right], \end{aligned} \quad (19)$$

where $K_j^*(\mu) = aF^{-1}(1 - \frac{\nu}{m_j})$ is the lowest observed skill hired by a manager evaluating skill j , and $E_j(\mu) = a \int_{K_j^*(\mu)}^{\bar{s} + \bar{z}} s f(s) ds$ is the average observed skill of a hire by this manager. Thus, $E_j(\mu) - K_j^*(\mu)$ is the expected gain from observing another potential hire who has acquired skill j : the firm replaces workers at the productivity standard with those whose expected productivity exceeds the standard. The gain from investigating the predominant skill is captured by the difference between the first two terms: as μ rises, m_A rises and m_B falls, causing workers to shift their investments toward skill A , so that managers see more workers who acquire skill A , and less who acquire skill B . The opportunity cost of investigating the predominant skill is captured by $a \frac{m_{AB}}{m_A m_B} \bar{s}$; because applicants skew investments toward the predominant skill, an applicant who is known to have acquired the non-predominant skill is more likely to be multi-skilled. The greater is \bar{s} — the expected skill component that the manager cannot evaluate — the greater is this opportunity cost of identifying fewer multi-skilled hires.

At least when skills are uniformly distributed, $\frac{\partial \left(\frac{m_{AB}}{m_A m_B} \right)}{\partial \mu} \bar{s} (m_A - m_B)$ is almost zero (slightly

positive), so that it can be ignored, leaving

$$\frac{\partial \Delta}{\partial \mu} \sim \frac{\frac{\partial m_A}{\partial \mu}}{\nu} [E_A(\mu) - K_A^*(\mu)] - \frac{\frac{\partial m_B}{\partial \mu}}{\nu} [E_B(\mu) - K_B^*(\mu)] - \left(\frac{\partial m_A}{\partial \mu} - \frac{\partial m_B}{\partial \mu} \right) a \frac{m_{AB}}{m_A m_B} \bar{s} \quad (20)$$

In the neighborhood of $\mu = \frac{1}{2}$, $m_A \sim m_B \sim m$, so that $\frac{\partial \Delta}{\partial \mu}$ simplifies to

$$\frac{\partial \Delta}{\partial \mu} \Big|_{\mu=0.5} = \left[\frac{E_A(0.5) - K_A^*(0.5)}{\nu} - \frac{m_{AB}(0.5)}{m^2} a \bar{s} \right] \left(\frac{\partial m_A}{\partial \mu} - \frac{\partial m_B}{\partial \mu} \right) \quad (21)$$

If there is little dispersion in acquired skills relative to the average acquired skill, \bar{s} , then for μ slightly greater than $\frac{1}{2}$, the second term dominates the first. This implies that $\Delta < 0$ so that it is optimal to use managers who can investigate the non-predominant skill: *i.e.*, identifying talent is more important than distinguishing among talented people. If, instead, there is sufficient dispersion in the realized skill distribution relative to \bar{s} then $\Delta > 0$ so that it is optimal to go with the flow and use managers who can evaluate the predominant skill.

Now consider what happens to $\frac{\partial \Delta}{\partial \mu}$ as μ is increased sufficiently past $\frac{1}{2}$. Because $\frac{m_{AB}}{m_A m_B}$ almost does not vary with μ , the derivative of the second term of Δ with respect to μ is essentially

$$- \left(\frac{\partial m_A(\mu)}{\partial \mu} - \frac{\partial m_B(\mu)}{\partial \mu} \right) \frac{m_{AB}(0.5)}{m^2} a \bar{s}.$$

In contrast the first term of Δ is a more convex function of μ , rising at a more rapid rate than

$$\left(\frac{\partial m_A(\mu)}{\partial \mu} - \frac{\partial m_B(\mu)}{\partial \mu} \right) \left(\frac{E_A(0.5) - K_A^*(0.5)}{\nu} \right).$$

This is because, as μ rises, $E_A(\mu)$ and $m_A(\mu)$ rise, while $E_B(\mu)$ and $m_B(\mu)$ fall, so that the positive derivatives in the first term of Δ , $\frac{\partial m_A(\mu)}{\partial \mu}$ and $\frac{\partial E_A(\mu)}{\partial \mu}$ are multiplying ever bigger terms, while the negative derivatives, $\frac{\partial m_B(\mu)}{\partial \mu}$ and $\frac{\partial E_B(\mu)}{\partial \mu}$ are multiplying ever smaller terms. As a result, as μ rises, eventually the first term of Δ dominates the second. Thus, if dispersion in realized skills is not too great, it is optimal to investigate the non-predominant skill as long as μ is not too great. However, if too many managers have expertise in the same skill, applicants' human capital investments are sufficiently distorted in that direction so that few acquire the non-predominant skill, consequently making it optimal to investigate the predominant skill. If, instead, employees differ substantially in their realized skills (relative to the mean), then it is optimal to investigate the predominant skill, independent of the value of μ , as Δ rises monotonically with μ .

When realized skills are uniformly distributed on $[\bar{s} - \bar{y}, \bar{s} + \bar{y}]$, we can highlight formally the impact of increasing the skill dispersion on the range of μ for which it is optimal to investigate the non-predominant skill. Define μ^* such that $\Delta(\mu^*) \equiv 0$ for $\mu^* > 0.5$. That is:

$$m_A(\mu^*) \left[\frac{E_A(\mu^*)}{\nu} - \frac{m_{AB}(\mu^*)}{m_A(\mu^*) m_B(\mu^*)} a \bar{s} \right] - m_B(\mu^*) \left[\frac{E_B(\mu^*)}{\nu} - \frac{m_{AB}(\mu^*)}{m_A(\mu^*) m_B(\mu^*)} a \bar{s} \right] \equiv 0.$$

We now show that as dispersion in realized skills is increased, it becomes optimal to investigate the predominant skill for ever lower values of μ . Equivalently, the range of $\mu > \frac{1}{2}$ for which it is optimal to investigate the non-predominant skill shrinks as the skill dispersion is increased.

