Learning and Self-Reinforcing Behavior

Margaret A. Meyer*

and

Jeffrey Zwiebel†

March 2007

Abstract

This paper examines history-dependence in decision-making by individuals or organizations facing a sequence of related decisions. A decision-maker must, in each period, choose whether to accept or reject a new opportunity, of uncertain attractiveness, on the basis of a noisy observation. We analyze how acceptance decisions are affected by past decisions, in a setting in which previous acceptances yield increasingly precise information on which to base future decisions but each rejection causes the stock of accumulated information to revert to a baseline level. We identify two potentially opposing informational effects of acceptances, one of which, the "stock effect", makes future acceptances more likely and the other of which, the "precision effect", can make them either more or less likely. We explore how the relative signs and magnitudes of these two effects vary with changes in the decision-making environment, specifically, the ex ante expected cost of acceptances, the discount factor, the baseline level of precision, and the amount of additional precision obtained with each successive acceptance. We thus provide conditions under which acceptance decisions are self-reinforcing (i.e., past acceptances make future acceptances more likely) or instead self-limiting. Our framework and techniques can be used to analyze a wide range of related dynamic decision-making scenarios.

*Nuffield College and Department of Economics, Oxford University

†Graduate School of Business, Stanford University. We are grateful to Heski Bar-Isaac, Mark Garmaise, Alexander Gorbenko, Chris Harris, Markus Mobius, Peter Norman Sorensen, and seminar audiences at Stanford, U. C. San Diego, Tokyo, Northwestern, Munich Symposium on Economics and Psychology, and European Summer Symposium on Economic Theory. Any errors are our own.

1 Introduction

Organizations and individuals often face sequences of similar decisions, for example about the promotion of junior employees, the adoption of proposed projects, the purchase of goods or services, or the carrying out of a particular type of activity. The standards used by different organizations or individuals often appear to differ in their stringency. For example, promotion policies of otherwise similar organizations often appear to vary significantly. And apparently similar individuals often appear to make systematically different purchasing or consumption decisions. What accounts for such apparent variation in decision criteria? One obvious possibility is that the organizations or individuals in question are operating in different environments or have different objectives. An alternative set of explanations starts from the premise that the decision-makers are using the same decision policy, but that this policy depends, for a variety of potential reasons, on the history of past decisions. For example, past promotions might predispose an organization to be more lenient in current promotion decisions. And past purchases might predispose an individual to set a lower reservation price for current purchasing decisions. When such "self-reinforcing" decision policies are in use, even very similar organizations and individuals could appear to differ systematically over a period of time in the decision criteria they employ.

Why might sequential decision-making by organizations or individuals display such history-dependence? And specifically, why might the form of history-dependence be that decisions appear to be "self-reinforcing"? Psychologists, organizational theorists, and more recently economists have developed a number of types of explanations for why individuals and groups might display "excessive adherence to past states of mind or behaviors" (Hirshleifer and Welch (2002, p. 403)), tendencies also referred to as inertia or status quo bias. Some of these explanations center on group dynamics.[1] Others focus on cognitive distortions—for example, both Festinger's (1957) theory of cognitive dissonance reduction and Rabin and Schrag's (1999) model of confirmatory bias emphasize that initial beliefs and/or choices may distort the subsequent gathering and processing of information and by this route cause subsequent choices to resemble previous choices to an inefficient degree.

[1]See, for example, Kuran's (1987) model of preference falsification and its consequences.

In contrast to models of inertia based on group interactions or on cognitive distortions, we explore models in which self-reinforcing behavior arises even with a single decision-maker who is Bayesian and forward-looking. Our decision-maker, however, can only imperfectly evaluate decision opportunities, and his past decisions directly affect the quality of future ones. The decision-maker faces a sequence of binary decisions, whether to accept or reject available opportunities. We analyze whether acceptances early in the sequence can make later acceptances more likely, even when the intrinsic characteristics of the opportunities in the sequence (for example, employees, projects, or merchandise) are, in a probabilistic sense, unrelated. We show that under a wide range of conditions, this type of "self-reinforcing behavior" in decision-making can arise from the accumulation of a stock of information that is valuable for future decisions.

In the models we study, a new opportunity emerges each period, and its profitability (or more generally, attractiveness) is uncertain. The profitabilities of opportunities emerging in different periods are independently distributed. Nevertheless, the decisions in the sequence are linked by the fact that the choice made today affects the quality of tomorrow's decision. To fix ideas, consider, for example, an organization's decision, each period, whether to fire a newly-hired employee or whether to retain her for one more period, on the basis of a noisy signal about her uncertain productivity. (We assume that after being retained for one more period, the employee must leave, so the organization makes only one decision about each employee.) The organization tries to retain employees of high ability and fire those of low ability. If the employee is retained, the observation of her subsequent performance improves the decision about next period's new hire, because the two individuals work in similar environments, and the knowledge already acquired about the retained employee helps to interpret the noisy signal about the new hire. Furthermore, if the new hire is then also retained, her subsequent performance allows an even more precise inference about the productivity of the following period's new hire, for the same reason, and so on. Thus retention of an employee, by making available comparative performance information, improves the next decision, and the greater the number of successive retentions, the more valuable is the comparative performance observation, so the greater is the quality of the following decision. On the other hand, if a newly-hired employee is ever fired, then the quality of the decision about the next period's new hire is reduced to a baseline level, since it must be made without the benefit of any comparative performance information. We ask: Under the optimal decision policy, how does the probability of acceptance today (i.e. retention of the current new hire) depend on past decisions?

Past decisions affect the organization's stock of information, which is useful as an input into the current decision. Specifically, the larger the number of acceptances since the last rejection, the greater is the precision with which the current new opportunity can be evaluated, and therefore the higher is the expected present discounted value of the decision-maker's payoff stream. This accumulation of information through repeated acceptances, however, has two distinct, potentially conflicting effects on the optimal probability of another acceptance. We term these effects the "stock effect" and the "precision effect". The former unambiguously increases the probability of continued acceptance, while the latter may raise or lower it.

At first glance, it might appear straightforward that the longer the firm has gone without firing, the smaller the probability that it will presently terminate its employee. Since the larger the stock of precision the more valuable it is, a larger stock makes the decision-maker more reluctant to reject and thereby dissipate the stock. More precisely, the critical value of the posterior evaluation of the new opportunity below which it is optimal to reject falls with the stock of information. We refer to this as the "stock effect" of accumulated information.

However, at the same time, the accumulated stock of information also affects the way in which the organization forms its posterior evaluation from the noisy observation it receives. To understand how this effect, which we term the "precision effect", operates, note that the larger the stock of information, the more precise is the noisy observation about the employee, so the greater is the weight the organization places on this observation in updating its beliefs about the employee. Consequently, the distribution of the posterior evaluation of the employee is more variable. How this greater variability per se affects the optimal probability of retaining the employee depends upon whether the critical threshold for the posterior evaluation is greater or less than the prior mean of the employee's ability. If the critical threshold is greater than the prior mean, then the more variable the posterior evaluation, the more likely it is to exceed the threshold and hence the more likely the employee is to be retained. In this case, the precision effect reinforces the stock effect, and we say that acceptance decisions are *strongly self-reinforcing*. On the other hand, if the critical threshold is below the prior mean, then the more variable the posterior evaluation, the more likely it is to fall below the threshold and thus, ceteris paribus, the less likely the employee is to be retained. In this case, the precision effect opposes the stock effect, and we need to determine which effect dominates. If the stock effect is stronger, then the optimal probability of accepting the new opportunity is increasing in the number of acceptances since the last rejection, and we say that acceptances are *self-reinforcing*, while if the precision effect dominates, we say that acceptances are *self-limiting*.

In Section 3, we study these potentially conflicting effects of information accumulation in a simple version of the model sketched above. An organization must decide each period whether to promote a junior employee to a senior position, or whether to fill the senior position with an untested outside hire. All that is observed each period is the joint output of the current junior and senior, and the organization tries to maximize the discounted present value of this joint output. Filling the senior slot with a promoted junior, who has already been evaluated once, allows a better evaluation of next period's junior than filling it with an untested outsider. We study how the history dependence of decisions under the optimal policy depends upon two parameters, the discount factor and the ex ante " acceptance hurdle", formally defined as the expected cost, to flow payoff, of choosing to retain the current junior (i.e. choosing the option that enhances the stock of information). We show that for all discount factors, retention behavior is strongly self-reinforcing for sufficiently positive values of the acceptance hurdle, self-reinforcing for all non-negative values of this hurdle, and self-limiting for at least some levels of accumulated information for sufficiently negative values of this hurdle. We provide a heuristic interpretation of the dependence of the relative sizes of the stock and precision effects on the acceptance hurdle.

Section 4 studies a more general model of history-dependence in decision-making and presents several different scenarios described by this general model. The first of these scenarios can be viewed as a "bandit" model with market wages, the second as a model of learning about environmental uncertainty, and the third formalizes the overlapping generations story above. In each of the scenarios, the decision-maker chooses each period between acceptance and rejection of an uncertain payoff opportunity, with successive acceptances gradually increasing the stock of precision and each rejection fully dissipating it. The more general model contains two additional parameters, one representing the increment in precision resulting from each successive acceptance and the other the baseline level to which the precision reverts after a rejection. Once again, we identify conditions under which acceptances are strongly self-reinforcing, self-reinforcing, or self-limiting.

It is helpful in clarifying the structure and objectives of our models to compare and contrast them with work on i) multi-armed bandit problems (optimal experimentation); ii) social learning; iii) the dynamics of promotion decisions; and iv) inertia or status quo bias. The next section discusses these related branches of literature.

2 Related Literature

2.1 Bandit Models

In a typical bandit model, a decision-maker chooses in each period which of several "arms" to pull; the payoff from each arm is random, and its realized value is informative about the underlying distribution from which it is drawn.[2] The strategy which maximizes the expected discounted present value of payoffs involves, at least initially, experimentation: the decision-maker may choose an arm with a lower expected payoff, because of the future value of the information in the payoff realization. Our model shares with bandit models this trading-off of the short-term expected payoff against the long-term value of accumulating information. However, whereas in bandit models each arm is available to the decision-maker in each period, in our analysis the decision-maker effectively cannot reap payoffs from each new (independently distributed) "payoff op-

[2]Berry and Fristedt (1985) and Bergemann and Valimaki (2006) provide surveys of bandit models. Economic applications are developed by Rothschild (1974), Easley and Kiefer (1988), Aghion, Bolton, Harris and Jullien (1991), Banks and Sundaram (1992, 1998), Bolton and Harris (1993), Rustichini and Wolinsky (1985), Schlag (1998), and Keller, Rady, and Cripps (2005), among others.

portunity" for more than one period.[3] In bandit models, past decisions affect current choices because they directly convey information about the profitability of the current options; in our model, such a direct link is absent, and past decisions affect *only* the precision with which the current opportunity can be evaluated. The focus in bandit models is typically on whether or not behavior under the optimal strategy converges to the full-information outcome, whether the optimal strategy can be characterized in terms of Gittins indices (Gittins and Jones (1974)), and whether and how much experimentation takes place. Because the focus is usually normative ("What should a decision-maker do?"), relatively less attention is typically paid to predictions about the nature of the history-dependence in the sequence of decisions under the optimal strategy. In our analysis, such predictions are a prime focus. Such predictions would be essential for any empirical tests designed to distinguish, say, models of rational learning from models of boundedly rational behavior.

One class of bandit models where predictions regarding the *form* of history-dependence in decision-making have received close attention is the job-matching paradigm initiated by Jovanovic (1979).[4] In Jovanovic's model, workers and firms learn over time about the quality of the worker-firm match, with the choice at each point being whether to continue the existing match or whether to terminate it, in which case new matches are drawn from a known prior distribution of match quality. Jovanovic analyzes how the probability that a match is terminated varies with its duration, a task that is considerably simplified by the use of normal distributions (as in our model). One could identify in Jovanovic's model analogs of our stock and precision effects—the stock effect would represent how a match's duration (a measure of how much is known about its quality) affects next period's optimal reservation value for estimated match quality, and the precision effect how duration affects the variance of next period's estimate, conditional on the current estimate. Nevertheless, in Jovanovic's model, as in any bandit model, there is an important additional effect, the "selection effect", which is not present in

[3]Even in the "bandit" version of our model, presented in Section 4.1, the firm's payoff opportunity each period is the difference between the incumbent factor's firm-specific value and its market wage, and the unconditional distribution of this difference is iid across time. For other variants of bandit models in which the payoff from each arm each period depends not only on the arm's intrinsic value but also on market competition, see Bergemann and Valimaki (1996) and Felli and Harris (1996).

[4]For a very recent application of this framework, see Marinescu (2006).

our model. The selection effect arises because a match's duration is informative not just about how much is known about its quality but also about how good the match is likely to be. In our model, by constrast, each opportunity is evaluated only once, so the history of past decisions is not directly informative about the quality of the current new opportunity. Whereas in our model, history-dependence in decision-making reflects only the interplay of the stock and precision effects, in bandit models such history-dependence will always reflect the additional presence of the selection effect.

2.2 Social Learning Models

Models of social learning examine the phenomenon of herding, which appears to be a form of self-reinforcing behavior at the level of group rather than individual decision-making. In models such as Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1992), and Smith and Sorensen (2000), short-lived individuals choose, one after the other, among several uncertain options, basing their choices both on their private signals and on the decisions of those who acted before them. Attention has been focused on the conditions under which individuals eventually "herd"; that is, mimic the decisions of their predecessors, regardless of their private information.[5] While it might seem that herding is fundamentally due to individuals' myopia, Smith and Sorensen (2006) have recently shown that such outcomes would arise (though with lower probability) even if individuals were to act as part of a forward-looking team, solving a discounted dynamic optimal experimentation problem.[6] Smith and Sorensen thus recast social learning models as (generalized) bandit models, recasting outcomes with incorrect herds as failures of complete learning in experimentation models. The fundamental difference between the self-reinforcing behavior in social learning models and in our model is therefore not that social learning models involve multiple decision-makers while ours has only one. Rather, it is exactly the difference emphasized in Section 2.1: In social learning models, as in bandit models, past decisions affect current ones because they

[5]Similarly, the models of reputational herding of Sharfstein and Stein (1990) and Zwiebel (1995) examine when the interaction between the actions of others and considerations of reputation-building lead individuals to ignore their own information and copy the actions of others.

[6]See also Ali and Kartik (2006), who show that herding arises in a sequential voting model even when voters are sophisticated and forward-looking.

reveal information about the payoffs from the current options; whereas in our model, payoff opportunities are independent across time, and past decisions influence only the precision with which the current opportunity can be evaluated.

2.3 Models of the Dynamics of Promotion Standards

Analyses of career design as in Meyer (1991, 1992), Prendergast (1992), and Bernhardt (1995) present various reasons why it may be efficient for organizations to structure promotion ladders so that success is self-reinforcing—that is, succeeding in the early stage of her career gives an employee an increased chance of succeeding in the next stage. An important difference between these papers and ours is that these papers focus on self-reinforcing outcomes *within* an individual career; our model, when interpreted in terms of promotions, examines self-reinforcing outcomes *across* the careers of different individuals.

Sobel (2000, 2001) studies two models of the evolution over time of standards for promotion or for entry into a club. Like our simple model in Section 3, Sobel's models involve overlapping generations of individuals, with (some of) today's candidates becoming tomorrow's members of the elite, and with the characteristics of the current elite influencing the standards employed in deciding on today's candidates. In Sobel's models, in contrast to ours, the candidates behave strategically, and the judges (who decide on promotions or entry) are not forward-looking.

