Managerial Conservatism and Rational Information Acquisition

Eric Rasmusen

Published: *Journal of Economics and Management Strategy* (Spring 1992), 1: 175-202.

Abstract

Conservative managerial behavior can be rational and profit- maximizing. If the valuation of innovations contains white noise and the status quo would be preferred to random innovation, then any innovation that does not appear to be substantially better than the status quo should be rejected. The more successful the firm, the higher the threshold for accepting innovation should be, and the greater the conservative bias. Other things equal, more successful firms will spend less on research, adopt fewer innovations, and be less likely to advance the industry's best practice.

Draft: 9.1 (Draft 1.1, October 1989.)

Eric Rasmusen, Kelley School of Business, BU 456, 1309 E 10th Street, Bloomington, Indiana, 47405-1701. Office: (812) 855-9219. Fax: 812-855-3354. Erasmuse@indiana.edu.

I would like to thank David Hirshleifer, Steven Lippman, George Lowenstein, Edward Miller, Emmanuel Petrakis, Ivan Png, Steve Postrel, Ivo Welch, two anonymous referees, and participants in the Chicago GSB Uncertainty Workshop for helpful comments, and Michael Kim and George Michaelides for research assistance.

Introduction

American managers, it is commonly said, are conservative and short-sighted, passing up new ideas and avoiding risks. Hayes & Garvin (1982) reflect a common opinion when they complain in the *Harvard Business Review* of excessive managerial conservatism in the form of high hurdle rates:

> Such hurdle rates often bear little resemblance either to a company's real cost of capital (even after appropriate adjustment for differences in risk) or to the actual rates of return (net of deterioration replenishment) that the company can reasonably expect to earn from alternative investments. Again and again, we have observed the use of pretax hurdle rates of 30% or more in companies whose actual pretax returns on investment were less than 20%. (Hayes & Garvin, p. 76)

One should keep in mind that managerial conservatism and myopia may just be popular myth. A recent Wall Street Journal/NBC poll found that 51% of Americans thought that Japan had more of a long-term perspective, but only 10% of the Japanese did, suggesting that people everywhere like to decry their compatriots' shortsightedness.[1] Some evidence, however, does seem to support the idea. Ross (1986) interviewed twelve manufacturers about their decisionmaking process regarding investments in energy conservation and examined their internal records. He concluded that the hurdle rate for large projects was near the companies' cost of capital, but for small projects the hurdle rate was higher, and "Decisions are then based on the primary quantitative measure from the analysis supplemented by informal adjustments made in the minds of decisionmakers." Pruitt & Gitman (1987) sent questionnaires on project evaluation to financial officers in Fortune 500 firms, and found they believed that forecasts were consistently overoptimistic, implying a need for conservative use of forecasts. 80% of respondents felt that revenue forecasts of capital-budgeting proposals were overstated. 37% of them thought that this was intentional and 36% attributed it to inexperience. 59% of them agreed, and 20% disagreed with the statement: "In general, decisionmakers who evaluate forecasts consider them to be optimistic in their estimates and adjust forecasts to correct them."

The perception is common enough that economists have devoted considerable effort trying to explain it. Recent articles include Narayanan (1985), Stein (1988), Shleifer & Vishny (1989), and Hirshleifer & Thakor (1991). A variety of explanations for conservative bias have been developed based on incomplete rationality or rational but not profit-maximizing behavior. People, including managers, might simply make systematic mistakes, as discussed in Libby & Fishburn (1977). Or, bounded rationality might generate conservative behavior rules, as in Day (1987), Heiner (1983), and Kuran (1988). Even if managers are conventionally rational, principal-agent problems can generate conservative bias in a variety of ways. Most simply, agents may overestimate the value of new projects because they are rewarded more highly when projects are adopted. The principal then imposes a conservative bias to undo the overestimate. Agency problems can also have more complicated effects, of which I will describe just two to give the flavor of them. In Holmstrom & Ricart (1986), the agent initially is ignorant of his ability, but he and the rest of the world can discover it by undertaking projects. The agent can either ignore a potential project or truthfully

[1] "Which Land is Fairest of Them All? Japanese, Americans Talk to Pollsters," *Wall Street Journal*, 15 June 1990, p. A6.

report a signal of its quality to the principal to obtain approval for undertaking it. If the agent is simply paid his estimated marginal product, he will veto too many projects, because he is risk-averse and fears discovering that he has low ability. But if he is given a downward-rigid wage, which is part of the second-best optimal contract, then he overinvests unless the principal engages in conservative capital rationing. In Lambert (1986), a risk-averse agent must be motivated to expend effort in measuring the value of an innovation and to adopt it if it is superior to the status quo. The principal does not observe effort or the agent's report, only the cash flow from the decision. Lambert finds a conservative bias if the average innovation is superior to the status quo, because the agent must be given sufficient incentive to expend effort.