To see this, differentiate $\Delta(\mu^*)$ with respect to the dispersion parameter, \bar{y} . This yields:

$$\begin{aligned} & \left[\left[\frac{E_A(\mu^*) - K_A^*(\mu^*)}{\nu} - \frac{am_{AB}(\mu^*)\bar{s}}{m_A(\mu^*)m_B(\mu^*)} \right] \frac{\partial m_A}{\partial \mu^*} - \left[\frac{E_B(\mu^*) - K_B^*(\mu^*)}{\nu} - \frac{am_{AB}(\mu^*)\bar{s}}{m_A(\mu^*)m_B(\mu^*)} \right] \frac{\partial m_B}{\partial \mu^*} \right] \frac{\partial \mu^*}{\partial \bar{y}} \\ & \sim -\frac{m_A(\mu^*)}{\nu} \left(\frac{(\bar{s} + \bar{y})}{2\bar{y}} - \frac{(\bar{s} + \bar{y})^2 - [K_A(\mu^*)]^2}{4\bar{y}^2} \right) + \frac{m_B(\mu^*)}{\nu} \left(\frac{(\bar{s} + \bar{y})}{2\bar{y}} - \frac{(\bar{s} + \bar{y})^2 - [K_B(\mu^*)]^2}{4\bar{y}^2} \right) \\ & = -\frac{m_A(\mu^*)}{\nu} \left(\frac{\bar{y}^2 + [K_A(\mu^*)]^2 - \bar{s}^2}{4\bar{y}^2} \right) + \frac{m_B(\mu^*)}{\nu} \left(\frac{\bar{y}^2 + [K_B(\mu^*)]^2 - \bar{s}^2}{4\bar{y}^2} \right) < 0. \end{aligned}$$

The expressions inside the square brackets multiplying $\frac{\partial \mu^*}{\partial \bar{y}}$ must be positive because of the U -shaped relationship of Δ with μ . This, in turn, implies that $\frac{\partial \mu^*}{\partial \bar{y}} < 0$, as required.

To summarize the above analysis, two central theoretical conclusions obtain:

- Fixing the dispersion in realized skills, \bar{y} , there exists a $\hat{\mu}(\bar{y}) \geq 0.5$ such that it is optimal to investigate the predominant skill A that most workers acquire if $\mu \geq \hat{\mu}(\bar{y})$; and it is optimal to investigate the non-predominant skill B if $0.5 < \mu < \hat{\mu}(\bar{y})$.
- Fixing the fraction μ of managers in the economy with expertise in evaluating skill A , there exists a dispersion parameter $\bar{y}(\mu)$ such that for $\bar{y} \geq \bar{y}(\mu)$, it is optimal to investigate predominant skill A that most workers acquire; and it is optimal to investigate the non-predominant skill B if $\bar{y} < \bar{y}(\mu)$.

To understand the quantitative tradeoffs, we next numerically solve for the optimal skill to investigate for different parameters. Table 2 details Δ for different values of μ and different supports of the distribution of realized skill levels when $\beta = 0.5$ and $a_A = a_B = 1$. Fixing μ , the proportion of managers skilled in evaluating skill A , as we increase the dispersion parameter \bar{y} , it becomes more attractive to investigate the predominant skill A . As a result, with sufficient dispersion, $\Delta > 0$, so that if $\mu > 0.5$, it is optimal to have a personnel department that evaluates skill A . Conversely, evaluating the non-predominant skill is better if there is sufficiently little dispersion in realized skills. It is worth observing that our analysis assumes that there are no complementarities between skills A and B . Obviously, such complementarities would serve to raise the attraction of investigating the skill that most workers do not acquire.

If we instead fix the dispersion in realized skills and vary μ , we see that as long as the dispersion in realized skills is not too high, then Δ is a U -shaped function of μ , decreasing initially in μ , but

then increasing for μ sufficiently high (if μ exceeds 0.9, eventually $\Delta > 0$, even if there is minimal dispersion). Note also, as revealed by our theoretical analysis, that as the dispersion in realized skills is increased, the interval of values of μ exceeding $\frac{1}{2}$ for which it is optimal to investigate the non-predominant skill shrinks.

Table 2: Δ : symmetric environment.

\bar{y}	μ : Proportion of managers skilled in A					
	0.5	0.51	0.6	0.7	0.8	0.9
0.1	0	-0.0009	-0.009	-0.020	-0.034	-0.053
	(0.335, 0.335)	(0.337, 0.333)	(0.355, 0.314)	(0.374, 0.288)	(0.396, 0.250)	(0.423, 0.183)
	[0.072, 0.640]	[0.072, 0.640]	[0.071, 0.643]	[0.07, 0.652]	[0.066, 0.670]	[0.054, 0.702]
0.2	0	-0.0007	-0.007	-0.016	-0.023	-0.011
	(0.335, 0.335)	(0.338, 0.332)	(0.361, 0.306)	(0.385, 0.270)	(0.409, 0.221)	(0.433, 0.144)
	[0.072, 0.640]	[0.072, 0.640]	[0.071, 0.645]	[0.068, 0.659]	[0.062, 0.684]	[0.045, 0.716]
0.3	0	-0.0003	-0.003	-0.005	0.001	0.058
	(0.335, 0.335)	(0.338, 0.332)	(0.364, 0.302)	(0.391, 0.260)	(0.415, 0.205)	(0.438, 0.125)
	[0.072, 0.640]	[0.072, 0.640]	[0.071, 0.646]	[0.067, 0.665]	[0.059, 0.692]	[0.040, 0.722]
0.4	0	0.0002	0.003	0.009	0.031	0.140
	(0.335, 0.335)	(0.339, 0.332)	(0.366, 0.299)	(0.394, 0.254)	(0.419, 0.195)	(0.440, 0.114)
	[0.072, 0.640]	[0.072, 0.640]	[0.071, 0.647]	[0.067, 0.668]	[0.057, 0.696]	[0.036, 0.725]
0.5	0	0.0008	0.009	0.024	0.064	0.231
	(0.335, 0.335)	(0.339, 0.332)	(0.368, 0.297)	(0.396, 0.250)	(0.421, 0.189)	(0.441, 0.107)
	[0.072, 0.640]	[0.072, 0.640]	[0.071, 0.648]	[0.066, 0.670]	[0.056, 0.699]	[0.034, 0.727]

The table details Δ as a function of μ and the dispersion in skill distribution. Below Δ are the measures acquiring each skill, (m_A, m_B) ; and the measures $[m_{AB}, \frac{m_{AB}}{m_A m_B}]$ in square brackets. The probability of acquiring a skill is $P(t, \theta) = \theta \sqrt{t}$, and $\theta \in \{0.3, 0.6\}$. We assume $a_A = a_B = 1$, $\nu = 0.1$ and $\beta = 0.5$. Realized skills are distributed uniformly on $[1 - \bar{y}, 1 + \bar{y}]$.