2.4 Models of Inertia or Status Quo Bias

Cornell and Welch (1996) develop an information-based theory to explain why employers may be systematically more likely to hire candidates who belong to the same group (racial, social, etc.) that they do, even when it is common knowledge that distributions of quality are the same across groups and employers have no intrinsic preference for candidates who belong to their own group. The key elements of their model are that job applicants' qualities can be evaluated only with noise, that this noise is less when an applicant belongs to the same group as the employer, and that the standard for hiring is "tough", that is, above the prior expected quality. With "tough" standards, the more precisely an applicant can be evaluated, the more likely it is that the posterior evaluation of his quality will exceed the critical threshold and therefore the more likely that he will be hired. What we have termed the "precision effect" of greater information is thus the driving force behind the "status quo bias" that emerges in the Cornell and Welch model.

Hirshleifer and Welch (1994, 2002) develop an information-based model to study organizational inertia vs. organizational impulsiveness. Specifically, they analyze how the loss of "institutional memory" resulting from the replacement of an experienced manager by a new one affects the likelihood that the old manager's policies will be continued or reversed. Our model is similar in spirit, since our stock of precision is a measure of institutional memory, and since our focus on self-reinforcing behavior is a focus on how decision-making varies with the level of this stock. However, in one important respect, our analysis goes beyond Hirshleifer and Welch's. Whereas they treat the loss of institutional memory as exogenous, focusing purely on its consequences, in our model the organization's stock of information evolves endogenously. Our analysis has at its core an idea they mention only in passing (1994, p.29): "Supervisors may trade off the benefit of obtaining a more competent manager against the cost of losing institutional memory."

3 A Simple Model

An infinitely-lived firm employs two individuals in each (discrete) period, one in the "junior" position and one in the "senior" position. Each period's output is given by the sum of the junior's and the senior's abilities. Only this team output is observed; individual ability is unobservable. After producing in the senior position for one period, an employee must leave the firm, and the firm must decide whether to fill the senior position by promoting the current junior or by hiring a new "senior" employee from outside. If the firm chooses the latter option, the current junior is dismissed. In either case, the junior position is then filled with a new "junior" employee from outside. The firm's problem is to choose each period whether to promote or fire the current junior, in order to maximize the expected discounted present value of team output, given discount factor δ.

The abilities of newly hired juniors are independently and normally distributed, with mean normalized to 0 and precision (the inverse of the variance) normalized to 1. An individual's ability, and hence contribution to team ouptut, remains constant over time. Newly hired seniors are drawn from a potentially different pool of individuals, with abilities also independently and normally distributed, also with precision 1, but with mean $m \in (-\infty, \infty)$. The parameter m represents the difference in expected contribution to team output between a "senior" hired from outside and a "senior" who, regardless of team performance while a junior, is promoted from within. This parameter can also be given a more general interpretation: it reflects the net ex ante expected cost to flow payoff (direct cost and/or opportunity cost) of choosing to fill the senior slot by promoting the current junior rather than hiring from outside. If, for example, there were transactions costs of firing and hiring, these would per se make m negative. On the other hand, higher average ability in the pool of people available for the senior slot than in the pool available for the junior slot would per se make m positive.[7] We will henceforth refer to m as the (ex ante) "acceptance hurdle".

The observation of team output in period t allows the firm to update its beliefs about the ability of the current junior, given what it already knows about the current senior's ability. The decision whether to promote or fire the current junior is made on the basis of these updated beliefs.

Suppose that in any period, the abilities of the junior and senior are both believed to be normally distributed. Then it follows from the standard "normal learning model" (see DeGroot (1970, Ch. 9)) that after the observation of team output, the posterior distribution of the junior's ability will also be normal. Since next period's senior will be either the promoted junior or a new hire, our initial supposition of normality is justified. Furthermore, the normal learning model also implies that if the mean and precision of the distribution of the current senior's ability at the start of period t are λ_t

[7] Note that this simple model ignores wages, so "abilities" should be interpreted as net contribution to firm profit. For a scenario which does incorporate market wages, see Section 4. The more general model of Section 4 also incorporates the case where the prior precision on newly hired seniors' abilities is different from 1, the prior precision on newly hired juniors' abilities. Finally, note that this is a partial equilibrium analysis, ignoring the influence of a firm's own past decisions on the composition of the pool of potential new hires.

and k_t, respectively, then after the observation of team output x_t, the posterior mean of the junior's ability will be

$$\tilde{\mu}(k_t, x_t - \lambda_t) = \frac{k_t(x_t - \lambda_t)}{k_t + 1}, \tag{1}$$

and the posterior precision will be

$$k_{t+1} = k_t + 1. \tag{2}$$

Intuitively, the posterior mean ability of the junior after team output is observed is a weighted average of the prior mean (given here by 0) and $x_t - \lambda_t$, the observation of team output with the senior's expected contribution netted out. The weight given to the observation relative to the prior mean depends on the "signal to noise ratio" in the observation. Since the prior precision of the junior's ability (the signal) is 1, while the precision of the current senior's ability (the noise in the observation) is k_t, the weight on the observation is k_t times the weight on the prior. That is, the observation receives weight $\frac{k_t}{k_t+1}$, and the prior receives weight $\frac{1}{k_t+1}$, thereby yielding equation (1). Furthermore, under the normal learning model, the precision of the posterior estimate of the junior's ability is simply the sum of the precision of the prior, 1, and the precision of the noise in the observation, k_t, thus yielding equation (2).

Denote $x_t - \lambda_t$ by $\bar{x}_t$, and note that conditioning on the information available at the beginning of period t, this random variable $\bar{x}_t$ is the sum of a normal random variable with mean 0 and precision 1 (the junior's contribution to team output) and an independent normal random variable with mean 0 and precision k_t (the senior's contribution, net of expected ability λ_t). Consequently, $\bar{x}_t$ is normally distributed with mean 0 and precision $\frac{k_t}{k_t+1}$. It then follows from equation (1) that, conditioning on the information available at the beginning of period t, the random variable $\tilde{\mu}(k_t, \bar{x}_t)$ is distributed according to

$$\tilde{\mu}(k_t, \bar{x}_t) \sim N\left(0, \frac{k_t}{k_t + 1}\right), \tag{3}$$

where $N(\mu, v)$ denotes a normal distribution with mean μ and variance v.[8]

[8] We will use $\tilde{\mu}(k_t, \bar{x}_t)$ and $\mu(k_t, \bar{x}_t)$ to denote this random variable and its realization, respectively.

We can now derive the Bellman equation for the firm's dynamic decision problem. The firm chooses each period, after observing x_t, whether to retain or fire the junior. If the junior is retained and moves to the senior slot, the firm's expected payoff next period is $0+\mu(k_t, \bar{x}_t)$, and the precision on the ability of the new senior is $k_{t+1} = k_t + 1$. If the junior is fired and the senior slot is filled with an outside hire, next period's expected payoff is $0+m$, and the precision on the new senior reverts to 1.

Note that from (3), the conditional distribution of $\tilde{\mu}(k_t, \bar{x}_t)$ is independent of the expected ability of the current senior λ_t. It follows that, under the optimal decision rule, the probabilities of retaining and firing the junior, prior to observing output x_t, must be independent of λ_t. Consequently, in this model, the relevant state (on which the probability of termination depends) can be summarized solely by the precision, k_t, of the ability distribution of the current senior. From (2), the precision k_t is always a positive integer and represents the number of periods since the most recent firing of a junior. We can view k_t as a stock of information possessed by the organization which is valuable for future decision-making.

The firm maximizes the expected discounted present value of team output. Since the junior's expected contribution to output (under any policy) is 0 in each period, we can simply define the firm's value function $V : \mathbf{Z}^+ \mapsto \mathbf{R}$ as the expected discounted present value of the senior's contribution under the optimal policy, given the state k. By convention, we compute $V(k)$ *before* the current period's team output is observed but *exclude* from it the contribution of the current senior (whose ability distribution has precision k). With this convention, the Bellman equation is given by

$$V(k) = \delta E_{\bar{x}} \left[\max \left\{ \tilde{\mu}(k, \bar{x}) + V(k+1), m + V(1) \right\} \right], \tag{4}$$

where $\tilde{\mu}(k, \bar{x})$ is normally distributed with mean 0 and variance $(\sigma_k)^2 \equiv k/(k+1)$. If, having observed the realization $\bar{x}$, the firm chooses to retain the current junior, then next period the expected contribution from the senior will be $\mu(k, \bar{x})$, and the precision on her ability will be $k+1$; if instead the current junior is fired, then next period's expected contribution from the (newly drawn) senior will be m, and the precision on her ability will be 1.

Since, in a single-person decision problem, better information is always more valuable, $V(k)$ is increasing in k. Furthermore, $V(k)$ is bounded, since

$$V(k) \leq \frac{\delta}{1-\delta} E_{\bar{x}} \left[\max \left\{ \tilde{\mu}(\infty, \bar{x}), m \right\} \right].$$

This bound on $V(k)$ equals the value the firm would reap from choosing between retaining and firing the current junior if, today and every period in the future, the precision of the ability distribution of the current senior were infinite, regardless of which choices the firm made.

A stationary decision policy for the firm is a rule which for each state k assigns a cutoff level $\bar{\mu}_k$ such that a junior is fired in state k if and only if her estimated ability $\mu(k, \bar{x})$ is less than $\bar{\mu}_k$. It is straightforward to see that if an optimal decision policy exists, it must be a stationary one which satisfies equation (4). It is apparent from (4) that for the optimal policy, $\bar{\mu}_k = m + V(1) - V(k+1)$. Since $V(k)$ is increasing in k, $\bar{\mu}_k < m$ for all $k \geq 1$. Thus, in all states the firm retains some juniors whose posterior expected ability is below that of a senior randomly chosen from outside; doing so allows it to benefit in future periods from the increase in the stock of precision.

A stationary policy $\bar{\mu}_k$ uniquely defines a function $p : \mathbf{Z}^+ \mapsto [0,1]$, giving the probability, prior to observing team output, that the current junior will be retained in state k. Specifically,

$$p(k) = Pr[\tilde{\mu}(k, \bar{x}) \geq \bar{\mu}_k] = 1 - \Phi\left[\frac{\bar{\mu}_k}{\sigma_k}\right] = 1 - \Phi\left[\frac{m + V(1) - V(k+1)}{\sigma_k}\right], \qquad (5)$$

where the second equality uses the fact that $\tilde{\mu}(k, \bar{x}) \sim N\left(0, (\sigma_k)^2\right)$. With the definition

$$A(V(k)) \equiv \frac{m + V(1) - V(k+1)}{\sigma_k}, \quad \forall k \in \mathbf{Z}^+, \qquad (6)$$

the probability of retention can be expressed as

$$p(k) = 1 - \Phi(A(V(k))), \qquad (7)$$

and $p(k)$ rises as $A(V(k))$ falls.

Our interest is in how $p(k)$ varies with k—that is, how the probability of retaining the current junior varies with the stock of information, which measures the number of periods since the last decision to fire a junior. It is easy to identify precisely what we referred to in the introduction as the "stock effect" and the "precision effect". On the one hand, since $V(k)$ is increasing in k, the optimal $\bar{\mu}_k$ is decreasing in k—for a larger k, the firm compares the junior's estimated ability with a lower standard. This is the stock effect: since the continuation value following a decision to retain, $V(k+1)$, is increasing in k, while the continuation value following a firing, $V(1)$, is independent of k, the firm is, ceteris paribus, more reluctant to fire, and thereby dissipate its stock, the larger is k. Formally, the stock effect is captured by the decrease in the numerator of $A(V(k))$ with k. The stock effect per se unambiguously acts to make $p(k)$ increase with k. (See Figure 1a.)

On the other hand, holding the cutoff $\bar{\mu}$ fixed, the accumulation of informaion, measured by k, also affects the way in which beliefs about the current junior's ability are updated. Specifically, the ex ante variance of the posterior evaluation of the junior, $(\sigma_k)^2$, is increasing in k. This is the precision effect, and it is formally captured by the appearance of σ_k in the denominator of $A(V(k))$. Whether the precision effect reinforces or opposes the stock effect can be seen from (5) to depend on whether the optimal cutoff $\bar{\mu}_k$ is above 0 ("stringent") or below 0 ("lenient"), where 0 is the prior mean of the junior's ability. If $\bar{\mu}_k \equiv m + V(1) - V(k+1) > 0$, then the increase in $(\sigma_k)^2$ with k per se raises the likelihood that the posterior evaluation of the junior will surpass the stringent cutoff: if $\bar{\mu}_k > 0$, an increase in σ_k per se causes an increase in $p(k)$. In contrast, if $\bar{\mu}_k < 0$, then the precision effect opposes the stock effect, and an increase in σ_k per se makes the posterior evaluation more likely to fall below the lenient cutoff: with $\bar{\mu}_k < 0$, a rise in σ_k per se causes a reduction in $p(k)$. (See Figure 1b.)

Figure 1a. Stock effect: decrease in critical threshold $\bar{\mu}_k$

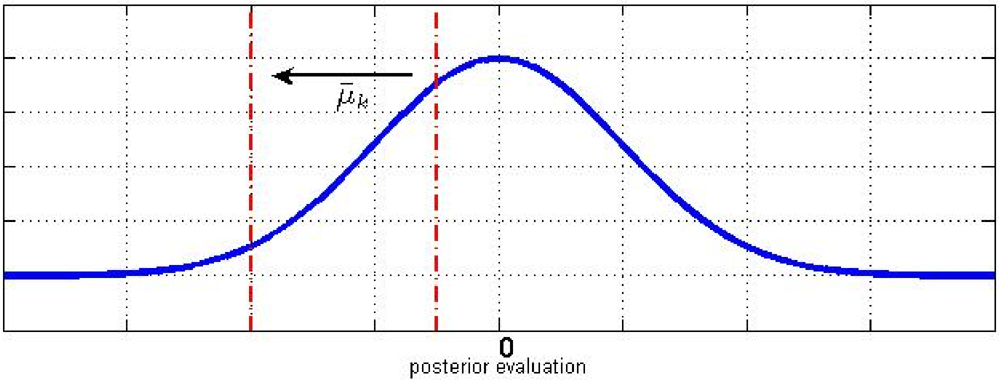
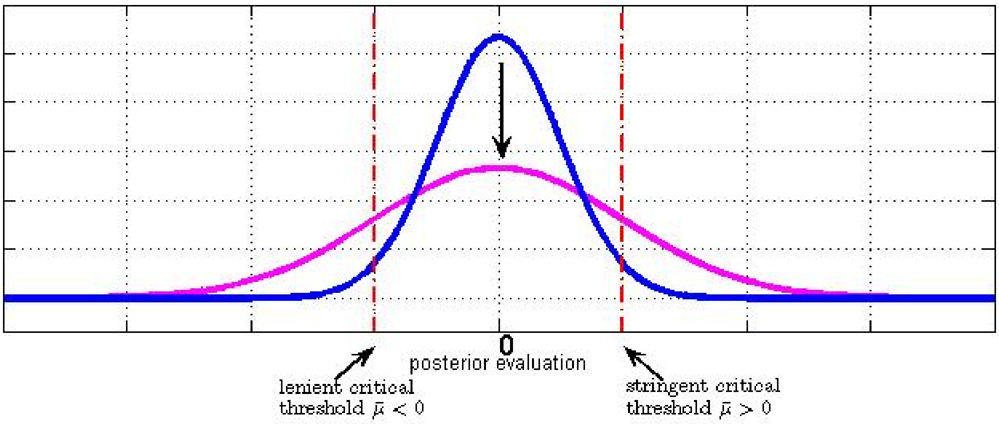


Figure 1b. Precision effect: increase in variance of posterior evaluation



If the precision effect is either outweighed by the stock effect or reinforces it, we will say that retentions are *self-reinforcing*, while if the precision effect opposes and outweighs the stock effect, we will say that retentions are *self-limiting*:

Definition 1. *Retention decisions are self-reinforcing in state k if under the optimal decision rule $p(k+1) > p(k)$ and self-limiting in state k if $p(k+1) < p(k)$.*

If the precision effect reinforces the stock effect, we will say that retentions are *strongly self-reinforcing*:

Definition 2. *Retention decisions are strongly self-reinforcing in state k if under the optimal decision rule $\bar{\mu}_k > 0$, or equivalently $p(k) < \frac{1}{2}$.*

We now state this section's main result, which identifies conditions on the parameters m and δ under which retentions are i) self-reinforcing in all states; ii) strongly self-reinforcing in all states; or iii) self-limiting in at least some states.