It is possible, however, to explain conservatism without either irrationality or agency problems. Most simply, those who see conservatism might be ignoring adjustment costs, especially the cost of management time. This explanation would not explain, however, why financial officers think that project revenue forecasts are consistent overestimates. The present article explores a different argument, valid even in the absence of adjustment costs, which suggests that managers process information rationally, but in a way that seems overcautious to outsiders, and that the borderline refused project, which seems clearly profitable, in fact would yield zero profits. The argument will be based on regression to the mean: if profitability estimates are made with error, even if the error is unbiased, then a project measured to be unusually profitable is probably not as profitable as it appears. Regression to the mean is symmetric with respect to unusually unprofitable projects, but such projects would be rejected anyway. Put somewhat differently, the distribution of profitabilities can be treated as a Bayesian prior, and the measurement of a particular new project's profitability is one data point of information. Under conditions specified below, the mean of the posterior distribution will lie between the prior mean and the information, and if the information is greater than the prior mean, the posterior mean is less than the information, giving rise to conservatism. The model follows a line of inquiry begun in the corporate finance literature, where Brown (1974, 1978) and Smidt (1979) noted that measurements of adopted projects will appear overoptimistic because projects with low measured values will not be adopted. Miller (1978, 1987) expanded the argument using examples which suggest that managers should apply a conservative bias in capital budgeting. The present article will focus on the assumptions behind the argument, and the implications for which firms will collect information and innovate.

Section 1 will lay out the model and establish that a conservative bias is rational (Proposition 1). It will also give an example with normal distributions to show the magnitude of the bias, and compare its size with the effect of risk aversion. Section 2 will show that as firms progress they should become increasingly conservative (Proposition 2), and that even with conservative decisionmaking most projects will prove disappointing (Proposition 3). Section 3 will discuss empirical implications, and Section 4 will suggest applications to a variety of other contexts. Section 5 concludes.

1. The Model

A risk-neutral manager whom we shall call "the boss" chooses a single "policy" for each of two periods to maximize his firm's profits. A policy is a method of using the resources available to the firm, and it may be an investment project, a technological innovation, or an organizational form. Many policies are possible, but the per-period profitability of policy i, θ_i, is unknown to the boss. He does know that θ_i is drawn from a continuous distribution with a symmetric density $f_\theta(\theta)$ which has mean $\overline{\theta} > 0$ and is either unimodal or uniform. Once a policy is adopted, θ_i becomes known, but the policy cannot be reversed until the following period.[2] At the end of the first period, the policy can be changed. The policy chosen for the first period will be called the "status quo," with profitability denoted θ_0, and our focus will be on the boss's incentive to innovate in the second period.

In each period, the boss may commission n staffers to estimate the profitability of n different policies, at a cost of c per policy estimated.[3] Estimates from previous periods grow stale and are no longer available.[4] Staffers do not know f_θ, and staffer i reports his measurement of θ_i, denoted y_i, where $y_i = \theta_i + u_i$. The variable u_i is a random error with mean zero, independent of θ_i and u_j for $j \neq i$ and distributed according to a symmetric continuous density f_u which is unimodal or uniform. Assume also, for reasons explained below, that the support of f_u is at least as wide as the support of f_θ. For convenience, let $f(x)$ denote the marginal density for any variable x; $f(x|z)$, the conditional density; $f(x,z)$, the joint density; and $E(x)$ the expected value; where f and E are derived from whatever functions are appropriate to their arguments. The following lemma will play a central role:

LEMMA 1: (i) $E(\theta|y)$ is increasing in y and (ii) $\overline{\theta} < E(\theta|y) < y$ for $y > \overline{\theta}$, $E(\theta|y) = y$ for $y = \overline{\theta}$, and $y < E(\theta|y) < \overline{\theta}$ for $y < \overline{\theta}$.

Proof: See Appendix.

Lemma 1 says that the expected value of the new policy conditional on the estimate is higher if the estimated value is higher, and that the expected value lies between the estimate and the value of the average new policy.

Initially, the firm has no status quo, and the boss must decide how to choose a policy. He could blindly accept an uninvestigated policy, which has expected profitability $\overline{\theta}$, or accept one of the investigated policies, which has measured profitability y and expected profitability $E(\theta|y)$. The optimal decision rule is to choose an uninvestigated policy if all the investigated policies have estimated profitabilities less than $\overline{\theta}$, and to choose the policy with the greatest estimated profitability

[2] Otherwise, the boss might adopt a policy simply to discover its profitability, which is not purely an adoption decision, but rather a way to acquire information. The problem of when to switch policies to acquire information is the "multi-armed bandit" problem discussed in Weitzman (1979).

[3] This is distinct from the problem analyzed by Sah & Stiglitz (1988) of combining the reports of n staffers on *one* policy.