Qualitatively identical conclusions arise if we instead introduce asymmetry through the measure of job applicants with innate ability in developing a particular skill, or through the relative productivities of the two skills. Panel A of Table 3 provides the values of Δ when $\beta = 0.55$, (keeping $a_A = a_b = 1$); and panel B reports Δ when skill A is more productive, $1 = a_A > a_b = 0.99$. If skill dispersion is not too great, then even though $\beta = 0.55$ implies that more workers have an innate ability to acquire skill A , it is optimal to investigate skill B , as long as not too many firms employ managers with expertise in evaluating skill A . So, too, even though skill A may be more productive, if μ and skill dispersion are not too great, it is optimal to investigate the less productive skill that fewer workers acquire. Again, in this case identifying identifying workers who are more likely to be

multi-skilled contributes more to profits than identifying the *level* of the more productive skill.

Table 3: Δ : asymmetric environment.

\bar{y}	μ : Proportion of managers skilled in A					
	0.5	0.51	0.6	0.7	0.8	0.9
	Panel A: More agents with innate ability in A : $\beta = 0.55$					
0.1	-0.005	-0.004	-0.014	-0.024	-0.038	-0.056
0.2	-0.004	-0.004	-0.011	-0.019	-0.024	-0.009
0.3	-0.002	-0.001	-0.004	-0.005	0.003	0.066
0.4	0.001	0.001	0.004	0.012	0.037	0.156
0.5	0.005	0.004	0.014	0.031	0.076	0.255
	Panel B: Skill A more efficient: $a_A = 1 > a_B = 0.99$					
0.1	0.006	0.005	-0.003	-0.014	-0.027	-0.045
0.2	0.009	0.008	0.001	-0.007	-0.013	0.001
0.3	0.011	0.010	0.007	0.006	0.013	0.073
0.4	0.011	0.012	0.015	0.022	0.045	0.158
0.5	0.013	0.014	0.022	0.038	0.080	0.251

The table details Δ , as a function of μ and the dispersion in skill distribution. In panel A, $\beta = 0.55$; and in panel B, skill A is more efficient, $a_A = 1 > a_B = .99$. Also, $P(t, \theta) = \theta\sqrt{t}$, and $\theta \in \{0.3, 0.6\}$. The probability of acquiring a skill is $P(t, \theta) = \theta\sqrt{t}$, and $\theta \in \{0.3, 0.6\}$. There is a measure $\nu = 0.1$ of slots and realized skills, s , are distributed uniformly on $[1 - \bar{y}, 1 + \bar{y}]$.

6.1.1 Equilibrium when all firms design personnel departments

If all firms design personnel departments optimally, then there are, essentially, three equilibria, *regardless* of whether managerial type is determined primarily by nature or nurture.³ Two equilibria feature all firms having all one managerial type or all the other. It does not pay for a single firm to deviate and evaluate the other skill because by Lemma 1(a), potential employees, facing a distribution of managers that is all one type, put all their efforts into developing that type of skill. As highlighted before, these extreme equilibrium distributions of managerial expertise resulting when firms choose personnel departments optimally lead to workforces with low expected productivity. In the third equilibrium, $\mu = \mu^c$, so that firms are indifferent about which skill to observe — the average employee productivity does not vary with the skill that is evaluated.

³A caveat is that the firms must have enough of the appropriate type of manager to fill their personnel departments.

6.1.2 Equilibrium when some firms design personnel departments.

Suppose finally that some fraction $\rho < 1$ of firms design their personnel departments strategically. If ρ is sufficiently large, then coordination equilibria of the type discussed in section 6.1.1 obtain. This is because the managerial expertise that the strategic firms choose to observe affects which skill is predominant in equilibrium (which skill agents skew skill acquisitions toward). However, as long as nature plays a role in the selection of managerial type, the equilibrium limiting distribution of managerial types, μ_∞ , features managers with both skills.

When ρ is not too large to affect which skill is predominant, and the weak regularity conditions provided above are satisfied (proposition 6 holds), the results can be counterintuitive. If the current distribution of managers, μ , is sufficiently unbalanced, all optimizing firms have their managers observe the predominant skill. When there are multiple steady states, the presence of these optimizing managers reduces the long-run quality of the workforce. This follows because the presence of these optimizing firms causes applicants to skew skill investment toward the predominant skill still further, resulting in stable steady states that become more and more extreme as the fraction of optimizing managers grows. However, if the current distribution of managerial types is sufficiently close to the symmetric “nature” equilibrium, then optimizing firms observe the “other” skill, and there is convergence to the “nature” equilibrium.

6.1.3 Affirmative Action and Personnel Department Design

Consider now the implications of our results for the efficacy of temporary affirmative action policies. Suppose that both analytical and business sense skills contribute equally to output, but that a fraction $\beta_w > 0.5$ of women have an innate ability to develop analytical skills, and a lesser fraction $\beta_m = 1 - \beta_w$ of men have an innate ability to develop business skills. Suppose further that historically more men than women were employed, so that the distribution of managerial skills is skewed toward evaluating the business sense skills that men are more likely to develop, $\mu < 0.5$. If nurture underlies managerial expertise, then a historical bias toward hiring men can be entrenched, as yesterday’s male hires are today’s evaluators, and male managers tend to have expertise at evaluating the business sense skills that men have a relative advantage at developing.

In this context, it is important to understand the long-run consequences of temporary affirmative action plans that mandate firms increase the measure of female hires for a *finite* number of periods. Suppose first that personnel department designs at each firm correspond to the distribution of expertise determined by their hiring decisions. Then, temporary affirmative action policies have no long run impact. If nature drives the expertise of management, a temporary affirmative action plan

may speed the economy toward its long-run steady state, but it cannot alter long-run outcomes. If, instead, nurture predominates, then the “good” balanced steady state remains unstable, so that affirmative action would have to be targeted perfectly in order to alter long-run outcomes.⁴

If, however, enough firms strategically design their personnel departments, then temporary affirmative action policies can have permanent impacts. A policy that mandates the increased hiring of women would increase the measure of managers who can evaluate the skills that most women develop. In turn, this more-balanced distribution of managerial expertise reduces investment distortions, *possibly* causing firms to specialize in the evaluation of the non-predominant skill, *i.e.* the skill that most women develop. In turn, this would lead to an even more balanced distribution of managerial expertise and investments so that the affirmative action policy becomes self-enforcing.

6.2 Managerial compensation

In our model, each generation of managers receives the profit generated by its new hires. Consequently, managers make hiring decisions to maximize the expected immediate productivity of their workforce, and do not consider the future implications for hiring decisions that will be made by the new hires. Suppose instead, that at the end of each period, the managers sell the firm to the current young workers who will be next period’s managers. Young workers have a discount factor $\delta < 1$. Then managers will hire to maximize the sum of current period profit plus the share price that they can get from the current young workers. Since future generations of managers will have the same considerations, this implies that managers will hire to maximize the discounted sum of profits (using discount factor δ): managers will hire to maximize firm value. Would it then make sense to hire workers based on the managerial expertise they might have, rather than on single-period productivity? Put another way, would an optimal hiring practice show concern not only for expected productivity, but also for expected managerial hiring skill?