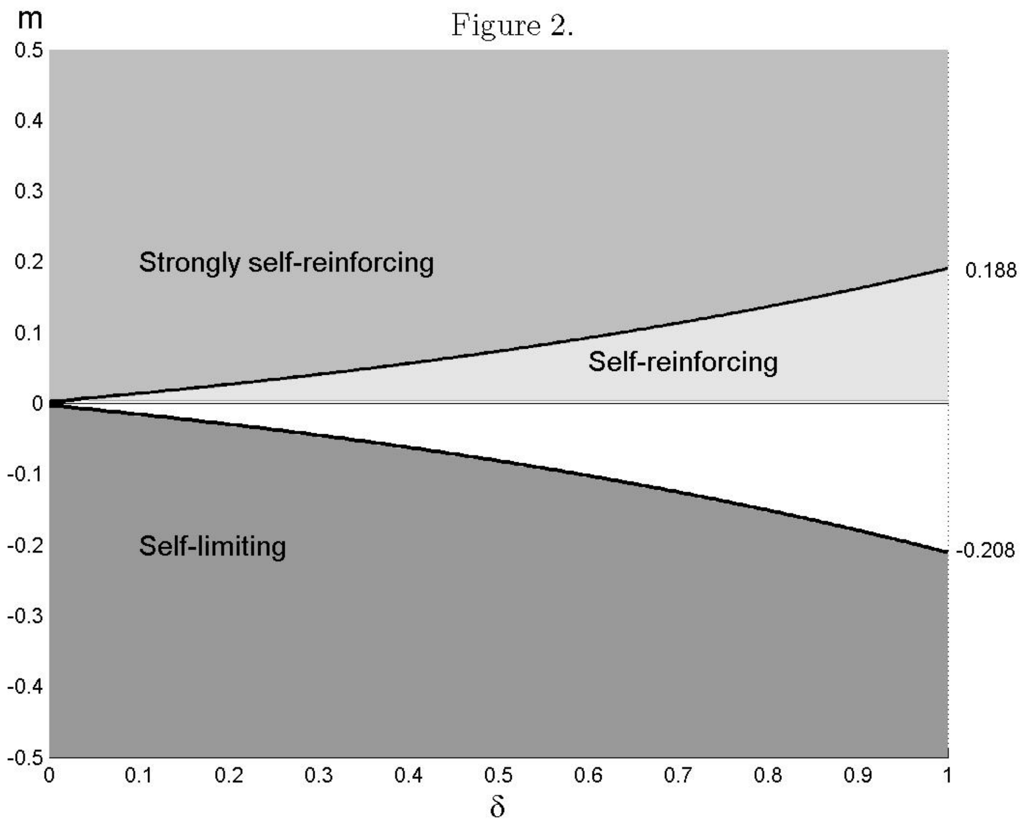
Proposition 1. i) *For all $m \geq 0$ and for all $\delta \in (0,1)$, under the optimal decision rule $p(k)$ is strictly increasing in k for all $k \in \mathbf{Z}^+$, that is, retentions are self-reinforcing in all states.*

ii) *For all $\delta \in (0,1)$, there exists $\bar{m}(\delta) > 0$ such that for $m > \bar{m}(\delta)$, under the optimal decision rule $p(k) < \frac{1}{2}$ for all $k \in \mathbf{Z}^+$, that is, retentions are strongly self-reinforcing in all states. Furthermore, $\bar{m}(\delta)$ is strictly increasing in δ, $\lim_{\delta \to 0} \bar{m}(\delta) = 0$, and $\lim_{\delta \to 1} \bar{m}(\delta) \approx .188$.*

iii) *For all $\delta \in (0,1)$, there exists $\underline{m}(\delta) < 0$ such that for $m < \underline{m}(\delta)$, there exists at least one $k \in \mathbf{Z}^+$ such that under the optimal decision rule $p(k+1) < p(k)$, that is there is at least one state in which retentions are self-limiting. Furthermore, $\underline{m}(\delta)$ is strictly decreasing in δ, $\lim_{\delta \to 0} \underline{m}(\delta) = 0$, and $\lim_{\delta \to 1} \underline{m}(\delta) \approx -.208$.*

Proposition 1 indicates that for positive m, retentions are always self-reinforcing under the optimal decision rule. Further, for any δ there exists an $\bar{m} > 0$ and an $\underline{\mathrm{m}} < 0$ such that retentions are strongly self-reinforcing whenever m exceeds $\bar{m}$ (i.e., whenever m is sufficiently positive) and are self-limiting in at least some states whenever m is less than $\underline{\mathrm{m}}$ (i.e., whenever m is sufficiently negative). These findings are represented in Figure 2.

Figure 2.



The detailed proof of Proposition 1 is in the Appendix, but we describe the strategy of the proof here, before providing some intuition for the results. Defining the transformation $\mathbf{G} : B \mapsto B$, in accord with the Bellman equation (4), by

$$\mathbf{G}(V(k)) = \delta E_{\bar{x}} \left[\max \left\{ \tilde{\mu}(k, \bar{x}) + V(k+1), m + V(1) \right\} \right], \tag{8}$$

where $B \equiv \{V \mid V : \mathbf{Z}^+ \mapsto \mathbf{R}, \text{bounded}\}$, we show that $\mathbf{G}$ is a contraction mapping and therefore has a unique fixed point, which corresponds to the optimal decision rule. Since the property of the optimal decision rule in which we are interested, namely the variation of $p(k)$ with k, depends on the value function only through the expression $A(V(k))$ defined in (6), we define a transformation $\mathbf{H} : B \mapsto B$ which operates directly on functions $A(V) : \mathbf{Z}^+ \mapsto \mathbf{R}$, as follows:

$$\mathbf{H}(A(V)) \equiv A(\mathbf{G}(V)). \tag{9}$$

The usefulness of this transformation follows from the observation that if V^* is the unique fixed point of $\mathbf{G}$, then

$$\mathbf{H}(A(V^*)) = A(\mathbf{G}(V^*)) = A(V^*),$$

so $A(V^*)$ is a fixed point of $\mathbf{H}$. We will use the properties of $\mathbf{H}$ to show that $A(V^*(k))$ must be monotonically decreasing in k. This will then imply, through equation (7), that $p(k)$ is monotonically increasing in k.

To derive the form of $\mathbf{H}$, we write

$$\begin{aligned} \mathbf{H}(A(V(k))) &= A(\mathbf{G}(V(k))) \\ &= \frac{m + \mathbf{G}(V(1)) - \mathbf{G}(V(k+1))}{\sigma_k}, \end{aligned} \tag{10}$$

and then, using (6) and (7) in equation (8), we express $\mathbf{G}(V(i))$, as[9]

$$\begin{aligned} \mathbf{G}(V(i)) &= \delta p(i) E[\tilde{\mu}(i,\bar{x}) \mid \tilde{\mu}(i,\bar{x}) > m + V(1) - V(i+1)] \\ &+ \delta[p(i)V(i+1) + (1-p(i))(m+V(1))] \end{aligned} \tag{11}$$

Since $\frac{\tilde{\mu}(i,\bar{x})}{\sigma_i}$ has a standard normal distribution and since for a standard normal random variable $\tilde{x}$,

$$E(\tilde{x} \mid \tilde{x} > c) = \frac{\phi(c)}{1-\Phi(c)},$$

it follows that

$$\begin{aligned} E\left[\tilde{\mu}(i,\bar{x}) \mid \tilde{\mu}(i,\bar{x}) > m + V(1) - V(i+1)\right] &= \sigma_i \frac{\phi\left[\frac{m+V(1)-V(i+1)}{\sigma_i}\right]}{1-\Phi\left[\frac{m+V(1)-V(i+1)}{\sigma_i}\right]} \\ &= \sigma_i \frac{\phi(A_i)}{1-\Phi(A_i)}. \end{aligned}$$

Using this and (7), we can rewrite (11) as

$$\mathbf{G}(V(i)) = \delta\left[\sigma_i \phi(A_i) + p(i)V(i+1) + (1-p(i))(m+V(1))\right]. \tag{12}$$

[9]We will henceforth use $V(k)$ and V_k interchangeably, and will let $A(k)$ or A_k represent $A(V(k))$.

Evaluating (12) for $i = 1$ and $i = k + 1$, substituting into (10), and again using (7) finally yields

$$(\mathbf{H}A)_k = -\frac{\delta}{\sigma_k}\left[\sigma_{k+1}R(A_{k+1}) - \sigma_1 R(A_1)\right] + \frac{m}{\sigma_k}, \quad \forall k \in \mathbf{Z}^+, \tag{13}$$

where $R : \mathbf{R} \mapsto \mathbf{R}$ is defined by

$$R(x) \equiv \phi(x) - (1 - \Phi(x))x. \tag{14}$$

Next we define, for all $m \in (-\infty, \infty)$ and for all $\delta \in (0, 1)$, a closed subset $\mathcal{S}(m, \delta)$ of B as follows:

$$\mathcal{S}(m, \delta) \equiv \{V \in B \,|\, \text{A(V) satisfies Conditions i,ii, and iii}\}, \tag{15}$$

where Conditions i,ii, and iii are given by:

Condition i (Monotonicity) $\forall k \in \mathbf{Z}^+$, $A_k \geq A_{k+1}$;

Condition ii (Upper Bound) $\forall k \in \mathbf{Z}^+$, $A_k \leq U(m, \delta)$;

Condition iii (Lower Bound) $\forall k \in \mathbf{Z}^+$, $A_k \geq L(m, \delta)$;

where $U : (-\infty, \infty) \times (0, 1) \mapsto \mathbf{R}$ and $L : (-\infty, \infty) \times (0, 1) \mapsto \mathbf{R}$ are defined implicitly by

$$U(m, \delta) = -c\delta R(U(m, \delta)) + \frac{m}{\sigma_1}, \tag{16}$$

$$L(m, \delta) = -\delta\left[R(L(m, \delta)) - \sigma_1 R(U(m, \delta))\right] + m, \tag{17}$$

and where c is a constant defined by $c \equiv \frac{\sigma_2 - \sigma_1}{\sigma_1}$.

To prove part i) of Proposition 1, we show that if $m \geq 0$, then for all $\delta \in (0, 1)$, $\mathbf{G}$ maps $\mathcal{S}(m, \delta)$ into itself. We do this by using the expression (13) for $\mathbf{H}$ to show that if $A(V)$ satisfies Conditions i, ii, and iii, then $\mathbf{H}(A(V))$ does so as well. Consequently, for the fixed point V^* of $\mathbf{G}$, $A(V^*)$ must satisfy these conditions, or equivalently, the fixed point V^* of $\mathbf{G}$ is in $\mathcal{S}(m, \delta)$. Hence, $A(V^*(k))$ is decreasing in k, so the optimal retention probability $p(k)$ is increasing in k.

Similarly, to show part ii), we show that if $L(m, \delta) > 0$ in addition to $m \geq 0$, then for the fixed point V^* of $\mathbf{G}$, not only are Conditions i, ii, and iii satisfied but also $A(V^*(k)) > 0$ and hence $p(k) < \frac{1}{2}$. For $m \geq 0$, the condition $m > \bar{m}(\delta)$ is equivalent to $L(m, \delta) > 0$.

Finally, to prove part iii), we show that if V^* is a fixed point of $\mathbf{G}$ and $A(V^*(k))$ is monotonically decreasing in k, then $L(m, \delta) \leq U(m, \delta)$. That is, $L(m, \delta) \leq U(m, \delta)$ is a necessary condition for $p(k)$ to be increasing in k for all k. The condition $m < \underline{\mathrm{m}}(\delta)$ is equivalent to $U(m, \delta) < L(m, \delta)$.

To understand these results qualitatively, recall that the probability of retention in state k, $p(k)$, can be expressed as

$$p(k) = 1 - \Phi \left[\frac{\bar{\mu}_k}{\sigma_k} \right], \tag{18}$$

where $\bar{\mu}_k = m + V(1) - V(k+1)$. The precision effect of an increase in k on $p(k)$ operates through the increase in σ_k and the stock effect through the reduction in $\bar{\mu}_k$. Thus the precision effect is proportional to

$$\frac{dp(k)}{d\sigma_k} = \frac{d}{d\sigma_k} \left(1 - \Phi(\frac{\bar{\mu}_k}{\sigma_k}) \right) = \frac{\bar{\mu}_k}{(\sigma_k)^2} \phi(\frac{\bar{\mu}_k}{\sigma_k}). \tag{19}$$

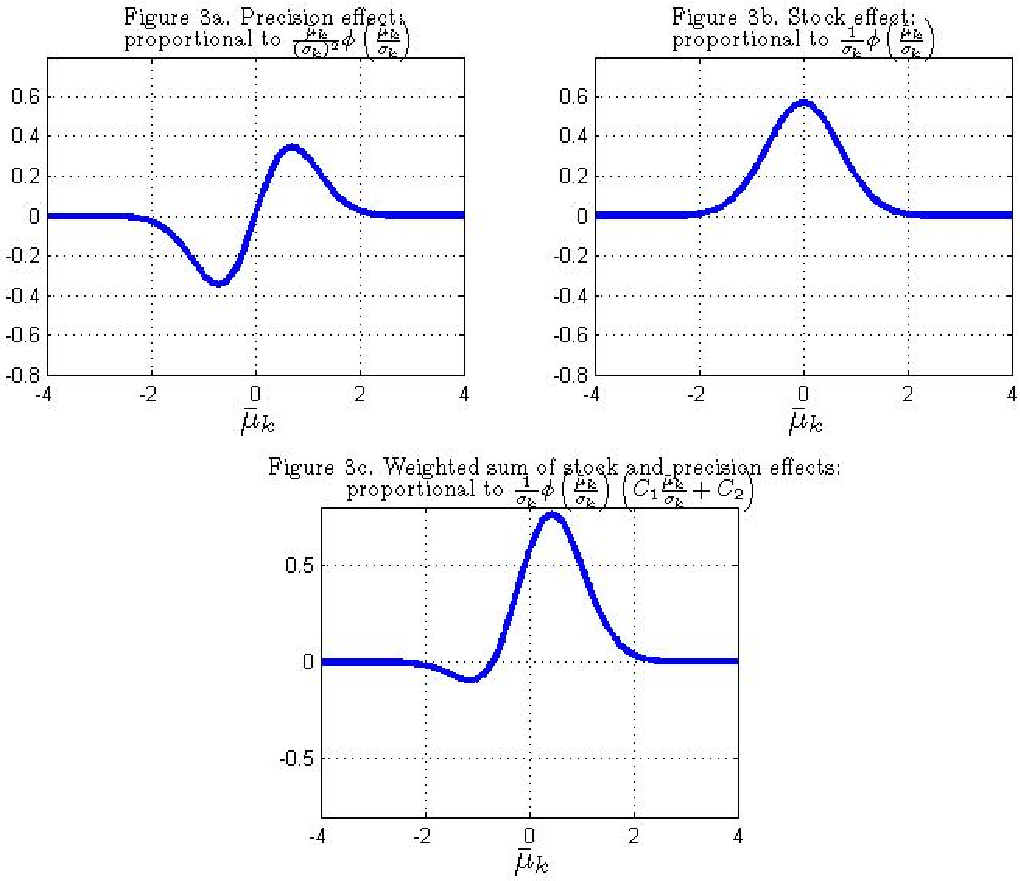
Equation (19) confirms that whether the precision effect acts to make $p(k)$ rise or fall with k depends on whether $\bar{\mu}_k$ is greater than or less than 0. It also shows how the *magnitude* of the precision effect varies with $\bar{\mu}_k$: when $\bar{\mu}_k = 0$, the precision effect is zero, and as $|\bar{\mu}_k|$ goes from 0 to ∞, the magnitude of the precision effect first rises and then falls. (See Figure 3a.) Using similar reasoning, the stock effect of an increase in k on $p(k)$ is proportional to[10]

$$-\frac{dp(k)}{d\bar{\mu}_k} = -\frac{d}{d\bar{\mu}_k} \left(1 - \Phi(\frac{\bar{\mu}_k}{\sigma_k}) \right) = \frac{1}{\sigma_k} \phi(\frac{\bar{\mu}_k}{\sigma_k}). \tag{20}$$

Equation (20) confirms that the stock effect always acts to make $p(k)$ rise with k. Moreover, it shows that the magnitude of the stock effect is proportional to $\phi(\frac{\bar{\mu}_k}{\sigma_k})$, so

[10]The negative sign in front of the derivative reflects the fact that as k rises, $\bar{\mu}_k$ falls.

as $|\bar{\mu}_k|$ goes from 0 to ∞, this magnitude falls. (See Figure 3b.)