[4] This assumption would become relevant if a firm adopted a bad policy in the first period which would, ex post, be abandoned for a different policy whose value had been estimated in the first period. Such a possibility does not affect Proposition 1, but it would require extra care in Section 2's analysis of the optimal n.

otherwise. If we define $y_m = Max\{y_i\}_{i=1}^n$, the firm adopts policy m if $y_m \geq \overline{\theta}$, and an uninvestigated policy otherwise.

Let us assume that c is low enough that it is worth investigating at least one policy:

$$E[(\theta - \overline{\theta})|(y \geq \overline{\theta})] \int_{\overline{\theta}}^{\infty} f(y)dy \geq c. \tag{1}$$

Equation (1) says that the expected gain in profitability from having the option to accept the investigated policy if its estimated value exceeds $\overline{\theta}$ equals at least c.

If the status quo were chosen blindly, it would be true that $E(\theta_0) = \overline{\theta}$. But the expected value can only increase, given the option of adopting an investigated policy. It takes the value

$$E(\theta_0) = \overline{\theta} + E[(\theta_m - \overline{\theta})|(y_m \geq \overline{\theta})] \int_{\overline{\theta}}^{\infty} f(y_m)dy_m. \tag{2}$$

The second term of equation (2) is positive because there is a positive probability that $\theta_m > \overline{\theta}$, and Lemma 1 then tells us that $E(\theta_m|y_m) > \overline{\theta}$. Thus we have

LEMMA 2: For the average new firm, the status quo is superior to blind innovation: $\theta_0 > \overline{\theta}$.

The question of what adoption rule the boss should follow regarding innovations can now be addressed. For the average firm, Lemma 2 says that the status quo is $\theta_0 > \overline{\theta}$. We will see that for this average firm, the optimal decision procedure is to adopt the new policy if y is greater than some threshold. Let $\overline{y}$ denote the adoption threshold and let y^* denote the optimal value of $\overline{y}$. In the absence of estimation error, the boss's optimal adoption rule would be $\overline{y} = \theta_0$: "Accept the new policy if and only if $y \geq \theta_0$." With estimation error, on the other hand, the optimal rule is described by Proposition 1.

PROPOSITION 1: The average firm uses a threshold rule with a conservative bias: $y^* > \theta_0$.

Proof: Lemma 2 says that for the average firm, $\theta_0 > \overline{\theta}$. Thus, the boss will reject innovations with $E(\theta|y) < \overline{\theta}$, and since Lemma 1 says that if $y > \overline{\theta}$ then $E(\theta|y)$ increases in y, a threshold rule is optimal. The boss will choose the threshold y^* so that $E(\theta|y \geq y^*) \geq \theta_0$, so the threshold is such that

$$E(\theta|y^*) = \theta_0. \tag{3}$$

Lemma 1 says that if $y > \overline{\theta}$ then $E(\theta|y) < y$. It follows that $E(\theta|y^*) < y^*$, and if equation (3) is to be satisfied, it must be that $y^* > \theta_0$. $\square$

As an example, let us assume normality of the distributions of new policies and measurement errors, f_θ and f_u. The boss is trying to use the observed variable $y = \theta + u$ to estimate the unobserved variable $\theta = \overline{\theta} + \epsilon$, where u and ϵ are independent random variables with zero mean, and $\overline{\theta}$ is a constant. If $u \sim N(0, \sigma_u^2)$ and $\epsilon \sim N(0, \sigma_\epsilon^2)$, then $\theta \sim N(\overline{\theta}, \sigma_\epsilon^2)$ and $y \sim N(\overline{\theta}, \sigma_u^2 + \sigma_\epsilon^2)$. What the boss cares about is the conditional distribution $f(\theta|y)$, which is also normal, with parameters that can be calculated (see, e.g., Casella [1985]). The mean is

$$E(\theta|y) = \left(\frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2} \right) \overline{\theta} + \left(\frac{\sigma_\epsilon^2}{\sigma_u^2 + \sigma_\epsilon^2} \right) y, \tag{4}$$

in which case $E(\theta|y) < y$ if $y > \theta_0$. The variance is

$$Var(\theta|y) = \frac{\sigma_u^2 \sigma_\epsilon^2}{\sigma_u^2 + \sigma_\epsilon^2}. \tag{5}$$

To find the optimal threshold, y^*, the boss solves for y in the equation

$$E(\theta|y) = \theta_0. \tag{6}$$

Using equation (4), equation (6) becomes

$$\left(\frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2}\right)\overline{\theta} + \left(\frac{\sigma_\epsilon^2}{\sigma_u^2 + \sigma_\epsilon^2}\right) y = \theta_0. \tag{7}$$