Again our analysis of the optimal personnel department design characterizes when it is optimal to shade hiring of workers toward the skill that more workers will develop (*i.e.*, sufficient dispersion in realized skill levels); and when it is optimal to hire against the grain (*e.g.*, lesser dispersion

⁴This discussion is not altered if managers internalize the fact that women are more likely to have innate ability at skill B, and men are more likely to have innate ability at skill A. Then, a manager evaluating skill A will set a lower standard for women, and a higher one for men, reflecting the fact that a woman with skill A is more likely to have also developed skill B, than is a man with skill A. The fewer managers evaluating skill B will do the opposite. In turn, this will lead women to increase their investments in skill A and lead men to increase their investments in skill B. Clearly, this standard difference will lead to more women being hired, and hence moderate the stationary stable states (if nurture is sufficiently important). However, these observations in no way alter the efficacy of temporary affirmative action: again when nature predominates, affirmative action is irrelevant in the long run, and if nurture predominates, affirmative action would have to be targeted perfectly at the unstable steady state to matter.

in realized skill levels; managerial expertises sufficiently balanced in the economy, so that human capital investments are not too distorted). If the current distribution of worker investments is skewed in one direction, then next period's distribution of managerial evaluation skills will be similarly skewed, causing the investments of workers in the next period to be skewed. Anticipating this, as long as there is sufficient dispersion in realized skill levels, managers who receive equity compensation will favor workers who are better able to evaluate the predominant skill in the next period, sacrificing worker productivity for the evaluation gains. In this case, compensating managers with equity, while optimal for the individual firm, have the collective effect of increasing the (in)efficiencies in skills investments: investments into skill acquisition is further skewed resulting in these strategic firms hiring less productive employees than they might.

If, instead, skill levels are less dispersed (see Tables 2 and 3), and the distribution of managerial expertise in the economy is not too skewed, then the firm will distort hiring toward those workers who are better able to evaluate the non-predominate skill. That is, firms will give up a little in worker productivity in order to develop a personnel department that will do a better job in identifying employees who will acquire both skills.

7 Conclusion

This paper illustrates the consequences of partial evaluation of skills for skill investment and hiring. In a dynamic framework where workers become managers, we show that if managerial expertise in a skill depends to a large extent on the time invested in that skill, managerial skills become more and more skewed over time. At best, when the good skill predominates in the long run, average worker quality declines. At worst, the bad skill predominates, and average worker quality falls even further. Over time, from an efficiency standpoint, agents make investment choices that are increasingly inefficient, reducing the quality of the entire job applicant pool.

Surprisingly, efforts by firms to constitute their personnel departments to arrest the decline in worker quality can be counter-productive. Whenever the initial distribution of managerial skills is sufficiently unbalanced, the ability of firms to choose their personnel departments to maximize expected employee productivity harms expected worker productivity throughout the economy.

Still, when the initial distribution of managerial skills is sufficiently balanced, then firms *can* constitute their personnel departments to arrest the decline in worker quality. In this case, a one-time intervention to balance the distribution of managerial types could result in a long-run outcome with higher expected worker productivity. That is, the intervention to balance managerial types “takes,” because it is reinforced over time by the equilibrium behavior of the optimizing firms.

In a more general firm hierarchy, the paper potentially explains why engineers are not promoted beyond a certain level without acquiring an MBA. It also potentially explains the existence of management training programs and the necessity of changing young managers' jobs relatively frequently so that many superiors have a chance to evaluate them.⁵

The paper can be viewed as a parable for American hiring practices over the years: the greater and ever-increasing emphasis on hiring MBAs. One conjectures that for some idiosyncratic reason, hiring was initially done by people in personnel departments that had few technical people. So to get hired, technically-minded students consequently acquired MBAs to develop business, rather than technical, skills. Those with business skills were hired in greater numbers and subsequently, personnel departments became more and more specialized in business skills. This further reinforced students' incentives to acquire MBAs. In the end, technical skills were weeded out, not only from the population of those hired, but also from the applicant pool, as even the technically-minded ceased to acquire engineering degrees.

⁵Granick (1972) observed that many superiors share in the evaluation of young managers. Young managers' jobs are changed sufficiently frequently so that each manager has different superiors over relatively short periods.

Appendix

Proof of Lemma 1:

(a) If all evaluators are of type A (B), utility maximization implies that agents invest only in skill A (B). Hence, if $\mu = 0$, $t_A = 0$ and $t_B = 1$. Similarly, if $\mu = 1$, $t_B = 0$ and $t_A = 1$.

(b) For $0 < \mu < 1$, subtracting the second first-order condition from the first yields:

$$\mu Pr(A \text{ hires}) \left[H(1 - t_A)^{1-\psi} - Lt_B^{1-\psi} \right] = (1 - \mu) Pr(B \text{ hires}) \left[Lt_A^{1-\psi} - H(1 - t_B)^{1-\psi} \right] \quad (Ap.1)$$

Suppose that $t_A < 1 - t_B$ ($\Rightarrow t_B < 1 - t_A$). Then $H(1 - t_A)^{1-\psi} - Lt_B^{1-\psi} > 0$ and $Lt_A^{1-\psi} - H(1 - t_B)^{1-\psi} < 0$. That is, the left-hand and right-hand sides of (Ap.1) have opposite signs. From part (a), if μ equals zero or one, $t_A = 1 - t_B$.

We prove the second part of (b) by contradiction. Suppose t_A is decreasing in μ ($\Rightarrow t_B$ is increasing in μ). Then, from the first first-order condition, on the left-hand side, μ and $1 - t_A$ both increase. On the right-hand side, $1 - \mu$ and t_A both decrease. For equality to hold as μ increases, $Pr(A \text{ hires})$ must decrease and $Pr(B \text{ hires})$ must increase. This follows because the number of workers hired must stay the same.

But if t_A decrease as μ increases, agents spend less time developing skill A and more time developing skill B . This means that for a given skill realization, s , a type- A manager must set a slacker standard, because given the skill realization, an agent is now more likely to have developed skill B as well. Hence, the probability that A hires rises, a contradiction.