Figures plotted for $\sigma_k = 0.7, C_1 = 0.5, C_2 = 0.5$

The reason for focusing on how the magnitudes of the stock and precision effects vary with the optimal cutoff $\bar{\mu}_k$ is that increases in the exogenous (ex ante) acceptance hurdle m translate into increases in $\bar{\mu}_k$. To see how the relative sizes of the stock and precision effects vary with m, take a positively weighted sum of the expressions in (19) and (20), with the weights $C_1 > 0$ and $C_2 > 0$ arbitrary constants (arbitrary to reflect the fact that (19) and (20) are merely *proportional* to the strengths of the precision and stock effects, respectively). The weighted sum is then:

$$C_1 \frac{dp(k)}{d\sigma_k} + C_2 \left(-\frac{dp(k)}{d\bar{\mu}_k} \right) = \frac{1}{\sigma_k} \phi\left(\frac{\bar{\mu}_k}{\sigma_k}\right) \left[C_1 \frac{\bar{\mu}_k}{\sigma_k} + C_2 \right]. \tag{21}$$

Now observe that, regardless of the precise values of C_1 and C_2, this sum is negative for $\bar{\mu}_k$ less than some strictly negative critical value and positive for $\bar{\mu}_k$ greater than this critical value. (See Figure 3c.) This observation explains why for m sufficiently negative, retention decisions are (at least in some states) self-limiting (i.e. the precision effect outweighs the stock effect), whereas as m rises, retention decisions become self-reinforcing (i.e. the precision effect is outweighed by the stock effect), with further increases in m leading to strongly self-reinforcing behavior (i.e. the precision effect reinforcing the stock effect).

Proposition 1 also shows that as the discount factor δ falls, the region in which retention decisions are self-reinforcing but not strongly self-reinforcing shrinks. In the limit as δ goes to 0, retentions are strongly self-reinforcing in all states for values of m greater than 0 and self-limiting in all states for values of m less than 0. These results are easily explained. As δ shrinks to 0, the stock effect disappears: the optimal cutoff for the posterior evaluation, $\bar{\mu}_k$, becomes independent of k, approaching the value m. Consequently, only the precision effect remains, and whether the precision effect makes $p(k)$ rise or fall with k depends on the sign of the optimal cutoff, hence on the sign of m.

4 The General Model

We now present three more complex scenarios, in each of which, as in Section 3, a decision-maker chooses each period between acceptance and rejection, with successive acceptances gradually increasing the stock of precision and each rejection fully dissipating it. Whereas the simple model of Section 3 generates a Bellman equation containing only two exogenous parameters, m and δ, the Bellman equation deriving from our more complex scenarios contains two additional parameters. The first represents the increment in precision resulting from each successive retention and the second the baseline level to which the precision reverts after a termination. These parameters are measured relative to the prior precision of each period's uncertain payoff opportunity, which we continue to fix at 1. After describing our three scenarios, we then analyze how these new parameters affect the form of history-dependence in optimal decision-making, once again identifying conditions under which acceptances are strongly self-reinforcing, self-reinforcing, or self-limiting.

4.1 A Bandit Model with Market Wages

First, consider the following "bandit" problem of a firm deciding when to replace a piece of equipment or an employee, where this factor must be paid a noisy estimate of its "market value". Suppose that the factor's unknown intrinsic value is η. In period t, the firm observes the realization of revenue

$$x_t = \eta + \epsilon_t$$

that will accrue at the beginning of period $t+1$ if it continues to employ the factor in period t. The realization of ϵ_t represents an interaction of the factor with the environment that is specific both to the firm and to period t, such as how suited an employee or machine is for that period's task. Having observed x_t, the firm chooses whether to retain or terminate this factor. If the factor is retained, the firm must pay at the beginning of period $t+1$ a wage

$$w_t = \eta + u_t$$

whose realization is not observed until the end of period t, after the retention decision is made. This "market wage" represents the estimated value of the factor to some competing firm with a different environmental interaction u_t. If instead the factor is terminated, a new factor is hired from a population with the same prior distribution as that from which η was originally drawn, but one period elapses before the new factor generates any revenue; during that period, the flow payoff is the constant $m \in (-\infty, \infty)$. Thus $-m$ represents the known fixed cost of replacing a factor.

Draws of ϵ, η, and u are independent of one another and across time and are distributed respectively according to $N(0,1)$, $N(0,1/b)$, and $N(0,1/(s-1))$.

Note that unlike the overlapping generations model in Section 3, here retention implies keeping the *same* factor η. It is for this reason that we refer to such a model as a "bandit model". However, unlike in most bandit problems, here the firm must pay the incumbent factor a market wage each period. As a result, the firm's flow payoff from retention is the difference between the firm-specific shock to revenue and the noise in the market wage, i.e. $\epsilon_t - u_t$, and the unconditional distribution of this payoff is iid across time.[11] Nevertheless, because the retention decision is based on knowledge of the potential revenue $x_t = \eta + \epsilon_t$, the more the firm knows about η the better able it is to estimate ϵ_t. And since the firm's stock of information about the incumbent factor is increased each time it chooses to retain, the longer the unbroken sequence of retentions, the better the quality of the next decision.

Let λ_t and k_t be the posterior mean and precision, respectively, of the distribution of η, the intrinsic value of the incumbent factor, computed at the start of period t (i.e. after observation of x_{t-1} and, if the incumbent factor was retained in period $t-1$, of w_{t-1}). Denote $x_t - \lambda_t$ by $\bar{x}_t$, as in Section 3. After the observation of revenue x_t, the posterior mean of ϵ_t will be

$$E(\epsilon_t|k_t, \bar{x}_t) = \frac{k_t \bar{x}_t}{k_t + 1}.$$

This posterior mean, regarded as a random variable, is, conditional on information available at the start of period t, distributed according to

$$E(\epsilon_t|k_t, \bar{x}_t) \sim N\left(0, \frac{k_t}{k_t + 1}\right).$$

That is, it has the same distribution (and for the same reason) as $\tilde{\mu}(k, \bar{x})$ in equation (3) of Section 3. Since this distribution is independent of λ_t, just as in Section 3 the relevant state can be summarized solely by a value of precision, here the precision k_t of the beliefs about the incumbent factor η.

If the firm chooses to retain the incumbent factor in period t, the expected flow profit that will accrue in period $t+1$ is

$$E\left(x_t - w_t\,|k, \bar{x}_t\right) = E\left(\epsilon_t - u_t\,|k, \bar{x}_t\right) = E\left(\epsilon_t\,|k, \bar{x}_t\right).$$

Also, since both x_t and w_t are observed before the end of period t, the precision of the beliefs about η at the start of period $t+1$ will be $k_{t+1} = k_t + 1 + (s-1) = k_t + s$. If instead the factor is terminated in period t, then a new factor is installed in period $t+1$ (with the first revenue accruing in period $t+2$), and the flow payoff in period

[11] See Section 2.1 for more discussion of the contrast between our models and the bandit literature.

$t+1$ is the constant m. The precision of beliefs about the new factor assumes its prior value, so in this case the state variable will become $k_{t+1} = b$.

Let $V(k)$ once again represent the expected discounted present value of future payoff in a state with precision k under the optimal decision policy. To write the Bellman equation, we use the fact that $E(\epsilon|k, \bar{x})$ has the same distribution as $\tilde{\mu}(k, \bar{x})$, so that $\forall k \in \mathbf{Q}$,

$$V(k) = \delta E_{\bar{x}} \left[\max \left\{ \tilde{\mu}(k, \bar{x}) + V(k+s), m + V(b) \right\} \right], \tag{22}$$

where

$$\mathbf{Q} \equiv \left\{ k \in \mathbf{R}_{+} \mid k = (b + js \text{ for some } j \in \{0 \cup \mathbf{Z}^{+}\} \right\}.$$

Equation (22) is the generalized Bellman equation we consider in this section. The parameter s represents the increment in precision resulting from each successive retention and the parameter b the baseline level to which the precision reverts after a termination. Equation (4) of the basic model of Section 3 corresponds to the case where $s = 1$ and $b = 1$.

4.2 Accumulating Information about an Uncertain Environment

Consider now a decision-maker that, in each period t, evaluates a new potential opportunity (say a consumption opportunity or an investment opportunity), whose unknown intrinsic value is ν_t. If the period-t opportunity is undertaken, the decision-maker receives in period $t+1$ a once-off payoff $y_t = \nu_t + u_t$, whose realization is observed only at the end of period t. The decision whether or not to undertake the opportunity is based on a noisy signal $x_t = \nu_t + \epsilon$, where ϵ represents uncertainty about the current "environment". The environment ϵ remains the same unless an opportunity is rejected, in which case the flow payoff next period is the constant $m \in (-\infty, \infty)$ and there is a new draw of environmental uncertainty. With each successive acceptance of a new opportunity, the decision-maker's stock of information about the current environment increases, thus improving the quality of his next decision; rejection of an opportunity, by contrast, reduces the quality of the next decision to a baseline level corresponding to the prior information about the new environment. If the opportunities under consideration are, for example, investment projects internally generated within a firm, then the environmental uncertainty contaminating the observed signals could represent the unknown biases of the manager who proposes the projects. The assumption that a new ϵ is drawn following a rejection could reflect the manager's quitting if ever her recommended project is turned down and then being replaced by someone with different unknown biases.

Draws of ν, ϵ, and u are independent of one another and across time and are distributed respectively according to $N(0,1)$, $N(0,1/b)$, and $N(0,1/(s-1))$.

Comparing this scenario with the bandit model in the previous subsection, here, when there is a string of continued acceptances, the environment ϵ remains fixed, while the intrinsic value ν_t of the opportunity varies across periods. In contrast, in the bandit model, continued acceptances keep fixed the incumbent factor η, while the environmental interaction ϵ_t varies across periods. In both cases, continued acceptances enable the decision-maker to build up a stock of knowledge about the fixed component, which is valuable insofar as, and only insofar as, it allows for a more precise estimate of the variable component. Specifically, in both cases, only the variable component (drawn anew each period), and not the fixed component, directly enters the decision-maker's payoff function: in the present model, the fixed environment ϵ affects the signal but not the payoff from the opportunity ν_t, while in the bandit model, the firm's need to pay the market wage implies that its net payoff depends on the environmental interaction ϵ_t but not on the fixed value of the incumbent factor η.

While the present scenario therefore differs from the bandit scenario in the interpretations of the fixed and variable components, formally, the analysis of this scenario is very similar to that in Section 4.1. Let λ_t and k_t be the mean and precision, respectively, of the distribution of ϵ, the current environment, computed at the start of period t (i.e. after observation of x_{t-1} and, if the period-$t-1$ opportunity was undertaken, of y_{t-1}). As before, define $\bar{x}_t \equiv x_t - \lambda_t$. Then the posterior mean of the period-t opportunity's value, ν_t, given $\bar{x}_t$, has a distribution, computed at the start of period t, which is the same as that of $\tilde{\mu}(k, \bar{x})$ in Section 3, given by equation (3). Once again, the state variable in the value function is just k_t.

If, after observing x_t, the decision-maker accepts the period-t opportunity, then in the next period the expected flow payoff will be $\tilde{\mu}(k_t, \bar{x}_t)$; the precision of beliefs about the current environment will increase to $k_{t+1} = k_t + s$, since by the end of the period the decision-maker observes both x_t and y_t.[12] If instead the period-t opportunity is rejected, the next period's flow payoff will be m, and the precision of beliefs about the (new) environment will revert to the baseline level $k_{t+1} = b$. Just as in Section 4.1, therefore, the Bellman equation is given by (22).

4.3 An Overlapping Generations Model with Comparative Performance Evaluation

Our third scenario formalizes the overlapping generations story in the introduction, in which each successive retention of a new employee further improves the subsequent decision by providing more valuable comparative performance information. This scenario generalizes Section 3's model by relaxing the assumption that only the joint output of the junior and senior employees is observed.[13] To further increase generality, we will describe the scenario here in terms of general factors of production rather than employees specifically.

In each period t, the firm observes a noisy signal $x_t = \nu_t + \epsilon_t$ of the intrinsic productivity ν_t of a new factor. If it chooses not to retain the factor, then its flow payoff in period $t+1$ is a constant m. If it chooses instead to retain the factor, then it receives in period $t+1$ a payoff $y_{t+1} = \nu_t + \epsilon_{t+1}$, after which the factor is of no further use. Draws of ν and ϵ are independent of one another and across time and are distributed respectively according to $N(0,1)$ and $N(0,1/b)$.

The crucial feature of this scenario is that the noise ϵ_{t+1} affecting the payoff from the retained factor ν_t is the same as that affecting the signal $x_{t+1} = \nu_{t+1} + \epsilon_{t+1}$ about

[12] The posterior precision on ν_t, given y_t, is s, so the posterior precision on ϵ, given x_t and y_t, is $k_t + s$.

[13] An alternative generalization of Section 3's model, that would yield Bellman equation (22)with $s = 1$, would simply involve the abilities of newly-hired seniors being distributed according to $N(m, 1/b)$. The case $b > 1$ would reflect the firm's ability to see some signal of a new senior's performance at a prior job, thus giving it a more precise evaluation of ability when hiring a new senior than a new junior.

the next factor ν_{t+1}. This feature represents the fact that the productivity of the new factor is evaluated in the same environment in which the retained factor produces.[14] If the firm decides to retain the factor ν_t, the observation of its subsequent performance y_{t+1} provides information about ϵ_{t+1} which improves the quality of the decision about the next factor ν_{t+1}. Furthermore, because in this case more is learned about ν_{t+1} than about ν_t, retention of ν_{t+1} allows an even more precise inference about the productivity of the following period's factor ν_{t+2}, and so on. Thus, retention of a factor, by making available comparative performance information, improves the next decision, and the greater the number of successive retentions, the more valuable is the comparative performance information, so the greater is the quality of the following decision. On the other hand, if a factor is ever rejected, then the quality of the next period's decision is reduced to a baseline level, since it must be made without the benefit of any comparative performance information, solely on the basis of that period's x signal and the prior information about the ϵ noise term.