Solving (7) for y gives

$$y^* = \left(\frac{\sigma_u^2 + \sigma_\epsilon^2}{\sigma_\epsilon^2}\right)\theta_0 - \left(\frac{\sigma_u^2}{\sigma_\epsilon^2}\right)\overline{\theta}, \tag{8}$$

or, equivalently,

$$y^* = \theta_0 + \frac{\sigma_u^2}{\sigma_\epsilon^2}(\theta_0 - \overline{\theta}). \tag{9}$$

Equation (9) confirms Proposition 1. If a random draw is likely to be worse than the status quo because $\theta_0 > \overline{\theta}$, then equation (8) says that $y^* > \theta_0$. The conservative bias increases in the variance of the measurement error, σ_u^2, and decreases in the variance of the possible new-policy values, σ_ϵ^2.[5]

Table 1 shows different values calculated from equation (8), which may give some idea of how empirically relevant the conservative bias might be. The maintained assumptions are that $\sigma_\epsilon = 15$ and $\theta_0 = 100$, while σ_u and $\overline{\theta}$ take various values.

[5]The effect of the variance here is not general. With the normal density, mean and variance fully characterize the distribution, but what really matters is how much of the new-policy density is for policy values superior to the status quo. Probability mass between the new-policy mean and the status quo increases variance, but not the attractiveness of new policies. This is a little like the value of an option: it is not exactly variance of the underlying asset price that gives an option value; it is the probability that the asset value will be beyond the strike price.

Table 1
Acceptance Thresholds for Normal Distributions
$y^* = \theta_0 + \frac{\sigma_u^2}{\sigma_\epsilon^2}(\theta_0 - \overline{\theta})$

		Standard Error of Measurement (σ_u)			
		5	10	15	40
	50	106	122	150	456
New	80	102	109	120	242
Policy	90	101	104	**110**	171
Mean	95	100.6	102	105	136
($\overline{\theta}$)	98	100.2	101	102	114
	100.0	100.0	100.0	100.0	100.0
	110	99	96	90	29
Assumed: $\theta_0 = 100$, $\sigma_\epsilon = 15$			(fractions rounded)		

Consider, for example, the boldfaced entry in Table 1, $y^*(90, 15) = 110$. This means that if the average new policy has a value of 90, and about 1/3 of the staffer's measurements are wrong by more than 15, the new policy needs to look about 10% better than the status quo to be acceptable to the boss. Thus, the effect is sizeable for plausible parameter values.

Discussion of Proposition 1

The most intuitive interpretation of the conservative bias is as regression towards the mean. The staffer might measure the new policy's value to be high for either of two reasons: (1) the true value θ is above average, or (2) the error u is positive. Having observed a high measured value, the subjective probabilities of both (1) and (2) should rise. Ex ante, the probability of a positive error is no greater than that of a negative error. But positive errors tend to push the measured value above the average, so, ex post, positive errors are more common for above-average observations of y. In that limited range, the errors have a positive mean, not a zero mean, and the observed value overestimates the true value. Since positive errors are more likely given an observed value above average, if the staffer made a second measurement, on the same policy but with independent error, he would most likely make a less positive or a negative error, and the second measured value would be less than the first. The measurement "regresses towards the mean." Because the high first measurement might have been produced by a high true value, the expected value of the second measurement is still above $\overline{\theta}$: the measurement only regresses *towards* the mean, not all the way. Thus, for a particular y it might be true that $y > \theta_0 > E(\theta|y) > \overline{\theta}$.[6]

[6] All of this has disregarded adoption costs, another difference between a status quo and an innovation. Adoption costs very obviously skew the conclusions towards conservatism. But in fact they skew them even more than might be obvious. Suppose that $\theta_0 = 100$ and $\overline{\theta} = 100$, but the implementation cost of $c = 5$ is not included in the definition of θ. The obvious conservative bias is that the boss should refuse the new policy if $\theta = 104$. But if θ is not known, and the staffer reports that $y = 106$, so that $y - c > \theta_0$, the boss should still refuse the new policy. Because $(y - c)$ equals 101, which is uncomfortably close to 100, $(E(\theta|y) - c)$ is less than 100. Thus, adoption costs create more conservatism than one might think.

The language of Bayesian statistics provides a terser interpretation of Proposition 1. The boss has a prior distribution, $f(\theta)$, which he updates using the information y to obtain the posterior distribution $f(\theta|y)$. Given the assumptions of the model, the posterior mean lies between the prior mean and the information, and hence has a lower value than the information, which induces a conservative bias.

Two elements of the model are key to the result. The first is Lemma 2's statement that the status quo is superior to a blind draw from the pool of policies. In the present model, Lemma 2 was generated by the assumption that the status quo was itself generated by investigation of possible policies. Other "front-ends" to Proposition 1 are also possible and plausible; for example, that new firms are randomly assigned policies, but only new firms with policies of more-than-average profitability survive to the second period. The second key element is the assumption that the policy's profitability becomes known once it is adopted. The status quo has the advantage of being a known value, and hence not subject to regression to the mean. If the true profitability of the status quo were unknown, the conservative bias would disappear, because the status quo's estimated profitability would be completely symmetric to the estimated profitability of the newly investigated policies.