(c) From part (b), the fact that t_A and $1 - t_B$ are increasing in μ implies that $\frac{Pr(A \text{ hires})}{Pr(B \text{ hires})}$ is decreasing in μ . However, from the first order conditions, the ratio of the measures of workers hired by managers of each type is

$$\frac{\mu Pr(A \text{ hires})}{(1 - \mu) Pr(B \text{ hires})} = \frac{Lt_A^{1-\psi}}{H(1 - t_A)^{1-\psi}}$$

which is increasing in μ since t_A is increasing in μ . Therefore, because the number of applicants hired is constant, the probability of being hired with skill A is increasing in μ .

(d) $Pr(A \text{ is hired}) = \mu Pr(A \text{ hires}) H t_A^\psi + (1 - \mu) Pr(B \text{ hires}) L (1 - t_A)^\psi$. From part (c), $\mu Pr(A \text{ hires})$ increases relative to $(1 - \mu) Pr(B \text{ hires})$ when μ increases. Since t_A also increases in μ , it follows that $Pr(A \text{ is hired})$ increases in μ . Similarly, $Pr(B \text{ is hired})$ decreases in μ .

Proof of Lemma 2: Consider equations (10a) and (10b). Fix $Pr(A \text{ hires})$ and $Pr(B \text{ hires})$, so that t_A and t_B are continuous in μ . Then $K_A(\nu, \mu) = \frac{1}{a_A} (K(\nu, \mu) - \frac{m_{AB}}{m_A} a_B \bar{s})$ and $K_B(\nu, \mu) = \frac{1}{a_B} (K(\nu, \mu) - \frac{m_{AB}}{m_B} a_A \bar{s})$ are also continuous in μ since they are continuous functions of t_A and t_B . Hence $Pr(A \text{ is hired})$ and $Pr(B \text{ is hired})$ are continuous in μ , and μ_{t+1} is also continuous in μ .

Proof of Proposition 1: Define $G(\mu_t) = \mu_{t+1} - \mu_t$. From equation (11),

$$\begin{aligned} G(0) &= \frac{\beta(1-\gamma)\nu L}{\beta L + (1-\beta)H} > 0 \quad (= 0 \text{ if } \gamma = 1) \\ \text{and } G(1) &= -\frac{(1-\gamma)(1-\beta)\nu L}{\beta H + (1-\beta)L} < 0 \quad (= 0 \text{ if } \gamma = 1). \end{aligned}$$

Since μ_{t+1} is continuous in μ_t (Lemma 2), there exists some μ such that $G(\mu) = 0$.

Proof of Proposition 2: If $\gamma = 0$, then equation 11 becomes $\mu_{t+1} = \frac{\beta}{\nu} Pr_t(A \text{ is hired}) = \frac{\beta}{\nu} (\mu H t_A Pr(A \text{ hires}) + (1-\mu) t_B Pr(B \text{ hires}))$. In this case, $\mu_{t+1}(\mu = 0) > 0$ and $\mu_{t+1}(\mu = 1) < 0$. To prove uniqueness (and stability), it will suffice to show that $\frac{\partial \mu_{t+1}}{\partial \mu_t} < 1$ at a steady state.

Substituting t_A and t_B from equations (10a) and (10b) into the equation for μ_{t+1} above gives

$$\mu_{t+1} = \frac{\beta}{\nu} \left[(\mu H Pr(A \text{ hires}))^{\frac{1}{1-\psi}} + ((1-\mu) L Pr(B \text{ hires}))^{\frac{1}{1-\psi}} \right]^{1-\psi}.$$

Differentiating w.r.t μ_t at the steady state, we get

$$\begin{aligned} \frac{\partial \mu_{t+1}}{\partial \mu_t} &= \frac{\frac{\beta}{\nu} \left[(\mu H Pr(A \text{ hires}))^{\frac{1}{1-\psi}} \frac{\frac{\partial \mu Pr(A \text{ hires})}{\partial \mu}}{\mu Pr(A \text{ hires})} + ((1-\mu) L Pr(B \text{ hires}))^{\frac{1}{1-\psi}} \frac{\frac{\partial (1-\mu) Pr(B \text{ hires})}{\partial \mu}}{(1-\mu) Pr(B \text{ hires})} \right]}{\left((\mu H Pr(A \text{ hires}))^{\frac{1}{1-\psi}} + ((1-\mu) L Pr(B \text{ hires}))^{\frac{1}{1-\psi}} \right)^{-\psi}} \\ &= \mu \left[t_A \frac{\frac{\partial \mu Pr(A \text{ hires})}{\partial \mu}}{\mu Pr(A \text{ hires})} + (1-t_A) \frac{\frac{\partial (1-\mu) Pr(B \text{ hires})}{\partial \mu}}{(1-\mu) Pr(B \text{ hires})} \right] \end{aligned}$$

But, by Lemma 1(c), $\frac{\frac{\partial \mu Pr(A \text{ hires})}{\partial \mu}}{\mu Pr(A \text{ hires})} > \frac{\frac{\partial (1-\mu) Pr(B \text{ hires})}{\partial \mu}}{(1-\mu) Pr(B \text{ hires})}$. Therefore

$$\frac{\partial \mu_{t+1}}{\partial \mu_t} < \frac{Pr(A \text{ hires}) + \mu \frac{\partial Pr(A \text{ hires})}{\partial \mu}}{Pr(A \text{ hires})} \leq 1$$

since $\frac{\partial Pr(A \text{ hires})}{\partial \mu} \leq 0$ (see proof of Lemma 1(b)). Hence, for $\gamma = 0$, there is a unique and stable steady at some μ^* .

Proof of Proposition 3: The existence of steady states at $\mu = \{0, 1\}$ is shown in Lemma 1. To prove stability, it suffices to show that $\frac{\partial \mu_{t+1}}{\partial \mu_t} < 1$ when $\mu = 0$ and 1. We use the convention that the derivatives at $\mu = 0$ and $\mu = 1$ are the right-hand and left-hand derivatives respectively.