Here, let λ_t and k_t be the mean and precision, respectively, of the distribution of ϵ_t, computed before observation of x_t but after observation of y_t (if the factor ν_{t-1} was retained). Define $\bar{x}_t \equiv x_t - \lambda_t$. The posterior mean of the period-t factor's productivity ν_t, given $\bar{x}_t$, has a distribution, computed prior to obseration of $\bar{x}_t$, which is the same as that of $\tilde{\mu}(k, \bar{x})$ in Section 3, given by (3). As before, the relevant state variable in the value function is just k_t.

If, after observing x_t, the firm retains the factor ν_t, then the period-$t+1$ expected flow payoff will be $\tilde{\mu}(k_t, \bar{x}_t)$. Since retention allows the firm to observe both x_t and y_{t+1}, the period-$t+1$ value of the state variable will rise to $k_{t+1} = k_t + 1 + b$.[15] If instead the factor ν_t is rejected, then the period-$t+1$ flow payoff will be m, and the state variable will revert to the prior precision on ϵ_{t+1}, so $k_{t+1} = b$. The Bellman equation is therefore

$$V(k) = \delta E_{\bar{x}} \left[\max \left\{ \tilde{\mu}(k, \bar{x}) + V(k + 1 + b), m + V(b) \right\} \right]. \tag{23}$$

[14]For example, a firm may be using an information processing system that must be replaced the following period. It may be able in the present period to test run a new system, using the information that currently needs to be processed. However, if adopted, the new system will be used to process future information, which differs in nature from current information.

[15]The posterior precision on ν_t, given x_t, is $k_t + 1$, so the posterior precision on ϵ_{t+1}, given x_t and y_{t+1}, is $k_t + 1 + b$.

This is clearly the special case of (22) where $s = 1 + b$.[16]

4.4 Analysis of the General Model

We again want to characterize the form of history-dependence in optimal decision-making, specifically how the probability of an acceptance (retention) depends on the number of consecutive acceptances since the last rejection. Formally, we are again analyzing the dependence of the acceptance probability $p(k)$ on the stock of accumulated information, k. Given the generalized Bellman equation (22), the probability of an acceptance (retention) in state k is

$$p(k) = Pr[\tilde{\mu}(k, \bar{x}) \geq m + V(b) - V(k+s)] = 1 - \Phi\left[\frac{m + V(b) - V(k+s)}{\sigma_k}\right]. \quad (24)$$

With the definition

$$A(V(k)) \equiv \frac{m + V(b) - V(k+s)}{\sigma_k}, \quad \forall k \in \mathbf{Q}, \quad (25)$$

it is once again true that

$$p(k) = 1 - \Phi(A(V(k))), \quad (26)$$

so $p(k)$ increases in k if and only if $A(V(k))$ decreases in k.

It is easy to see that the stock and precision effects of accumulated information continue to operate, and in the same directions as before. The optimal cutoff for the posterior evaluation, $\bar{\mu}_k = m + V(b) - V(k+s)$, is decreasing in k: this is the stock effect, and by itself it acts to make $p(k)$ rise with k. But the variance of the posterior evaluation, $(\sigma_k)^2$, itself rises with k, so if the optimal cutoff is below (above) the prior mean of zero, this variation per se acts to make $p(k)$ fall (rise) with k: this is the precision effect.

Our focus now is on how the increment in precision induced by an acceptance, s, and the baseline level of precision following a rejection, b, affect the likelihood that acceptance decisions are self-reinforcing ($p(k)$ increasing in k) or self-limiting ($p(k)$

[16] A more complex version of this overlapping generations scenario generates (22), but we have chosen the version above for expositional simplicity.

decreasing in k). Recall that the parameters s and b are measured relative to 1, where 1 is the value to which we have normalized the prior precision of the uncertain payoff "opportunity" which the decision-maker is deciding each period whether or not to accept.[17]

The transformation $\mathbf{G}$ derived from the Bellman equation (22),

$$\mathbf{G}(V(k)) = \delta E_{\bar{x}} \left[\max \left\{ \tilde{\mu}(k, \bar{x}) + V(k+s), m + V(b) \right\} \right], \tag{27}$$

is a contraction mapping and so has a unique fixed point. Define $\mathbf{H}$ as in Section 3 by

$$\mathbf{H}(A(V)) \equiv A(\mathbf{G}(V)), \tag{28}$$

so that if V^* is a fixed point of $\mathbf{G}$, then $A(V^*)$ is a fixed point of $\mathbf{H}$. Evaluating $A(\mathbf{G}(V))$ in the same way as we did in Section 3 allows us to write the generalized version of the transformation $\mathbf{H}$ as

$$(\mathbf{H}A)_k = -\frac{\delta}{\sigma_k} \left[\sigma_{k+s} R(A_{k+s}) - \sigma_b R(A_b) \right] + \frac{m}{\sigma_k}, \quad \forall k \in \mathbf{Q}. \tag{29}$$

The intuitive argument we developed following Proposition 1 can be extended to this more general model and suggests that, for given values of the parameters s and b, acceptance decisions are more likely to be self-reinforcing the larger is the "acceptance hurdle" m. This intuitive reasoning is formally confirmed below. To determine how variations in s and b affect the size of the region in (δ, m) space where acceptances are self-reinforcing, it is enlightening to examine the extreme cases of $s \to \infty$ and $s \to 0$. In these cases, we can fully characterize the dynamics of optimal decision-making, for all values of $m \in (-\infty, \infty)$, $\delta \in (0, 1)$, and $b \geq \frac{1}{3}$, and for all states $k \in \mathbf{Q}$.[18]

[17]This uncertain payoff opportunity is the current junior's ability in Section 3, the current firm-specific shock ϵ_t in Section 4.1, and the value of the new project or factor ν_t in Sections 4.2 and 4.3.

[18]We restrict attention to $b \geq \frac{1}{3}$ in order to ensure that the generalization of the bound $U(m, \delta)$ defined in (16) is well-defined.

Consider first the limiting case of $s \to \infty$. In this case,

$$A(V(k)) = \frac{m + V(b) - V(\infty)}{\sigma_k}, \tag{30}$$

and the fixed point of the transformation H solves

$$A_k = -\frac{\delta}{\sigma_k}\left[R(A_\infty) - \sigma_b R(A_b)\right] + \frac{m}{\sigma_k}. \tag{31}$$

The stock effect is *absent* here, since regardless of the current state, one more acceptance causes the stock of precision to rise to ∞, and hence the optimal cutoff for the posterior evaluation is $m + V(b) - V(\infty)$, *independent* of the current state k. The sign of the precision effect alone therefore determines whether acceptances are self-reinforcing (in which case they are also strongly self-reinforcing) or self-limiting. From (30) it is clear that the boundary case, where A_k and hence $p(k)$ are independent of k, is where $A_k = 0$ for all $k \in \mathbf{Q}$, that is, $m + V(b) - V(\infty) = 0$. Substituting $A_k = 0$ into both sides of (31) then transforms it into

$$0 = -\delta(1 - \sigma_b)\phi(0) + m,$$

since $R(0) = \phi(0)$. These observations establish:

Proposition 2. *For any $b \in [\frac{1}{3}, \infty)$, consider the limit as $s \to \infty$.*

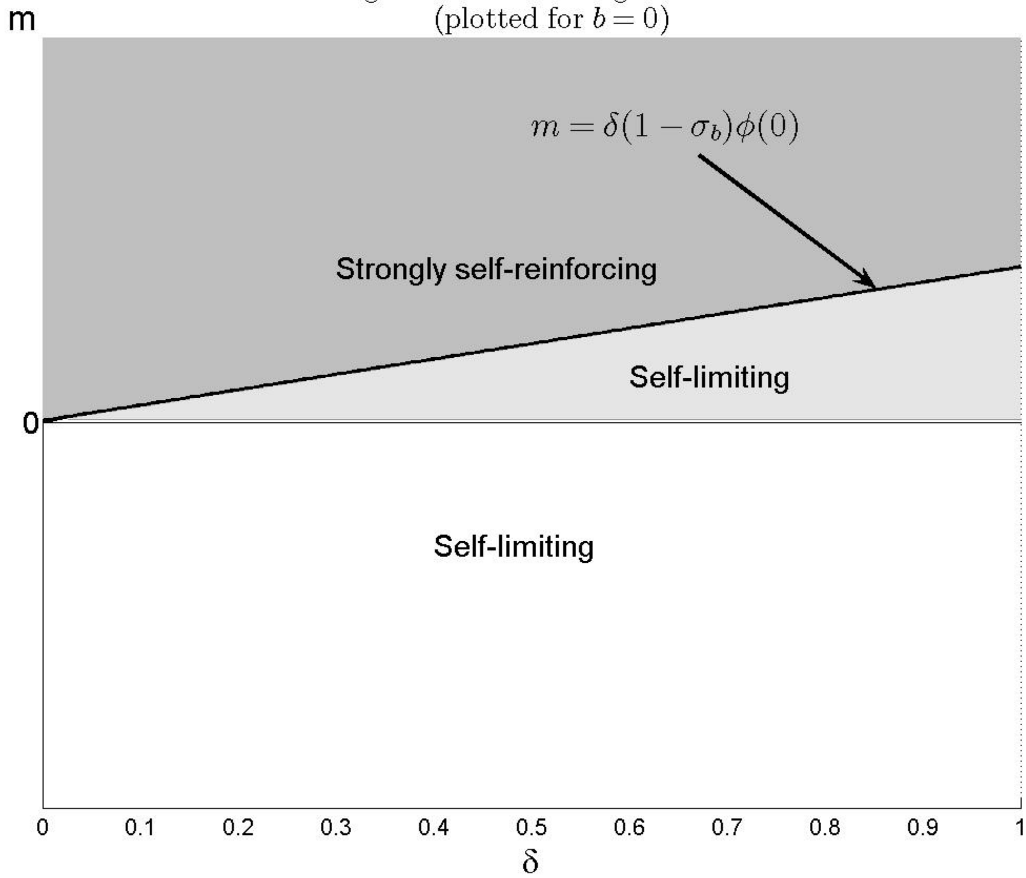
i) *For all $\delta \in (0, 1)$, if $m > \delta(1 - \sigma_b)\phi(0)$, then under the optimal decision rule $p(k)$ is strictly increasing in k and $p(k) < \frac{1}{2}$ for all $k \in \mathbf{Q}$, that is, retentions are strongly self-reinforcing in all states.*

ii) *For all $\delta \in (0, 1)$, if $m < \delta(1 - \sigma_b)\phi(0)$, then under the optimal decision rule $p(k)$ is strictly decreasing in k for all $k \in \mathbf{Q}$, that is, retentions are self-limiting in all states.*

iii) *The critical value of m at which $p(k)$ is independent of k, $\delta(1-\sigma_b)\phi(0)$, is increasing in δ and decreasing in b.*

These findings are represented in Figure 4.

Figure 4. The limiting case of $s \to \infty$
(plotted for $b = 0$)



Now consider the other limiting case of $s \to 0$. Here,

$$A(V(k)) = \frac{m + V(b) - V(k)}{\sigma_k},$$

and the fixed point of $\mathbf{H}$ solves

$$A_k = -\frac{\delta}{\sigma_k}\left[\sigma_k R(A_k) - \sigma_b R(A_b)\right] + \frac{m}{\sigma_k}. \tag{32}$$

This can be rewritten as

$$A_k + \delta R(A_k) = \frac{\delta \sigma_b R(A_b) + m}{\sigma_k}. \tag{33}$$

In this setting, both the stock effect and the precision effect continue to operate. These effects exactly offset each other when A_k is independent of k. Since the left-hand side of

(33) is strictly increasing in A_k, A_k is independent of k if and only if $\delta\sigma_b R(A_b)+m=0$. Now setting $k=b$ in (33) shows that $A_b=\frac{m}{\sigma_b}$, and therefore

$$A_k \text{ is independent of } k \quad \Longleftrightarrow \quad \frac{m}{\sigma_b}+\delta R(\frac{m}{\sigma_b})=0. \tag{34}$$

If $\frac{m}{\sigma_b}+\delta R(\frac{m}{\sigma_b})>0$ (respectively, <0), then (33) shows that A_k is strictly decreasing (respectively, strictly increasing) in k. Since $\frac{m}{\sigma_b}+\delta R(\frac{m}{\sigma_b})$ is increasing in m, it follows that for this limiting case of $s\to 0$, equation (34) implicitly defines the locus of points in (δ, m) space above which retentions are self-reinforcing in all states and below which they are self-limiting in all states. Furthermore, retentions are strongly self-reinforcing in all states if and only if $\bar{\mu}_k \equiv m+V(b)-V(k)>0$ for all k. Since $\bar{\mu}_k$ is decreasing in k, this is equivalent to $\bar{\mu}_\infty>0$. Now $\bar{\mu}_\infty=0$ if and only if $A_\infty=0$, and setting $k=\infty$ in (33) shows that A_∞ solves

$$A_\infty+\delta R(A_\infty)=\delta\sigma_b R(\frac{m}{\sigma_b})+m.$$

Hence

$$A_\infty=0 \qquad \Longleftrightarrow \qquad \frac{m}{\sigma_b}+\delta R(\frac{m}{\sigma_b})=\frac{\delta\phi(0)}{\sigma_b}.$$

Thus for $s\to 0$, retentions are strongly self-reinforcing in all states if and only if $\frac{m}{\sigma_b}+\delta R(\frac{m}{\sigma_b})>\frac{\delta\phi(0)}{\sigma_b}$. Proposition 3 summarizes and formalizes the above arguments:

Proposition 3. *For any $b\in[\frac{1}{3},\infty)$, consider the limit as $s\to 0$. For all $\delta\in(0,1)$, there exist functions $\underline{m}^0(\delta,b)<0$ and $\bar{m}^0(\delta,b)>0$ such that*

i) *if $m>\bar{m}^0(\delta,b))$, then under the optimal decision rule $p(k)$ is strictly increasing in k and $p(k)<\frac{1}{2}$ for all $k\in\mathbf{Q}$, that is, acceptances are strongly self-reinforcing in all states;*

ii) *if $m\in(\underline{m}^0(\delta,b),\bar{m}^0(\delta,b))$, then under the optimal decision rule $p(k)$ is strictly increasing in k for all $k\in\mathbf{Q}$, but for k sufficiently large, $p(k)>\frac{1}{2}$, that is, acceptances are self-reinforcing in all states but for sufficiently large states are not strongly self-reinforcing;*