It should also be noted that the conservative bias may not be the only effect of poor information about alternative policies. In particular, if the firm can discover the value of a policy relatively quickly and reverse the adoption decision, then an innovative bias may be appropriate, as was mentioned in footnote 2. In the extreme, the firm could briefly adopt each possible policy in turn, and then choose the best one for a permanent policy. In such a case, adoption has two purposes— information collection, for the initial adoptions, and direct profitmaking for the serious adoption. The initial adoptions replace the staffer's estimate, and if the information acquired by adoption was imperfect like the staffer's, then a conservative bias would still be applied in making the final adoption decision.

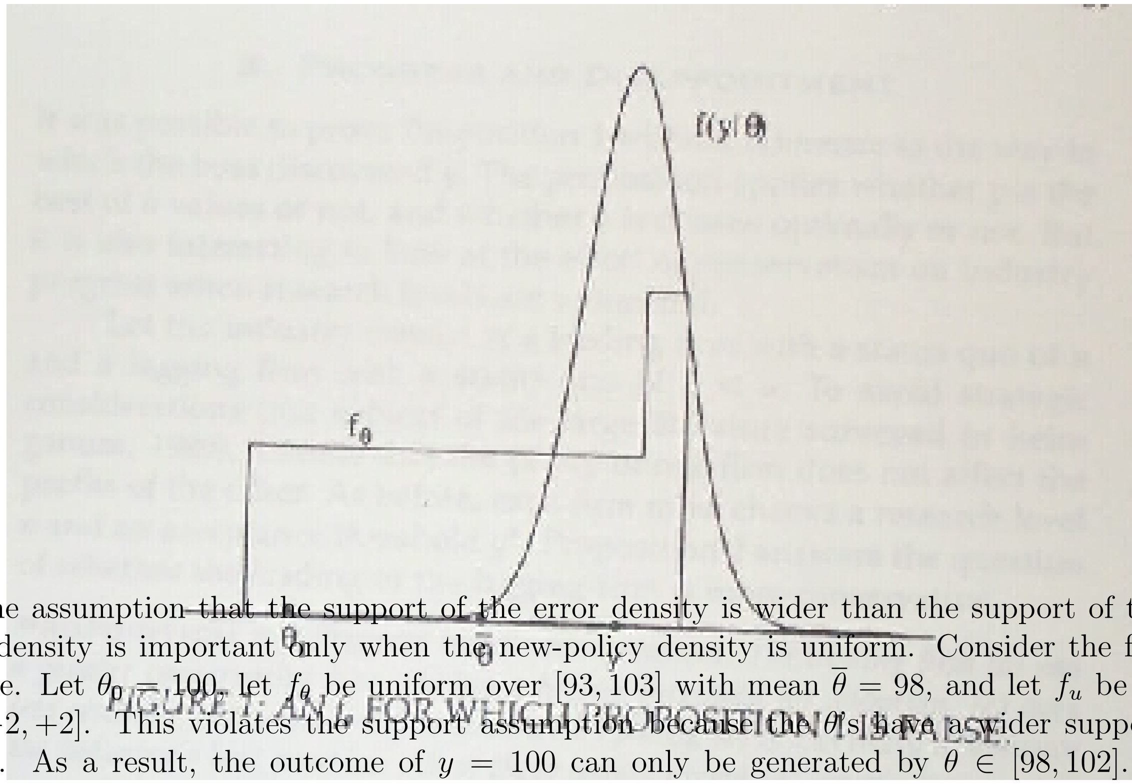
Proposition 1 applies to the average firm, not to every firm. A minority of firms will have been unlucky in the first period, and they will not have a conservative bias. They will innovate, but not due to a bias in the sense used here. Rather, their adoption procedure will be the same as that of a new firm: adopt the best investigated policy if its estimated profitability is greater than $\overline{\theta}$, and adopt a random uninvestigated policy otherwise. The unlucky firm will certainly abandon its status quo, but it will never adopt an investigated policy whose estimated value is less than the status quo— it would prefer to adopt an uninvestigated policy if the investigation proved disappointing. Although if $y < \theta_0 < \overline{\theta}$ it is true that $E(\theta|y) > y$, by Lemma 1, the lemma also says that if $y < \overline{\theta}$ then $E(\theta|y) < \overline{\theta}$, so the blind choice is preferable to adjusting an investigated policy's valuation upwards and accepting it. Thus, unlucky firms do not so much impose an innovative bias as simply start over. This extends to their choice of how many policies to investigate. Except that the profits from a policy can only last one period instead of two (which is a modelling artifact), the unlucky firm faces the same fallback position as the new firm, an expected profitability of $\overline{\theta}$, so status quo is irrelevant, because it is sure to be rejected.