Differentiating the first-order conditions with respect to μ gives,

$$\begin{aligned} H(1-t_A)^{1-\psi} [Pr(A \text{ hires}) + \mu \frac{\partial Pr(A \text{ hires})}{\partial \mu}] - H(1-\psi) \frac{\mu Pr(A \text{ hires})}{(1-t_A)^\psi} \frac{\partial t_A}{\partial \mu} = \\ -L t_A^{1-\psi} [Pr(B \text{ hires}) - (1-\mu) \frac{\partial Pr(B \text{ hires})}{\partial \mu}] + L(1-\psi) \frac{(1-\mu) Pr(B \text{ hires})}{t_A^\psi} \frac{\partial t_A}{\partial \mu} \end{aligned}$$

Multiplying through by t_A^ψ , we get

$$(1 - \psi)[\mu Pr(A \text{ hires})H\left(\frac{t_A}{1 - t_A}\right)^\psi + L(1 - \mu)Pr(B \text{ hires})]\frac{\partial t_A}{\partial \mu} = Ht_A^\psi(1 - t_A)^{1-\psi}\left[Pr(A \text{ hires}) + \mu\frac{\partial Pr(A \text{ hires})}{\partial \mu}\right] + Lt^\psi[Pr(B \text{ hires}) - (1 - \mu)\frac{\partial Pr(B \text{ hires})}{\partial \mu}]$$

Since we are interested in $\frac{\partial t_A}{\partial \mu}$ as $\mu \rightarrow 0$, we let μ be as close as possible to zero. The right-hand side of the last equation vanishes because $t_A \rightarrow 0$ as $\mu \rightarrow 0$. Hence, $\frac{\partial t_A}{\partial \mu} \rightarrow 0$ if $\mu \rightarrow 0$. Similarly,

$$Lt_B^{1-\psi}\left[Pr(A \text{ hires}) + \mu\frac{\partial Pr(A \text{ hires})}{\partial \mu}\right] + (1 - \psi)\mu Pr(A \text{ hires})\frac{L}{t_B^\psi}\frac{\partial t_B}{\partial \mu} = -H(1 - t_B)^{1-\psi}\left[Pr(B \text{ hires}) - (1 - \mu)\frac{\partial Pr(B \text{ hires})}{\partial \mu}\right] - (1 - \psi)(1 - \mu)Pr(B \text{ hires})\frac{H}{(1 - t_B)^\psi}\frac{\partial t_B}{\partial \mu}$$

Multiplying through by $(1 - t_B)^\psi$, we get

$$(1 - \psi)\left[\mu Pr(A \text{ hires})L\left(\frac{1 - t_B}{t_B}\right)^\psi + H(1 - \mu)Pr(B \text{ hires})\right]\frac{\partial t_A}{\partial \mu} = -Lt_B^{1-\psi}(1 - t_B)^\psi\left[Pr(A \text{ hires}) + \mu\frac{\partial Pr(A \text{ hires})}{\partial \mu}\right] - H(1 - t_B)[Pr(B \text{ hires}) - (1 - \mu)\frac{\partial Pr(B \text{ hires})}{\partial \mu}]$$

and letting μ be close enough to zero yields $\frac{\partial t_B}{\partial \mu} \rightarrow 0$ since $t_B \rightarrow 1$ as $\mu \rightarrow 0$.

When $\gamma = 1$, equation 11 becomes:

$$\mu_{t+1} = \frac{1}{\nu} [\beta Pr_t(A \text{ is hired}) - \beta(1 - t_A)Pr_t(A \text{ is hired}) + (1 - \beta)(1 - t_B)Pr_t(B \text{ is hired})]$$

and

$$\begin{aligned} \frac{\partial \mu_{t+1}}{\partial \mu_t} &= \frac{1}{\nu} \left[\beta \frac{\partial Pr(A \text{ is hired})}{\partial \mu} - \beta(1 - t_A) \frac{\partial Pr(A \text{ is hired})}{\partial \mu} + \beta Pr(A \text{ is hired}) \frac{\partial t_A}{\partial \mu} \right. \\ &\quad \left. + (1 - \beta)(1 - t_B) \frac{\partial Pr(B \text{ is hired})}{\partial \mu} - (1 - \beta) Pr(B \text{ is hired}) \frac{\partial t_B}{\partial \mu} \right]. \end{aligned}$$

Hence, if $\mu \rightarrow 0$, then $t_B \rightarrow 1$, $t_A \rightarrow 0$ and $\frac{\partial \mu_{t+1}}{\partial \mu_t} \rightarrow 0$. The above analysis can be repeated around $\mu = 1$ to show that $\frac{\partial \mu_{t+1}}{\partial \mu_t} \rightarrow 0$ when $\mu \rightarrow 1$, so that $\mu = \{0, 1\}$ are stable steady states.

Since $\mu = \{0, 1\}$ are stable and μ_{t+1} is continuous in μ_t , it must be that there exist some μ^s ($0 < \mu^s < 1$) such that if the initial distribution of managerial skills, $\mu_0 > \mu^s$, then $\lim\{\mu_t\} = 1$ and if $\mu_0 < \mu^s$, $\lim\{\mu_t\} = 0$. Finally if $\mu_t = \mu^s$, $\mu_{t+1} = \mu_t$ so that μ^s is an unstable steady state. That there are no more than three equilibria follows because $\forall \mu_t < \mu^s$, $\frac{\partial \mu_{t+1}}{\partial \mu_t} < 0$ and $\forall \mu_t > \mu^s$, $\frac{\partial \mu_{t+1}}{\partial \mu_t} > 0$.

Proof of Lemma 3: Let $H(\mu) = \beta(1 - t_A)Pr(A \text{ is hired}) - (1 - \beta)(1 - t_B)Pr(B \text{ is hired})$; then, by Lemma 2, $H(\mu)$ is continuous in μ . Differentiating equation 7 w.r.t μ , we get

$$m_A \left[Pr(A \text{ hires}) + \mu \frac{\partial Pr(A \text{ hires})}{\partial \mu} \right] + m_B \left[(1 - \mu) \frac{\partial Pr(B \text{ hires})}{\partial \mu} - Pr(B \text{ hires}) \right] = 0.$$

Hence,

$$\begin{aligned}
\frac{\partial H(u)}{\partial \mu} &= -\beta Pr(A \text{ is hired}) \frac{\partial t_A}{\partial \mu} + (1 - \beta) Pr(B \text{ is hired}) \frac{\partial t_B}{\partial \mu} \\
&+ \left[Pr(A \text{ hires}) + \mu \frac{\partial Pr(A \text{ hires})}{\partial \mu} \right] \left[\beta H t_A^\psi (1 - t_A) - (1 - \beta) L (1 - t_B)^\psi (1 - t_B) \right] \\
&+ \left[(1 - \mu) \frac{\partial Pr(B \text{ hires})}{\partial \mu} - Pr(B \text{ hires}) \right] \left[\beta L (1 - t_A)^\psi (1 - t_A) - (1 - \beta) H t_B^\psi (1 - t_B) \right] \\
&< m_A \left[Pr(A \text{ hires}) + \mu \frac{\partial Pr(A \text{ hires})}{\partial \mu} \right] + m_B \left[(1 - \mu) \frac{\partial Pr(B \text{ hires})}{\partial \mu} - Pr(B \text{ hires}) \right] \\
&- \beta Pr(A \text{ is hired}) \frac{\partial t_A}{\partial \mu} + (1 - \beta) Pr(B \text{ is hired}) \frac{\partial t_B}{\partial \mu} = -\beta Pr(A \text{ is hired}) \frac{\partial t_A}{\partial \mu} \\
&+ (1 - \beta) Pr(B \text{ is hired}) \frac{\partial t_B}{\partial \mu} < 0.
\end{aligned}$$

Hence, $H(\mu)$ is decreasing in μ .