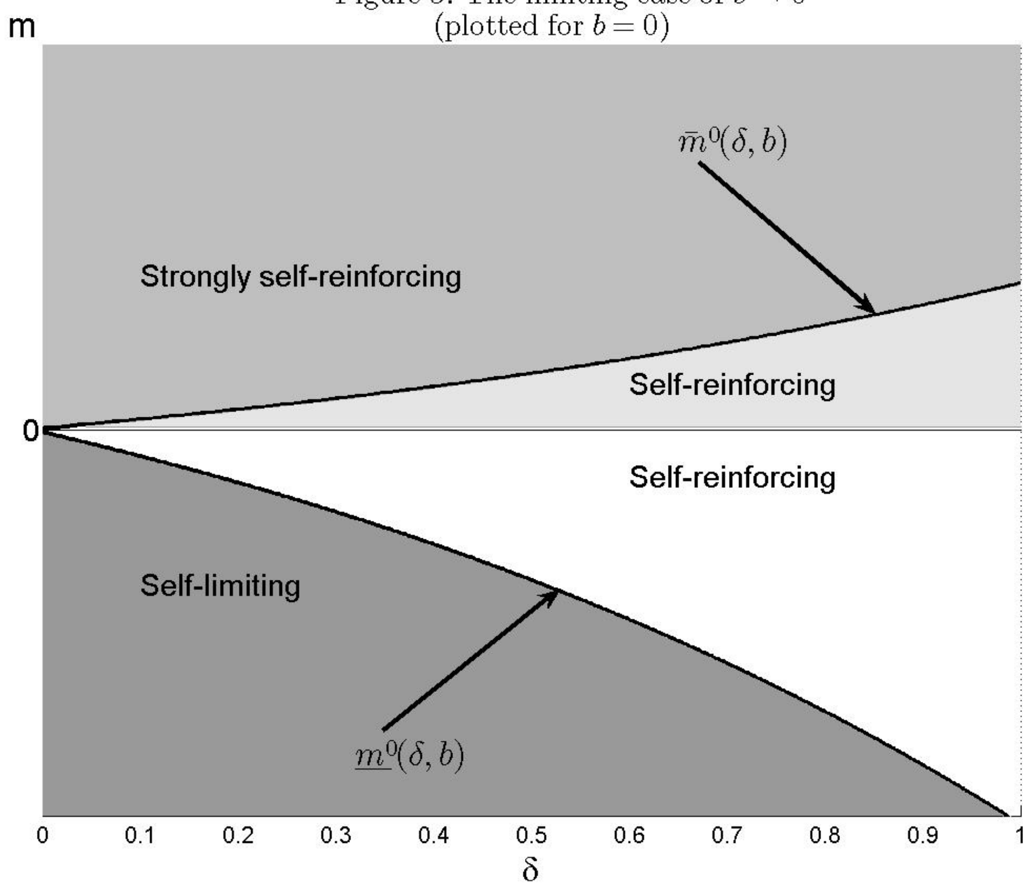
iii) *if* $m < \underline{m}^0(\delta, b)$*, then* $p(k)$ *is strictly decreasing in* k *for all* $k \in \mathbf{Q}$*, that is, acceptances are self-limiting in all states.*

The function $\bar{m}^0(\delta, b)$ *solves* $\frac{m}{\sigma_b} + \delta R(\frac{m}{\sigma_b}) = \frac{\delta\phi(0)}{\sigma_b}$*, and for all* $\delta \in (0, 1)$ *and for all* $b \in [\frac{1}{3}, \infty)$*,* $\bar{m}^0(\delta, b)$ *is strictly decreasing in* b *and strictly increasing in* δ*, with* $\lim_{\delta\to 0} \bar{m}^0(\delta, b) = 0$ *and* $\lim_{\delta\to 1} \bar{m}^0(\delta, \frac{1}{3}) \approx .320$*.*

The function $\underline{m}^0(\delta, b)$ *solves* $\frac{m}{\sigma_b} + \delta R(\frac{m}{\sigma_b}) = 0$*, and for all* $\delta \in (0, 1)$ *and for all* $b \in [\frac{1}{3}, \infty)$*,* $\underline{m}^0(\delta, b)$ *is strictly decreasing in* b *and in* δ*, with* $\lim_{\delta\to 0} \underline{m}^0(\delta, b) = 0$ *and* $\lim_{\delta\to 1} \underline{m}^0(\delta, b) = -\infty$*.*

These findings are represented in Figure 5. The final two sentences of Proposition 3 are proved in the Appendix.

Figure 5. The limiting case of $s \to 0$
(plotted for $b = 0$)



Comparing Propositions 2 and 3 shows that for any discount factor $\delta \in (0, 1)$ and for any $b \geq \frac{1}{3}$, the critical value of the acceptance hurdle m above which acceptance decisions are self-reinforcing is *smaller* when $s \to 0$ than when $s \to \infty$. And it is not hard to show that, fixing $b = 1$ and for any δ, the range of m values where acceptances are self-reinforcing in the model of Section 3 (where $s = 1$, $b = 1$) is strictly larger than for $s \to \infty$ and strictly smaller than for $s \to 0$. Thus, in environments where s, the increment in precision from each successive acceptance, is large, acceptances are less likely to be self-reinforcing: the critical m above which the stock effect will outweigh the precision effect will be larger. This result can be understood by recalling that

$$A_k = \frac{m + V(b) - V(k+s)}{\sigma_k},$$

where the numerator of A_k is $\bar{\mu}_k$, the optimal cutoff for the posterior evaluation of the new opportunity. While a change in s will of course indirectly affect the value of the value function $V(k)$ in all states k, the first-order effect of a change in s on the variation of A_k with k is to reduce the rate at which $\bar{\mu}_k$ falls with k, that is, to reduce the magnitude of the stock effect. (Recall that for $s \to \infty$, $V(k+s)$ is independent of k, and hence the stock effect is completely absent.) This weakening of the stock effect explains why the range of values of m where acceptances are self-reinforcing shrinks as the parameter s grows large.

Propositions 2 and 3 show that, both when $s \to 0$ and when $s \to \infty$, for any $\delta \in (0, 1)$, the critical value of m above which acceptances are self-reinforcing is smaller the larger is b, the value to which the stock of precision falls after a rejection. To understand why acceptances are more likely to be self-reinforcing in environments where b is large, observe that changes in b, unlike changes in s, do not have a direct effect on the rate at which $\bar{\mu}_k = m + V(b) - V(k+s)$ falls with k. The first-order effect of an increase in b is the rise it causes in $\bar{\mu}_k$, via the increase in $V(b)$. Recalling the discussion following Proposition 1, this rise in the cutoff has the same qualitative effect on the variation of $p(k)$ with k as an increase in m. Hence, an increase in b will increase the range of m values in which acceptances are self-reinforcing.

Propositions 2 and 3 also reveal how changes in s and b affect the likelihood that acceptances are strongly self-reinforcing, that is, that the precision effect *reinforces* the stock effect. Since the critical locus for $s \to \infty$, $m = \delta(1 - \sigma_b)\phi(0)$, and that for $s \to 0$, $m = \bar{m}^0(\delta, b)$, both shift downwards as b increases, increases in b enlarge the region in (δ, m) space where acceptances are strongly self-reinforcing. Similarly, it can be checked that $\bar{m}^0(\delta, b) > \delta(1 - \sigma_b)\phi(0)$, so increases in s have a qualitatively similar effect to increases in b. To understand these results, recall that acceptances are strongly self-reinforcing in all states when $\bar{\mu}_k > 0$ for all k. Since $\bar{\mu}_k$ is decreasing in k, this is equivalent to $\bar{\mu}_\infty = m + V(b) - V(\infty) > 0$. It is clear that the first-order effect of an increase in b is to lower the critical m above which $\bar{\mu}_\infty > 0$. The effect of s is more subtle, since s affects $\bar{\mu}_\infty$ only indirectly, as a parameter shifting $V(b)$ and $V(\infty)$. An increase in s will raise $V(b)$ more than it raises $V(\infty)$, and hence, like b, will lower the critical m above which $\bar{\mu}_\infty > 0$. This can be seen by using (22) to compare $V(b)$ and $V(\infty)$: when the current state is $k = \infty$, neither the continuation value following an acceptance, $V(\infty)$, nor the continuation value following a rejection, $V(b)$, is directly raised by an increase in s, whereas when the current state is b, the continuation value following an acceptance, $V(b + s)$, does directly increase with a rise in s.

5 Conclusion

In this paper, we have characterized history-dependence in decision-making in several models involving sequences of decisions. In all of the models, the decision-maker must, in each period, choose whether to accept or reject a different opportunity, whose payoff is uncertain and is observed only with noise. The decisions may concern the promotion of junior employees, the adoption of proposed projects, the purchase of goods or services, or the carrying out of a particular type of activity. Even though the payoffs of the different opportunities are independently distributed, the quality of tomorrow's decision depends upon whether or not today's opportunity is accepted. Specifically, each time the organization accepts its current opportunity, it adds to the stock of precision with which future opportunities can be evaluated; each time the current opportunity is rejected, this stock of information reverts to a baseline level. We analyzed how the optimal probability of acceptance depends on the current information stock, or equivalently, on the number of acceptances since the last rejection.

We identified two distinct and sometimes conflicting effects of information accumulation on decision-making. The stock effect reflects the fact that information is a valuable stock which grows with the time since the last rejection, thereby making the organization more reluctant to reject and dissipate this valuable stock. On the other hand, the precision effect reflects the fact that this stock of information is valuable only insofar as it gives rise to a more informed rejection decision. Specifically, the bigger is the information stock, the more precise is the noisy observation about the current opportunity, and therefore the more variable ex ante is the posterior evaluation of its profitability. We showed that over a wide range of parameter values for the decision-making environment, the stock effect is either reinforced by the precision effect or dominates the precision effect, and as a result, decisions to accept the current opportunity are self-reinforcing: the probability of acceptance is increasing in the number of acceptances since the last rejection. Acceptances are more likely to be self-reinforcing the higher is the "acceptance hurdle" (the ex ante expected cost, to flow payoff, of accepting the current opportunity), the smaller is the increase in precision with each successive retention, and the larger is the baseline level of precision to which the information stock reverts after a rejection.

This paper's identification of the stock and precision effects of information accumulation, and the techniques used to analyze their relative sizes, can be applied to study other sequential decision-making environments in which current decisions affect the quality of future decisions. We leave such further applications of our framework to future research.

Appendix

Proof of Proposition 1, part i): We proceed via a series of lemmas.

Lemma 1. *The transformation* $\mathbf{G}$ *defined by (8) is a contraction mapping and therefore has a unique fixed point.*

Proof: We demonstrate that under the standard metric (the sup-norm metric) on $\mathbf{R}^\omega$, $\mathbf{G} : B \mapsto B$ satisfies the two conditions of Blackwell's theorem, and therefore is a contraction mapping.

1. *Monotonicity*: $\forall V^1, V^2 \in B$, $V^1 \leq V^2$, $\forall k \in \mathbf{Z}^+$,

$$\begin{aligned} (\mathbf{G}V^1)_k &= \delta E \left[\max\left\{\tilde{\mu}(k) + V^1_{k+1}, m + V^1_1\right\}\right] \\ &\leq \delta E \max\left[\left\{\tilde{\mu}(k) + V^2_{k+1}, m + V^2_1\right\}\right] = (\mathbf{G}V^2)_k. \end{aligned}$$

2. *Discounting*: $\forall V \in B$, $a \geq 0$, $k \in \mathbf{Z}^+$,

$$(\mathbf{G}(V + a))_k = \delta E \left[\max\left\{\tilde{\mu}(k) + V_{k+1} + a, m + V_1 + a\right\}\right] = (\mathbf{G}V)_k + \delta a$$

Since both monotonicity and discounting are satisfied by $\mathbf{G}$, it follows that $\mathbf{G}$ is a contraction mapping and therefore has a unique fixed point. Q.E.D.

Lemma 2.

a) *For all* $m \in (-\infty, \infty)$ *and* $\delta \in (0, 1)$, $U(m, \delta)$ *and* $L(m, \delta)$ *are uniquely defined by (16) and (17).*

b) *If* $m \geq 0$, *then for all* $\delta \in (0, 1)$, $U(m, \delta) > L(m, \delta)$.

c) *For all* (m, δ), $U(m, \delta)$ *is strictly decreasing in* δ *and, if* $U(m, \delta) \geq L(m, \delta)$, *then* $L(m, \delta)$ *is strictly decreasing in* δ.

d) *For all* $m \geq 0$ *and for all* $\delta \in (0, 1)$, $U(m, \delta)$ *and* $L(m, \delta)$ *are strictly increasing in* m.

Proof:

a) Define the functions $u : \mathbf{R} \mapsto \mathbf{R}$ and $l : \mathbf{R} \mapsto \mathbf{R}$ by

$$u(x) \equiv -c\delta R(x) + \frac{m}{\sigma_1} - x, \qquad l(x) \equiv -\delta\left[R(x) - \sigma_1 R(U)\right] + m - x.$$

Since $c < 1$ and since $\phi'(x) = -x\phi(x)$ implies $R'(x) = -(1 - \Phi(x))$, it follows that

$$u'(x) = c\delta(1 - \Phi(x)) - 1 < 0, \qquad l'(x) = \delta(1 - \Phi(x)) - 1 < 0.$$

Now noting that $\lim_{x \to -\infty} \frac{R(x)}{x} = -1$, it follows that for any a, $-1 < a < 0$,

$$\lim_{x \to -\infty} (aR(x) - x) = \lim_{x \to -\infty} x \left(a\frac{R(x)}{x} - 1 \right) = \infty;$$

and therefore, since both δ and $c\delta$ are between 0 and 1, it follows that,

$$\lim_{x \to -\infty} u(x) = \infty, \qquad \lim_{x \to -\infty} l(x) = \infty.$$

Also, since $\lim_{x \to \infty} R(x) = 0$,[19]

$$\lim_{x \to \infty} u(x) = -\infty, \qquad \lim_{x \to \infty} l(x) = -\infty.$$

Thus, since both u and l are continuous and monotonically decreasing, each has a unique root for any $m \in (-\infty, \infty)$ and for any $\delta \in (0, 1)$. And by construction, the roots of u and l uniquely define $U(m, \delta)$ and $L(m, \delta)$.

b) To show that for any $m \geq 0$ and for any $\delta \in (0, 1)$, $U(m, \delta) > L(m, \delta)$, we evaluate $l(x)$ at $U(m, \delta)$ and show that $l(U(m, \delta)) < 0$:

$$\begin{aligned}
l(U(m,\delta)) &= -\delta\left[R(U(m,\delta)) - \sigma_1 R(U(m,\delta))\right] + m - U(m,\delta) \\
&= -\delta\left[(1-\sigma_1)R(U(m,\delta))\right] + m + \delta c R(U(m,\delta)) - \frac{m}{\sigma_1} \\
&= -\delta(1 - \sigma_1 - c)R(U(m,\delta)) - m\left(\frac{1-\sigma_1}{\sigma_1}\right) \\
&< 0,
\end{aligned}$$

for all $m \geq 0$ and for all $\delta \in (0, 1)$, since $1 - \sigma_1 - c > 0$, $R(U) \geq 0$, and $\sigma_1 \in (0, 1)$.

c) Implicitly differentiating the definition of $U(m, \delta)$ with respect to δ yields

$$U_\delta = -\frac{R(U)c}{1 + \delta c R'(U)} < 0,$$

[19]This follows since l'Hôpital's rule implies that $lim_{x \to \infty}(1 - \Phi(x))x = 0$.

where the inequality follows since $\delta \in (0,1)$, $c \in (0,1)$, and $-1 < R' < 0$. Implicitly differentiating the definition of $L(m,\delta)$ with respect to δ yields

$$\begin{aligned}
L_\delta(1+\delta R'(L)) &= -[R(L)-\sigma_1 R(U)] + \delta\sigma_1 R'(U)U_\delta \\
&= -[R(L)-\sigma_1 R(U)] - \frac{\delta c\sigma_1 R'(U)R(U)}{1+\delta c R'(U)} \\
&\overset{\text{sgn}}{=} -(R(L)-\sigma_1 R(U))(1+\delta c R'(U)) - \delta c\sigma_1 R'(U)R(U) \\
&= -(R(L)-\sigma_1 R(U)) - \delta c R'(U)R(L) \\
&= \frac{L-m}{\delta} - \delta c R'(U)R(L) \\
&\overset{\text{sgn}}{=} L-m-\delta^2 c R'(U)R(L) \\
&\leq L-m+\delta^2 c R(L) \\
&\leq L-m+\delta c R(L) \\
&= L-m+\delta R(L)+\delta R(L)(c-1) \\
&= \delta\sigma_1 R(U)+\delta R(L)(c-1) \\
&\leq \delta R(L)(\sigma_1+c-1) \\
&< 0,
\end{aligned}$$

where the first inequality follows from $R'(U) \geq -1$, the second from $\delta \in (0,1)$, the third from the hypothesis that $L \leq U$, and the fourth from the fact that $1-\sigma_1-c>0$. Since $(1+\delta R'(L))>0$, it follows that $L(m,\delta)$ is strictly decreasing in δ when $L \leq U$.