In the case of either firm, average or unlucky, a virtue of the regression argument is that managers need not understand it to behave according to it: they can learn to be optimally conservative as a rule of thumb. The boss might well be conservative for the wrong reasons, thinking in terms of agency problems or irrationality—"My staff get emotionally attached to the projects they research." Even though he misunderstands the process, he could grasp that the conservative bias should be less for more accurate measurements: "Anderson is smarter than Brown, so he avoids exaggeration." Or, he might reach his conclusion by blind empiricism, adjusting the adoption threshold until he finds he is adopting only policies that are ex post superior to the status quo.

On the other hand, one might ask whether the regression argument is simple enough that the subordinate would apply it himself. This is a matter of interpreting the model, which divides the decision process into two stages: first a staffer measures the value of θ and delivers a report y, and then a boss calculates $(\theta|y)$ and makes the decision. This distinguishes two kinds of information processing. What business schools teach MBA students is how to use theory to measure the value of things. What they spend less time teaching, perhaps because it is harder to teach, perhaps because it is easier to learn, is a vague knowledge of how things actually are. Such "soft" knowledge comes with experience, and senior decisionmakers might positively prefer that their juniors not add noise to staff estimates by subjective adjustments from their tiny personal experience. Splitting the decision process divides hard knowledge from soft knowledge, measuring from deciding. The staffer watches the trees; and the boss, the forest. Were there no division of labor of this kind, the firm could dispense with the boss and simply have the staffer make the decisions. But even if there is just one person involved, the distinction between measurement and decision is useful. It could be the boss himself, not the staffer, who estimates the value. In one part of his mind, he believes his measurement of y to be an unbiased estimate of θ. If he is sophisticated, he steps back and realizes that he is probably overestimating.

The Technical Assumptions

Proposition 1 relied on the finding in Lemma 1 that $E(\theta|y)$ lies in the open interval $(E\theta, y)$; that is, the posterior mean lies between the prior mean and the data. This may seem obvious, but it is false for certain distributions which violate the assumptions of the model. The counterexample in Figure 1 shows why the assumptions exclude bimodal or skewed densities. In Figure 1, f_θ is weakly unimodal but right-skewed, and f_u has a small variance. Since f_u has a small variance, the observation of y was more likely generated by $\theta = m(\theta) > y$, where $m(\theta) > y$, than by smaller values of θ, so $\overline{\theta} < y < E(\theta|y)$. Symmetry without unimodality allows a similar perversity: if a probability peak were added at θ_1 to make f_θ symmetric, it would remain true that $\overline{\theta} < y < E(\theta|y)$.



The assumption that the support of the error density is wider than the support of the new-policy density is important only when the new-policy density is uniform. Consider the following example. Let $\theta_0 = 100$, let f_θ be uniform over $[93, 103]$ with mean $\overline{\theta} = 98$, and let f_u be uniform over $[-2, +2]$. This violates the support assumption because the θ's have a wider support than the u's. As a result, the outcome of $y = 100$ can only be generated by $\theta \in [98, 102]$. Hence, $E(\theta|(y = 100)) = 100$, and the new policy should be adopted if $y > 100$. The optimal threshold is $y^* = 100$, because values of y near θ_0 are as likely to be the result of negative as of positive measurement errors. Proposition 1 fails.

If, on the other hand, f_u is uniform, not over $[-2, +2]$, but over $[-10, +10]$, then the support assumption is satisfied. Then an outcome of $y = 100$ is generated by $\theta \in [90, 103]$, and $E(\theta|y = 100) = 96.5$. Thus, $E(y|\theta) < y$, and a conservative bias must be applied.[7]

Risk Aversion as an Altenative Explanation

The model assumes that the boss is risk-neutral, but risk aversion would also make a known status quo preferable to an uncertain alternative. This is true whether the uncertainty arises from exogenous events in the world or from the measurement error of the staffer, since in either case the boss cannot perfectly predict profitability. Which argument, regression or risk aversion, more reasonably explains the conservative boss?

Table 2 illustrates a variant model in which the decisionmaker is risk-averse. The status quo is $\theta_0 = 100$, and $\overline{\theta} = 100$ also, so Proposition 1 fails to apply and $E(\theta|y) = y$. As in Table 1, the standard deviation of the new policy population is $\sigma_\epsilon = 15$. The acceptance threshold is $100 + P\left(\frac{\sigma_\epsilon^2 \sigma_u^2}{\sigma_\epsilon^2 + \sigma_u^2}\right)$, because each unit of variance requires a premium of P.[8]

[7] If f_θ is not uniform, the support assumption is unnecessary. If, for example, f_θ is triangular over $[93, 103]$ and f_u is uniform over $[-2, +2]$, the support assumption is violated, and $y = 100$ could only be generated by $\theta \in [98, 102]$. The probability that $y = 100$ was generated by $\theta \in [98, 100]$, however, is greater than the probability of $\theta \in [100, 102]$ (which is farther from the mean of 98). Therefore, $y^* > 100$, and there is a conservative bias.