Also if $\mu = 0$, $H(0) = \beta Pr_{min}^A > 0$ and if $\mu = 1$, $H(1) = -(1 - \beta) Pr_{min}^B < 0$. Hence, there exist some $\mu = \mu^T$ such that $H(\mu^T) = 0$. It follows that

$$\begin{aligned}
\beta(1 - t_A) Pr(A \text{ is hired}) &> (1 - \beta)(1 - t_B) Pr(B \text{ is hired}) \quad \text{if } \mu < \mu^T \\
\text{and } \beta(1 - t_A) Pr(A \text{ is hired}) &< (1 - \beta)(1 - t_B) Pr(B \text{ is hired}) \quad \text{if } \mu > \mu^T
\end{aligned}$$

Proof of Corollary 1: Differentiating (8) w.r.t γ , gives

$$\frac{\partial \mu_{t+1}}{\partial \gamma} = -\beta(1 - t_A) Pr_t(A \text{ is hired}) + (1 - \beta)(1 - t_B) Pr_t(B \text{ is hired})$$

and the results follow from Lemma 3.

Note also that differentiating equation 11 w.r.t μ gives

$$\begin{aligned}
\frac{\partial \mu_{t+1}}{\partial \mu_t} &= \frac{1}{\nu} \left[\beta \frac{\partial Pr(A \text{ is hired})}{\partial \mu} - \beta \gamma (1 - t^A) \frac{\partial Pr(A \text{ is hired})}{\partial \mu} + \beta \gamma Pr(A \text{ is hired}) \frac{\partial t_A}{\partial \mu} \right. \\
&+ \left. (1 - \beta) \gamma (1 - t^B) \frac{\partial Pr(B \text{ is hired})}{\partial \mu} - (1 - \beta) \gamma Pr(B \text{ is hired}) \frac{\partial t_B}{\partial \mu} \right] > 0
\end{aligned}$$

since first term in the square bracket is positive (Lemma 1(d)) and the remaining terms is $-\frac{\partial H(\mu)}{\partial \mu}$ in the proof of Lemma 3, which is positive.

Proof of Proposition 4: By Lemma 3,

$$\beta(1 - t_A) Pr(A \text{ is hired}) \underset{(>)}{>} (1 - \beta)(1 - t_B) Pr(B \text{ is hired})$$

if $\mu \underset{(>)}{<} \mu^T$ and $\frac{\partial \mu_{t+1}}{\partial \gamma} = 0$ if $\mu_t = \mu^T$. Let μ^* be the unique steady state when $\gamma = 0$.

Case (1): Suppose $\mu^T < \mu^*$ when $\gamma = 0$. Then $\mu_{t+1}(\mu^T) > \mu^T$. By corollary 1, μ^* is increasing in γ . Consider a $\mu_t = \hat{\mu} < \mu^T$ such that $\mu_{t+1} < \mu_t$ when $\gamma = 1$ (such a $\hat{\mu}$ exists by Proposition 3). By Proposition 2, $\mu_{t+1} > \mu_t$ when $\gamma = 0$. Since μ_{t+1} is continuous in γ , there exists a $\hat{\gamma}$ such that $\mu_{t+1} = \hat{\mu}$. Analytically, $\hat{\mu}$ and $\hat{\gamma}$ are given by the solution to:

$$\hat{\mu} = \frac{1}{\nu} [\beta(1 - \hat{\gamma}) + \hat{\gamma} \hat{t}_A] \hat{P}r(A \text{ is hired}) + (1 - \beta) \hat{\gamma} (1 - \hat{t}_B) \hat{P}r(B \text{ is hired})$$

and $\frac{\partial \mu_{t+1}(\mu_t, \hat{\gamma})}{\partial \mu_t} = 1$ if $\mu_t = \hat{\mu}$.

Hence if $\gamma < \hat{\gamma}$, the economy has a unique and stable steady state at $\mu^* = \mu(\gamma)$. For $\gamma > \hat{\gamma}$, we have

$$\begin{aligned} \mu_{t+1} - \mu_t &> 0 \quad \text{if } \mu_t = 0, \\ \mu_{t+1} - \mu_t &< 0 \quad \text{if } \mu_t = \hat{\mu}, \\ \text{and } \mu_{t+1} - \mu_t &> 0 \quad \text{if } \mu_t = \mu^T. \end{aligned}$$

Hence, by continuity of μ_{t+1} in μ_t , there exists some $\mu^B(\gamma)$ such that $\mu_{t+1} = \mu_t = \mu^B(\gamma)$; $\mu^B(\gamma)$ will then be a (bad) steady state; and some $\mu = \mu^s(\gamma) \in (\mu^B(\gamma), \mu^T)$ such that $\mu_{t+1} = \mu_t = \mu^s(\gamma)$, so that $\mu^s(\gamma)$ is also a steady state. Since, $\mu^* > \mu^T$, by corollary 1, μ^* will increase to some $\mu = \mu^G(\gamma)$; the good steady state.

Since μ_{t+1} is monotonic, by arguments similar to the proof of Proposition 3, $\mu^B(\gamma)$ and $\mu^G(\gamma)$ are stable steady states and μ^s is an unstable steady state. By corollary 1, if $\gamma' > \gamma > \hat{\gamma}$, then $\mu^B(\gamma') < \mu^B(\gamma)$ and $\mu^G(\gamma') > \mu^G(\gamma)$.

Case (2): Suppose $\mu^T = \mu^*$ when $\gamma = 0$ and consider some $\mu_t < \mu^*$ (we will show that μ_t is a steady state for some γ). Then

$$\begin{aligned} \mu_{t+1} - \mu_t &< 0 \quad \text{if } \gamma = 0, \quad (\text{by Prop. 2}) \\ \text{and } \mu_{t+1} - \mu_t &> 0 \quad \text{if } \gamma = 1 \quad (\text{by Prop. 3}). \end{aligned}$$

Since μ_{t+1} is continuous in γ , there exist some $\hat{\gamma}$ such that $\mu_{t+1} = \mu_t = \mu^B$. Similarly, there exist another steady state $\mu^G(\gamma) > \mu^*$.