d) Implicitly differentiating the definition of $U(m,\delta)$ with respect to m yields

$$U_m = \frac{1}{\sigma_1(1+\delta c R'(U))} > 0.$$

Implicitly differentiating the definition of $L(m,\delta)$ with respect to m yields

$$\begin{aligned}
L_m(1+\delta R'(L)) &= 1+\delta\sigma_1 R'(U)U_m \\
&\overset{\text{sgn}}{=} 1+\delta(1+c)R'(U),
\end{aligned}$$

so L_m has the sign of $1+\delta(1+c)R'(U)$. Since $U_m>0$, $1+\delta(1+c)R'(U)$ is strictly increasing in m. For $\delta \leq \frac{1}{1+c}$, $1+\delta(1+c)R'(U)>0$ for all $m>-\infty$. For $\delta > \frac{1}{1+c}$, $1+\delta(1+c)R'(U)=0$ has a unique solution $m=\tilde{m}(\delta)$. Implicitly differentiating

$1 + \delta(1+c)R'(U(\tilde{m}(\delta), \delta)) = 0$ with respect to δ yields

$$\tilde{m}'(\delta) = -\frac{(R'(U) + \delta R''(U)U_\delta)}{\delta R''(U)U_m} > 0.$$

Now evaluate $\tilde{m}(\delta)$ at $\delta = 1$, as follows:

$$1 + (1+c)R'(U(\tilde{m}(1), 1)) = 0.$$

Since $R'(x) = -(1 - \Phi(x))$, $\Phi(U(\tilde{m}(1), 1)) = \frac{c}{1+c}$, so $U(\tilde{m}(1), 1) = (\Phi)^{-1}\left(\frac{c}{1+c}\right) = -1.11$. Therefore $\tilde{m}(1) = \sigma_1(U(\tilde{m}(1), 1) + cR(U(\tilde{m}(1), 1))) \approx -.655 < 0$. Since $\tilde{m}(1) < 0$ and $\tilde{m}'(\delta) > 0$, it follows that for all $\delta \in (0, 1)$ and for all $m \geq 0$, $L(m, \delta)$ is strictly increasing in m. Q.E.D.

Lemma 3. *The set $\mathcal{S}(m, \delta)$ is closed in $B \equiv \{V \mid V : \mathbf{Z}^+ \mapsto \mathbf{R}, \text{bounded}\}$ under the sup-norm.*

Proof: Assume there is a sequence of functions V^n in $S(m, \delta)$ such that $V^n \to V^\diamond$ as $n \to \infty$ in the sup-norm, where $V^\diamond \in B$. As $A(\cdot)$ is continuous, we know that $A(V^n) \to A(V^\diamond)$. Then clearly we have pointwise convergence $A(V^n(k)) \to A(V^\diamond(k))$. Since $V^n \in \mathcal{S}(m, \delta)$ for all n, $A(V^n(k)) \geq A(V^n(k+1))$ for all n. Taking the limit as $n \to \infty$ on both sides gives $A(V^\diamond(k)) \geq A(V^\diamond(k+1))$, so $A(V^\diamond)$ satisfies Condition i. In the same way, we can show that $A(V^\diamond)$ satisfies Conditions ii and iii. Therefore $A(V^\diamond) \in \mathcal{S}(m, \delta)$, so $\mathcal{S}(m, \delta)$ is closed under the sup-norm. Q.E.D.

Lemma 4. *For all $m \geq 0$ and for all $\delta \in (0, 1)$,*

$$\frac{R(U(m, \delta)) + \frac{m}{\sigma_1}}{R(L(m, \delta))} > \theta \equiv \frac{4}{3\sqrt{3}}.$$

Proof: Parts b) and c) of Lemma 2 imply that, for $m \geq 0$,

$$\frac{R(U(m, \delta)) + \frac{m}{\sigma_1}}{R(L(m, \delta))} > \frac{R(U(m, 0)) + \frac{m}{\sigma_1}}{R(L(m, 1))} = \frac{R(\frac{m}{\sigma_1}) + \frac{m}{\sigma_1}}{R(L(m, 1))}.$$

Now the numerator of this last expression is strictly positive and is an increasing function of m. From part d) of Lemma 2, the denominator is a decreasing function of

m. Therefore

$$\frac{R(\frac{m}{\sigma_1})+\frac{m}{\sigma_1}}{R(L(m,1))} \geq \frac{R(0)}{R(L((0,1))} = \frac{\phi(0)}{R(L(0,1))},$$

where $L(0,1)$ solves $L(0,1)+R(L(0,1)) = \sigma_1 R(U(0,1))$ and $U(0,1)$ solves $U(0,1)+\delta c R(U(0,1)) = 0$. Solving the second equation for $U(0,1)$ gives $U(0,1) = -.067$, and then using this value in the first equation to solve for $L(0,1)$ gives $L(0,1) = -.201$, from which it follows that

$$\frac{\phi(0)}{R(L(0,1))} = .786 > \theta. \qquad \text{Q.E.D.}$$

Lemma 5. *Define* $f : \mathbf{R}^2_+ \mapsto \mathbf{R}$ *by*

$$f(k,x) \equiv \left[\frac{1}{\sigma_k}(\sigma_{k+1} - \sigma_1 x)\right].$$

Then for all $x > \theta$ *and for all* $k \geq 1$, $f_k(k,x) > 0$.

Proof: Since

$$\frac{d\sigma_k}{dk} = \frac{1}{2}\left(\frac{1}{\sigma_k (k+1)^2}\right),$$

it follows that

$$f_k(k,x) = \left(\frac{-1}{\sigma_k^2}\right)\left(\frac{1}{2\sigma_k(k+1)^2}\right)(\sigma_{k+1} - \sigma_1 x) + \frac{1}{2\sigma_k \sigma_{k+1}(k+2)^2}.$$

With some straightforward (but tedious) manipulations, it then follows that $f_k(k,x) > 0$ iff

$$x > \frac{2\sqrt{2}}{k+2}\sigma_{k+1}. \tag{35}$$

Letting $g(k)$ represent the right-hand-side of this expression, further manipulations yield

$$g'(k) = -\frac{\sqrt{2}}{(k+2)^3 \sigma_{k+1}}(2k+1),$$

which is clearly less than 0 for all $k \geq 1$. It therefore follows that inequality (35) will be satisfied for all $k \geq 1$ provided that it is satisfied at $k = 1$. At $k = 1$ inequality (35) simply becomes

$$x > \frac{4}{3\sqrt{3}} \equiv \theta,$$

and this completes the proof. Q.E.D.

Lemma 6. *For all $m \geq 0$ and for all $\delta \in (0,1)$, if $V \in \mathcal{S}(m,\delta)$, then $A(\mathbf{G}(V)) = \mathbf{H}(A(V))$ satisfies Condition i.*

Proof: Using the expression for H we derived in (13), we have that $\forall k \in \mathbf{Z}^+$, $(\mathbf{H}A)_k \geq (\mathbf{H}A)_{k+1}$ if and only if

$$\frac{\delta}{\sigma_k}\left(\sigma_{k+1} - \sigma_1\frac{R(A_1)}{R(A_{k+1})}\right) - \frac{m}{\sigma_k R(A_{k+1})} \leq \frac{\delta}{\sigma_{k+1}}\left(\sigma_{k+2}\frac{R(A_{k+2})}{R(A_{k+1})} - \sigma_1\frac{R(A_1)}{R(A_{k+1})}\right) - \frac{m}{\sigma_{k+1}R(A_{k+1})}. \tag{36}$$

Since by hypothesis $A_{k+1} \geq A_{k+2}$, it follows that $\frac{R(A_{k+2})}{R(A_{k+1})} \geq 1$. Consequently, a sufficient condition for $(\mathbf{H}A)_k \geq (\mathbf{H}A)_{k+1}$, for $\delta \in (0,1)$, is that

$$\frac{1}{\sigma_k}\left[\sigma_{k+1} - \left(\sigma_1\frac{R(A_1)}{R(A_{k+1})} + \frac{m}{\delta R(A_{k+1})}\right)\right] \leq \frac{1}{\sigma_{k+1}}\left[\sigma_{k+2} - \left(\sigma_1\frac{R(A_1)}{R(A_{k+1})} + \frac{m}{\delta R(A_{k+1})}\right)\right]. \tag{37}$$

Since by hypothesis $L(m,\delta) \leq A_k \leq U(m,\delta)$, $m \geq 0$, and $\delta \in (0,1)$, it follows that for all $k \in \mathbf{Z}^+$,

$$\frac{R(A_1)}{R(A_{k+1})} + \frac{m}{\delta\sigma_1 R(A_{k+1})} \geq \frac{R(U(m,\delta)) + \frac{m}{\sigma_1}}{R(L(m,\delta))}$$

Lemma 4 then implies that $\forall k \in \mathbf{Z}^+$,

$$\frac{R(A_1)}{R(A_{k+1})} + \frac{m}{\delta\sigma_1 R(A_{k+1})} > \theta.$$

This observation, together with Lemma 5, demonstrates that equation (37) is satisfied $\forall\, k \in \mathbf{Z}^+$, thereby completing the proof. Q.E.D.

Lemma 7. *For all $m \in (-\infty,\infty)$ and for all $\delta \in (0,1)$, if $V \in \mathcal{S}(m,\delta)$, then $A(\mathbf{G}(V)) = \mathbf{H}(A(V))$ satisfies Condition ii.*

Proof: From Lemma 6, it follows that $\forall\, k \in \mathbf{Z}^+$, $\mathbf{H}A_1 \geq \mathbf{H}A_k$. Thus it is sufficient to show that $\mathbf{H}A_1 \leq U(m,\delta)$. Now

$$\mathbf{H}A_1 = -\frac{\delta}{\sigma_1}\left[\sigma_2 R(A_2) - \sigma_1 R(A_1)\right] + \frac{m}{\sigma_1};$$

and since $R'(\cdot) \leq 0$, it follows that

$$\mathbf{H}A_1 \leq -\delta\frac{(\sigma_2 - \sigma_1)}{\sigma_1}R(A_1) + \frac{m}{\sigma_1} \leq -\delta c R(U(m,\delta)) + \frac{m}{\sigma_1} = U(m,\delta),$$

where the second inequality uses the hypothesis that $A_1 \leq U(m,\delta)$. Q.E.D.

Lemma 8. *For all $m \in (-\infty,\infty)$ and for all $\delta \in (0,1)$, if $V \in \mathcal{S}(m,\delta)$, then $A(\mathbf{G}(V)) = \mathbf{H}(A(V))$ satisfies Condition iii.*

Proof: For $V \in \mathcal{S}(m,\delta)$, since $A(V)$ is bounded and monotonic, it follows that $\lim_{k\to\infty} A_k$ exists. Call this limit A_∞. Also, from Lemma 6, it follows that $\forall\, k \in \mathbf{Z}^+$, $\mathbf{H}A_k \geq \mathbf{H}A_\infty$. Thus it is sufficient to show that $\mathbf{H}A_\infty \geq L(m,\delta)$. Now, since $\mathbf{H}$ is a continuous operator and R is a continuous function, it follows that

$$\begin{aligned}
\mathbf{H}A_\infty &= lim_{k\to\infty}\mathbf{H}A_k = \lim_{k\to\infty} -\left[\frac{\delta}{\sigma_k}\left[\sigma_{k+1}R(A_{k+1}) - \sigma_1 R(A_1)\right]\right] + \frac{m}{\sigma_k} \\
&= -\delta\left[R(A_\infty) - \sigma_1 R(A_1)\right] + m \geq -\delta\left[R(L(m,\delta)) - \sigma_1 R(U(m,\delta))\right] + m = L(m,\delta),
\end{aligned}$$

where the inequality uses the hypotheses that $A_\infty \geq L(m,\delta)$ and $A_1 \leq U(m,\delta)$. Q.E.D.

Lemmas 6, 7, and 8 together show that for all $m \geq 0$ and for all $\delta \in (0,1)$, $\mathbf{G}(\mathcal{S}(m,\delta)) \subset \mathcal{S}(m,\delta)$. From this, together with Lemma 3, it follows that the unique fixed point V^* of $\mathbf{G}$ must be in $\mathcal{S}(m,\delta)$ and hence $A(V^*(k+1)) \leq A(V^*(k))$ for all $k \in \mathbf{Z}^+$, which is equivalent to $p(k+1) \geq p(k)$ for all $k \in \mathbf{Z}^+$. Strict monotonicity of $p(k)$ follows by noting from Lemmas 4 and 5 that since

$$\frac{R(U(m,\delta)) + \frac{m}{\sigma_1}}{R(L(m,\delta))} > \theta$$

holds strictly, inequality (37) in the proof of Lemma 6 is strict as well. This implies that $A(V(k+1)) = A(V(k))$ for all $k \in \mathbf{Z}^+$ is incompatible with V being a fixed point of $\mathbf{G}$. This completes the proof of part i) of Proposition 1.

Proof of Proposition 1, part ii): If, for $m \geq 0$ and for any $\delta \in (0,1)$, $L(m,\delta) > 0$, then the same arguments used to prove part i) establish that for the fixed point V^*, not only are Conditions i, ii, and iii satisfied but also $A(V^*(k)) > 0$ for all $k \in \mathbf{Z}^+$, which is equivalent to $p(k) < \frac{1}{2}$. The final steps are accomplished by:

Lemma 9. *In the region $m \geq 0$, the condition $L(m,\delta) = 0$ implicitly determines a function $m = \bar{m}(\delta)$ satisfying $\bar{m}(\delta) > 0$ and $\bar{m}'(\delta) > 0$ for $\delta \in (0,1)$, $\lim_{\delta\to 0}\bar{m}(\delta) = 0$, and $\lim_{\delta\to 1}\bar{m}(\delta) \approx .188$. For $m > \bar{m}(\delta)$, $L(m,\delta) > 0$.*

Proof: For $m = 0$, the condition defining $U(m, \delta)$ becomes $U(0, \delta) + \delta c R(U(0, \delta)) = 0$. Therefore $U(0, \delta) < 0$ and so, by part b) of Lemma 2, $L(0, \delta) < 0$. By part d) of Lemma 2, $L(m, \delta)$ is strictly increasing in m for $m \geq 0$. The condition defining $L(m, \delta)$ is

$$L(m, \delta) + \delta R(L(m, \delta)) = \delta \sigma_1 R(U(m, \delta)) + m,$$

and as $m \to \infty$, the right-hand side of this condition goes to ∞. Since the left-hand side is increasing in L and goes to ∞ as $L \to \infty$, $\lim_{m \to \infty} L(m, \delta) = \infty$. Therefore, for all $\delta \in (0, 1)$, in the interval $m \in (0, \infty)$ there exists a unique $m \equiv \bar{m}(\delta)$ such that $L(m, \delta) = 0$, and for $m > \bar{m}(\delta)$, $L(m, \delta) > 0$.