[8] Note that the value of the status quo and the measured value of the new policy are assumed to be certainty

Table 2
Acceptance Thresholds for a Risk-Averse Decisionmaker
$\left(100 + P\left(\frac{\sigma_\epsilon^2 \sigma_u^2}{\sigma_\epsilon^2 + \sigma_u^2}\right)\right)$

		Standard Error of Measurement (σ_u)			
		5	10	15	40
	0.104	102.3	107.2	**112**	121
Price	0.052	101.2	103.6	106	110
of	0.026	100.6	101.8	102.9	105
Risk	0.013	100.3	100.9	101.5	102.6
(P)	0.0	100.0	100.0	100.0	100.0
Assumed: $\theta_0 = 100$, $\overline{\theta} = 100$, $\sigma_\epsilon = 15$ (fractions rounded)					

What price of risk is reasonable? The situation being modelled is a boss facing idiosyncratic risk, rather than an investor facing market risk, so one might think that the price should be negligible and risk aversion is unimportant, but let us suppose that idiosyncratic risk does matter for some reason such as the necessity of tying managerial compensation to firm value. The stock market's risk premium provides a benchmark for the price of risk. The mean real return on the stock market from 1889 to 1978 was 7.0% and the standard deviation was 16.5, compared with a mean of 0.8 and a standard deviation of 5.7 for low-risk securities. An increase of 240 percent-squared in variance thus requires an increase of 6.2% in return, a price of $P = 0.026$ percent-squared.[9]

Consider again the boss who refuses a new policy with measured value of 110 in favor of a status quo with value 100. Suppose that $\sigma_u = 15$. From Table 1, the regression argument explains the boss's behavior if $P = 0$ and $\overline{\theta} = 90$. From Table 2, risk aversion explains it if $P = 0.104$ and $\overline{\theta} = 100$. The amount of risk aversion needed seems high—four times as high as the market price of risk. Moreover, since the project risk is idiosyncratic, it should actually be priced lower.

The regression argument has different empirical implications than simple risk aversion. One difference is that if the distribution of new policies has greater variance (σ_ϵ^2 increases), regression has a weaker effect, but risk aversion, a stronger. [10] A second difference is that the regression effect does not depend on the covariance of the random terms with other assets in the economy, unlike

equivalents, except for measurement error. If the status quo has a non-stochastic return and the new policy is risky (even beyond measurement error), then the problem is simply that the staffer has not used the proper discount rate. The problem analyzed in this section is different: it asks whether the risk due to measurement error is important relative to the regression effect.

[9] The returns are from p. 147 of Mehra & Prescott (1985), who constructed them from the annual average Standard and Poor's Composite Stock Price Index, the consumption deflator, and various low-risk short-term securities. The point of their article is that the amount of implied relative risk aversion is implausibly high, so 0.026 is most likely an overestimate.

[10] Regarding risk aversion, note that in the example with normal densities, equation (5) says that $Var(\theta|y) = \frac{\sigma_u^2 \sigma_\epsilon^2}{\sigma_u^2 + \sigma_\epsilon^2}$. σ_ϵ^2 is $\frac{dVar(\theta|y)}{d\sigma_\epsilon^2} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2} - \frac{\sigma_u^2 \sigma_\epsilon^2}{(\sigma_u^2 + \sigma_\epsilon^2)^2} > 0$.

risk aversion. Thus, the conservatism due to the regression argument should not depend on the beta of the measurement error or the degree to which the decisionmaker is diversified.

2. Progress and Disappointment

It was possible to prove Proposition 1 without reference to the way in which the boss discovered y. The proposition applies whether y is the best of n values or not, and whether n is chosen optimally or not. But it is also interesting to look at the effect of conservatism on industry progress when research levels are optimized.

Let the industry consist of a leading firm with a status quo of α and a lagging firm with a status quo of $\beta < \alpha$. To avoid the strategic considerations that are the subject of the large literature surveyed in Reinganum (1989), assume that the policy of one firm does not affect the profits of the other. As before, each firm must choose a research level n and an acceptance threshold y^*. Which firm is more conservative?

PROPOSITION 2: Progress instills conservatism. The leading firm
(a) has a greater conservative bias,
(b) has a higher threshold for adoption,
(c) does less research,
(d) is less likely to advance its policy,
(e) is less likely to advance the industry's best practice.
All of these points except (a) remain true even if research discovers a policy's value with perfect accuracy.

Proof: See Appendix.