The case of $\mu^T > \mu^*$ when $\gamma = 0$ is identical to case (1). In both cases (1) and (2), the stable steady states are on either side of μ^T and are further away from μ^T as γ increases.

Proof of proposition 5: The measure of type-A managers in a steady state, μ_∞ solves

$$\mu_\infty = \beta[1 - \gamma + \gamma t_A(\mu_\infty)] Pr(A \text{ is hired} | \mu_\infty) + (1 - \beta)\gamma(1 - t_B(\mu_\infty)) Pr(B \text{ is hired} | \mu_\infty). \quad (22)$$

Differentiating equation (22) with respect to γ , one can show that the measure of type- A managers varies with γ unless

$$\beta(1 - t_A(\mu_\infty))Pr(\text{A is hired}|\mu_\infty) = (1 - \beta)(1 - t_B(\mu_\infty))Pr(\text{B is hired}|\mu_\infty). \quad (23)$$

That is, a marginal increase in γ will lead to an increase (decrease) in the measure of managers that can evaluate skill A in the (local) equilibrium steady state if and only if

$$\beta(1 - t_A(\mu_\infty))Pr(\text{A is hired}|\mu_\infty) > (<)(1 - \beta)(1 - t_B(\mu_\infty))Pr(\text{B is hired}|\mu_\infty).$$

This follows because if at μ , an increase in γ leads to an increase in the measure of managers with skill- A expertise, then succeeding generations of workers will further increase their investment in skill A , reinforcing the consequences for the limiting measure of managerial expertise. In particular, in the long run steady state, if $a_A > a_B$ then, unless $\beta < .5$, an increase in the importance of nurture for managerial expertise, γ , will raise the equilibrium steady state mix of managers with expertise in skill A . Finally, note that, *ceteris paribus* individuals will devote more time to acquiring skill A if β is increased. That is, if (17) holds at equality for one value of β , then fixing μ_∞ , for $\beta' > \beta$,

$$\beta'(1 - t_A(\mu_\infty; \beta'))Pr(\text{A is hired}|\mu_\infty; \beta') > (1 - \beta')(1 - t_B(\mu_\infty; \beta'))Pr(\text{B is hired}|\mu_\infty; \beta'),$$

where we explicitly incorporate β' into the notation for clarity.

References

- [1] Arthur, Brian W. (1994), *Increasing Returns and Path Dependence in the Economy*, Ann Arbor, Michigan.
- [2] Arthur, Brian W. (1989), "Competing Technologies, Increasing Returns and Lock-in by Historical Events," *Economic Journal*, 99, 116-31.
- [3] Athey, Susan, Christopher Avery and Peter Zemsky (2000), "Mentoring and Diversity," *American Economic Review*, 90(4), 765-786.
- [4] Banerjee, Abhijit (1992), "A Simple Model of Herd Behavior", *Quarterly Journal of Economics*, 107(3), 797-817.
- [5] Bikhchandani, Sushil, David Hirshleifer and Ivo Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *Journal of Political Economy*, 100(5), 992-1026.
- [6] Coate, Stephen, and Glenn C. Loury (1993) "Will Affirmative Action Policies Eliminate Negative Stereotypes?", *American Economic Review*, 83, 1220-40.
- [7] Cornell, Bradford and Ivo Welch (1996), "Culture, Information, and Screening Discrimination", *Journal of Political Economy*, 104(3), 542-71.
- [8] Granick, David (1972), *Managerial Comparisons of Four Developed Countries: France, Britain, United States and Russia*, Cambridge, MIT Press.
- [9] Moro, Andrea and Peter Norman (1999), "Affirmative Action in a Competitive Economy" SSRI working paper 9914, Wisconsin University.
- [10] Norman, Peter (1999), "Statistical Discrimination and Efficiency", SSRI Working Paper 9916, Wisconsin University.
- [11] Phelps, Edmund (1972), "The Statistical Theory of Racism and Sexism", *American Economic Review*, 659-662.
- [12] Prendergast, Canice (1993), "A Theory of 'Yes Men'", *American Economic Review*, 83(4), 757-70.
- [13] Scharfstein, David and Jeremy C. Stein (1990), "Herd Behavior and Investment", *American Economic Review*, June 1990, 465-79.
- [14] Schelling, Thomas (1978), *Micromotives and Macrobehavior*, New York, W.W. Norton
- [15] Unseem, Michael (1989), *Liberal Education and the Corporation: The Hiring and Advancement of College Graduates*, Aldine de Gruyter, New York.
- [16] Zweibel, Jeffrey (1995), "Corporate Conservatism and Relative Compensation", *Journal of Political Economy*, 103(1), 1-25.

WORKING PAPER SERIES 2002

ase.tufts.edu/econ/papers

- 2002-01** FULLERTON, Don and Gilbert METCALF. The Distribution of Tax Burdens.
- 2002-02** EGLLESTON, Karen and Richard ZECKHAUSER. Government Contracting for Health Care.
- 2002-03** EGLLESTON, Karen, Robert JENSEN, and Richard ZECKHAUSER. Information and Communication Technologies, Markets and Economic Development.
- 2002-04** EGGLESTON, Karen, George NORMAN, and Lynne PEPALL. Managed Health Care and Provider Integration: a Theory of Bilateral Market Power.
- 2002-05** CHISHOLM, Darlene C. and George NORMAN. Heterogeneous Preferences and Location Choice with Multi-Product Firms.
- 2002-06** BROWN, Drusilla K., Alan V. DEARDORFF, and Robert M. STERN. Pros and Cons of Linking Trade and Labor Standards.
- 2002-07** BROWN, Drusilla K., Alan V. DEARDORFF, and Robert M. STERN. Impacts on NAFTA Members of Multilateral and Regional Trading Arrangements and Initiatives and Harmonization of NAFTA's External Tariffs.
- 2002-08** IOANNIDES, Yannis M. and Jeffrey E. ZABEL. Interactions, Neighborhood Selection and Housing Demand.
- 2002-09** METCALF, Gilbert E. and George NORMAN. Oligopoly Deregulation and the Taxation of Commodities.

- 2002-10** METCALF, Gilbert E. and George NORMAN. Oligopoly
Deregulation in General Equilibrium: A Tax Neutralization Result.
- 2002-11** BABIKER, Mustafa H., Gilbert E. METCALF, and John REILLY.
Tax Distortions and Global Climate Policy.
- 2002-12** BERNHARDT, Dan, Eric HUGHSON, and Edward KUTSOATI.
Survival of the Unfittest: How Dodos Become Managers.
- 2002-13** BERNHARDT, Dan, Murillo CAMPBELLO, and Edward
KUTSOATI. Who Herds?