The condition $L(\bar{m}(\delta), \delta) = 0$ is equivalent to

$$\delta \phi(0) = \delta \sigma_1 R(U(\bar{m}(\delta), \delta)) + \bar{m}(\delta), \tag{38}$$

and implicitly differentiating with respect to δ yields

$$\bar{m}'(\delta)(1 + \delta \sigma_1 R'(U) U_m) = \phi(0) - \sigma_1 R(U) - \delta \sigma_1 R'(U) U_\delta.$$

The proof of part d) of Lemma 2 showed that $1 + \delta \sigma_1 R'(U) U_m > 0$ for $m \geq 0$. Therefore

$$\begin{aligned}
\bar{m}'(\delta) &\overset{\text{sgn}}{=} \phi(0) - \sigma_1 R(U) - \delta \sigma_1 R'(U) U_\delta \\
&= \frac{m}{\delta} + \frac{\delta c \sigma_1 R(U) R'(U)}{1 + \delta c R'(U)} \\
&= \frac{m}{\delta} + \frac{(m - \sigma_1 U) R'(U)}{1 + \delta c R'(U)} \\
&\overset{\text{sgn}}{=} \frac{m}{\delta}(1 + \delta(1 + c) R'(U)) - \sigma_1 U R'(U).
\end{aligned}$$

Now for $m \geq 0$ the first term is positive (as shown in the proof of part d) of Lemma 2) and, by part b) of Lemma 2, $U(\bar{m}(\delta), \delta) > L(\bar{m}(\delta), \delta) = 0$, so $-\sigma_1 U R'(U) > 0$. Therefore $\bar{m}'(\delta) > 0$. From (38), $\lim_{\delta \to 0} \bar{m}(\delta) = 0$. As $\delta \to 1$, (38) becomes $\phi(0) = \sigma_1 R(U(\bar{m}(1), 1)) + \bar{m}(1)$, and $U(\bar{m}(1), 1)$ solves $U(\bar{m}(1), 1) + c R(U(\bar{m}(1), 1)) = \frac{\bar{m}(1)}{\sigma_1}$. Solving these two conditions simultaneously for $U(\bar{m}(1), 1)$ and $\bar{m}(1)$ yields $\bar{m}(1) \approx .188$. Q.E.D.

This completes the proof of part ii) of Proposition 1.

Proof of Proposition 1, part iii):

Lemma 10. *If for the fixed point V^* of $\mathbf{G}$, $A(V^*(k)) \geq A(V^*(k+1))$ for all $k \in \mathbf{Z}^+$, then the bounds $L(m,\delta)$ and $U(m,\delta)$ defined in (17) and (16) must satisfy $L(m,\delta) \leq U(m,\delta)$. Equivalently, if $L(m,\delta) > U(m,\delta)$, then for the fixed point V^* of $\mathbf{G}$ there must exist some $k \in \mathbf{Z}^+$ such that $A(V^*(k)) < A(V^*(k+1))$ and hence $p(k) > p(k+1)$.*

Proof: From the definition of the operator $\mathbf{H}$ in (9), if V^* is a fixed point of $\mathbf{G}$, then

$$\mathbf{H}(A(V^*)) = A(\mathbf{G}(V^*)) = A(V^*),$$

so $A(V^*)$ is a fixed point of $\mathbf{H}$, and hence from equation (13),

$$A(V^*)_k = -\frac{\delta}{\sigma_k}\left[\sigma_{k+1}R(A(V^*)_{k+1}) - \sigma_1 R(A(V^*)_1)\right] + \frac{m}{\sigma_k}. \tag{39}$$

Write A_k for $A(V^*)_k$, set $k=1$, and use the hypothesis that $A_k \geq A_{k+1}$ for all $k \in \mathbf{Z}^+$ to derive

$$A_1 = -\frac{\delta}{\sigma_1}\left[\sigma_2 R(A_2) - \sigma_1 R(A_1)\right] + \frac{m}{\sigma_1} \leq -\delta c R(A_1) + \frac{m}{\sigma_1}.$$

Since $\delta c < 1$, $A_1 + \delta c R(A_1)$ is strictly increasing in A_1. Therefore it follows from the definition of $U(m,\delta)$ that

$$A_1 \leq U(m,\delta). \tag{40}$$

Taking the limit as $k \to \infty$ in (39) and using (40) gives

$$A_\infty = -\delta[R(A_\infty - \sigma_1 R(A_1)] + m \geq -\delta[R(A_\infty) - \sigma_1 R(U(m,\delta))] + m.$$

Since $A_\infty + \delta R(A_\infty)$ is strictly increasing in A_∞, it follows from the definition of $L(m,\delta)$ that

$$L(m,\delta) \leq A_\infty. \tag{41}$$

The hypothesis that $A_{k+1} \leq A_k$ for all $k \in \mathbf{Z}^+$, along with (40) and (41), implies that

$$L(m,\delta) \leq A_\infty \leq A_1 \leq U(m,\delta). \qquad \text{Q.E.D.}$$

Lemma 11. *The condition $L(m,\delta) = U(m,\delta)$ implicitly determines a function $m = \underline{m}(\delta)$ satisfying $\underline{m}(\delta) < 0$ and $\underline{m}'(\delta) < 0$ for $\delta \in (0,1)$, $\lim_{\delta\to 0} \underline{m}(\delta) = 0$, and $\lim_{\delta\to 1} \underline{m}(\delta) \approx -.208$. For $m < \underline{m}(\delta)$, $L(m,\delta) > U(m,\delta)$.*

Proof: From the proof of part b) of Lemma 2, we know that $L < U$, $L > U$, or $L = U$ according to whether

$$\delta R(U(m,\delta))(1-\sigma_1-c) + m\frac{(1-\sigma_1)}{\sigma_1} \tag{42}$$

is strictly positive, strictly negative, or zero. At $m = 0$, expression (42) is strictly positive. For fixed δ, the derivative of (42) with respect to m is

$$\begin{aligned}\delta(1-\sigma_1-c)R'(U)U_m &+ \frac{1-\sigma_1}{\sigma_1} \\ &= \frac{(1-\sigma_1)(1+\delta R'(U)) - \sigma_1\delta c R'(U)}{\sigma_1(1+\delta c R'(U))} > 0,\end{aligned}$$

and as $m \to -\infty$, this derivative approaches

$$\frac{(1-\sigma_1)(1-\delta)+\sigma_1\delta c}{\sigma_1(1-\delta c)} > 0,$$

so as $m \to -\infty$, expression (42) approaches $-\infty$. Therefore for all $\delta \in (0,1)$, there exists a unique $m \equiv \underline{\mathrm{m}}(\delta) < 0$ such that $L(m,\delta) = U(m,\delta)$, and for $m < \underline{\mathrm{m}}(\delta)$, $L(m,\delta) > U(m,\delta)$. Setting expression (42) equal to 0 and implicitly differentiating with respect to δ gives

$$\begin{aligned}\underline{\mathrm{m}}'(\delta)[\delta(1-\sigma_1-c)R'(U)U_m &+ \frac{1-\sigma_1}{\sigma_1}] \\ &= -(1-\sigma_1-c)[R(U)+\delta R'(U)U_\delta],\end{aligned}$$

from which it follows that $\underline{\mathrm{m}}'(\delta) < 0$. Setting $\delta = 0$ in (42) shows that $\lim_{\delta\to 0}\underline{\mathrm{m}}(\delta) = 0$. Finally, taking $\delta \to 1$ in (42) and setting the expression equal to 0 gives

$$R(U(\underline{\mathrm{m}}(1),1))(1-\sigma_1-c) + \underline{\mathrm{m}}(1)\frac{(1-\sigma_1)}{\sigma_1} = 0,$$

where $U(\underline{\mathrm{m}}(1),1)$ solves $U(\underline{\mathrm{m}}(1),1)+cR(U(\underline{\mathrm{m}}(1),1)) = \frac{m}{\sigma_1}$. Solving these two equations simultaneously for $U(\underline{\mathrm{m}}(1),1)$ and $\underline{\mathrm{m}}(1)$ yields $\underline{\mathrm{m}}(1) \approx -.208$. Q.E.D.

Proof of Proposition 3: Parts i), ii), and iii) are proved in the text, so here we prove only the claims about the properties of $\bar{m}^0(\delta,b)$ and $\underline{\mathrm{m}}^0(\delta,b)$.

The function $\bar{m}^0(\delta, b)$ solves

$$m + \delta\sigma_b R(\frac{m}{\sigma_b}) = \delta\phi(0). \tag{43}$$

For all $\delta \in (0,1)$ and for all $b \in [\frac{1}{3}, \infty)$, the left-hand side of (43) is strictly increasing in m and, evaluated at $m = 0$, gives $\delta\sigma_b\phi(0)$, which is strictly less than the right-hand side of (43), so $\bar{m}^0(\delta, b) > 0$. Differentiating (43) implicitly with respect to σ_b yields

$$\frac{\partial m}{\partial \sigma_b}\left(1 + \delta R'(\frac{m}{\sigma_b})\right) = -\delta R(\frac{m}{\sigma_b}) + \frac{\delta m}{\sigma_b} R'(\frac{m}{\sigma_b}) < 0,$$

where the inequality uses the fact that $\bar{m}^0(\delta, b) > 0$. Since $1 + \delta R'(\frac{m}{\sigma_b}) > 0$, $\frac{\partial m}{\partial \sigma_b} < 0$, and since σ_b is strictly increasing in b, $\bar{m}^0(\delta, b)$ is strictly decreasing in b. Differentiating (43) implicitly with respect to δ yields

$$\frac{\partial m}{\partial \delta}\left(1 + \delta R'(\frac{m}{\sigma_b})\right) = \phi(0) - \sigma_b R(\frac{m}{\sigma_b}) > 0,$$

where once again the inequality uses the fact that $\bar{m}^0(\delta, b) > 0$. Hence $\bar{m}^0(\delta, b)$ is strictly increasing in δ. From (43) it is obvious that $\lim_{\delta \to 0} \bar{m}^0(\delta, b) = 0$. Substituting $\delta = 1$ and $b = \frac{1}{3}$ into (43) yields $2m + R(2m) = 2\phi(0)$, which when solved for m gives $m \approx .320$.

The function $\underline{\mathrm{m}}^0(\delta, b)$ solves $\frac{m}{\sigma_b} + \delta R(\frac{m}{\sigma_b}) = 0$, so it follows immediately from the definition of $R(\cdot)$ that for $\delta \in (0,1)$ and $b \in [\frac{1}{3}, \infty)$, $\underline{\mathrm{m}}^0(\delta, b) < 0$. It is straightforward to show, by implicit differentiation, that $\underline{\mathrm{m}}^0(\delta, b)$ is strictly decreasing in b and in δ, and the limiting cases as δ goes to 0 and 1 are similarly straightforward, given the definition of $R(\cdot)$. Q.E.D.

References

Aghion, P., Bolton, P., Harris, C., and Jullien, B., "Optimal Learning by Experimentation," *Review of Economic Studies* 58, 1991, 621-655.

Ali, S., and Kartik, N., "A Theory of Momentum in Sequential Voting," Working Paper, Stanford Univ. and U.C. San Diego, April 2006, http://ssrn.com/abstract=902697.

Banerjee, A., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics* 107, 1992, pp. 797-818.

Banks, J., and Sundaram, R., "Denumerable-Armed Bandits," *Econometrica* 60, 1992, 1071-1096.

Banks, J., and Sundaram, R., "Optimal Retention in Agency Problems," *Journal of Economic Theory* 82, 1998, pp. 293-323.

Bergemann, D., and Valimaki, J., "Learning and Strategic Pricing," *Econometrica* 64, 1996, pp. 1125-49.

Bergemann, D., and Valimaki, J., 22 "Bandit Problems," Cowles Foundation D.P. No. 1551, Yale University, Jan. 2006.

Bernhardt, D., "Strategic Promotion and Compensation," *Review of Economic Studies* 62, 1995, pp. 315-339.

Berry, D.A. and Fristedt, B. *Bandit Problems: Sequential Allocation of Experiments*, London: Chapman Hall, 1985.

Bikhchandani, S., Hirshleifer, D., and Welch, I., "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy* 100, 1992, pp. 992-1026.

Bolton, P., and Harris, C., "Strategic Experimentation," *Econometrica* 67, 1999, pp. 349-374.

Cornell, B., and Welch, I., " Culture, Information, and Screening Discrimination," *Journal of Political Economy* 104, 1996, pp. 542-71.

DeGroot, M., *Optimal Statistical Decisions*, New York: McGraw-Hill, 1970.

Easley, D., and Kiefer, N., "Controlling a Stochastic Process with Unknown Parameters," *Econometrica* 56, 1988, pp. 1045-1064.

Felli, L., and Harris, C. J., "Learning, Wage Dynamics, and Firm-Specific Human Capital," *Journal of Political Economy* 104, 1996, pp. 838-68.

Festinger, L., *A Theory of Cognitive Dissonance*, Evanston, IL: Row Peterson, 1957.

Gittins, J., and Jones, D., "A Dynamic Allocation Index for the Sequential Allocation of Experiments," in Gani, J., ed. *Progress in Statistics*, Amsterdam: North Holland, 1974, pp. 241-266.

Hirshleifer, D., and Welch, I., "Institutional Memory, Inertia, and Impulsiveness," Anderson Graduate School of Management Working Paper, UCLA, September 1994.

Hirshleifer, D., and Welch, I., "An Economic Approach to the Psychology of Change: Amnesia, Inertia, and Impulsiveness," *Journal of Economics and Management Strategy* 11, 2002, pp. 379-421.

Jovanovic, B., "Job Matching and the Theory of Turnover," *Journal of Political Economy* 87, 1979, pp. 972-990.

Keller, G., Rady, S., and Cripps, M., "Strategic Experimentation with Exponential Bandits," *Econometrica* 73, 2005, pp. 39-68.

Kuran, T., " Preference Falsification, Policy Continuity and Collective Conservatism," *Economic Journal* 97 (1987), pp. 642-665.

Marinescu, I. "Shortening the Tenure Clock: The Impact of Strengthened U.K. Job Security Legislation", mimeo, L.S.E., Jan. 2006.

Meyer, M. A., "Learning from Coarse Information: Biased Contests and Career Profiles," *Review of Economic Studies* 58, 1991, pp. 15-41.

Meyer, M. A., "Biased Contests and Moral Hazard: Implications for Career Profiles," *Annales d'Economie et de Statistique* 25/26 1992, pp. 165-187.

Prendergast, C., " Career Development and Specific Human Capital Collection," *Journal of the Japanese and International Economies* 6 (1992), pp. 207-227.

Rabin, M., and Schrag, J., " First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics* 114 (1999), pp. 37-82.

Rothschild, M., "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory* 9, 1974, pp. 185-202.

Rustichini, A., and Wolinsky, A., " Learning about Variable Demand in the Long Run," *Journal of Economic Dynamics and Control* 19, 1995, pp. 1283-1292.

Scharfstein, D. and Stein, J., "Herd Behavior and Investment," *American Economic Review* 80, 1990, pp. 465-479.

Schlag, K., "Why Imitate, and If So, How?", *Journal of Economic Theory* 78, 1998, pp. 130-156.

Smith, L. and Sorensen, P.N., "Pathological Outcomes of Observational Learning," *Econometrica* 68, 2000, pp. 371-398.

Smith, L. and Sorensen, P.N., "Informational Herding and Optimal Experimentation," Cowles Foundation D.P. No. 1552, Yale University, Jan. 2006.

Sobel, J., "A Model of Declining Standards," *International Economic Review*, 41, 2000, 295-303.

Sobel, J., "On the Dynamics of Standards," Rand Journal of Economics, 32, 2001, 606-623.

Zwiebel, J., "Corporate Conservatism and Relative Compensation," *Journal of Political Economy* 103, 1995, pp. 1-26.