Points (a) and (b) of Proposition 2 say that the leading firm will require a greater apparent advantage of the new policy over its status quo as well as having a higher threshold. As the policy in use improves, it becomes less and less likely that a new policy is genuinely better. Unless there is an exogenous shock to the system that improves the new-policy pool, firms should become more and more skeptical of apparently superior new policies. Points (c), (d) and (e) concern the amount of research done by the leading firm. Point (c) is the most difficult point to prove. The simple intuition is that for a given level of research the leading firm is more likely to reject every investigated policy, and so it is less willing to spend on research. The complications arise because the lagging firm is more likely to find an improvement with even a small amount of research, so additional research might be redundant and it is not clear without careful analysis which firm has the higher return from research at the margin. Since it does turn out that the leader does less research, in addition to rejecting a greater fraction of investigated policies, the leader is less likely to improve. Since the lagging firm adopts innovations more often, it is not only more likely to improve over its own current policy, but also more likely to improve over the leader's current policy.

Proposition 2 has implications for the life cycle of the industry. Even if the industry starts off only mildly conservative, it will become more conservative as time passes. New policies are adopted only if they are expected to be superior to the status quo, which becomes less probable as the status quo improves. Moreover, since firms give less thought to changing their policies as they improve, the lagging firm is both more likely to improve and more likely to advance the best practice of the industry. On the industry level, this generates the familiar notion that firms in a young industry will be bolder and innovate more than firms in a mature industry, and that struggling firms are more likely to take risks than successful firms.

Two caveats should be made. First, interactions between firms have been ignored, since such interactions have tremendous variety and can favor research by either the leader or the laggard.[11] Second, it was implicitly assumed that the two firms are equally good at the technology of innovation. If one firm has a superior ability to find good policies or to estimate policy values, that firm will more likely become the leader and will have a tendency to do more research which must be balanced against the conservative bias arising from its advanced technology.

The regression argument implies another form of apparently irrational behavior consistently adopting innovations that turn out to be disappointing. Let us define "disappointment" to be the decisionmaker's state of mind when he discovers that the new policy is less valuable than predicted by the forecast (i.e., θ is less than y). This definition is appropriate if one believes that a person can rationally appreciate his own bounded rationality and expect to be disappointed. A decision can lead to disappointment ($\theta < y$), but still be correct ($\theta > \theta_0$). From Proposition 1 and the fact that all policy changes accepted have $y > \theta_0$, it is easy to see why adoption will on average be followed by disappointment: that is simply another way to state that $E(\theta|y) < y$. This idea, which can essentially be found in Brown (1978) and Harrison & March (1984), is listed here as Proposition 3.

PROPOSITION 3: A rational decisionmaker will be disappointed on average when he adopts a new policy: $E(\theta|y) < y$.

3. Empirical Implications

Proposition 1 has strong empirical implications, because it cannot explain genuine managerial risk aversion, but it can explain *apparent* risk aversion, reconciling our usual belief that managers maximize profits with the common perception that in this particular aspect of decisionmaking they do not. The model explains this perception as the result of outsiders observing managers rejecting projects which unbiased analysts have stated are superior to the status quo. The opinion of analysts and outsiders may never be contradicted, because the true value of a rejected project is never discovered by the firm that rejects it. In the case of unused technological innovations, it may never be discovered by anybody, and the outsider will strongly suspect that the firm has done its research to acquire "sleeping patents" for strategic purposes. If the project is not patentable, then it might be adopted by another firm, but this too might serve to confirm the outside observer's suspicions, since sometime the rejected project will turn out to be profitable after all. Using the metric "number of successful adoptions," to compare managers, the naive manager who takes estimates at face value, or the manager of a firm that begins with an inferior status quo would win, because the sophisticated manager of an already-successful firm would indeed reject more truly profitable projects—as a simple consequence of having higher standards. The sophisticated manager also rejects more unprofitable projects, but this could easily be overlooked.

[11] Examples: The interaction favors the leader's research if consumer switching costs induce once-and-for-all switches when one firm acquires a large enough lead. The interaction favors the lagging firm's research if there are leakages from innovation so the lagging firm shares in the leader's progress.

The regression argument can also explain why many managers share the mistrust of Hayes & Garvin (1982) concerning academically uncontroversial capital budgeting methods such as discounted cash flow. If estimation of the discounted value is done without regard to the regression argument, the recommendations will be wrong. Schnall & Sundem (1980) expected to find that a survey of companies' capital budgeting techniques would find greater use of formal methods in firms with riskier environments. Instead, they found the opposite: risky firms are more informal. The regression argument has an explanation: where measurement error is greater, the naive use of discounted cash flow leads to more frequent unprofitable innovation. Firms with high measurement error might find that trained intuition works better than naive discounting, although a sophisticated use of discounted cash flow might work better still.

The regression argument explains perceptions of myopia in the same way as perceptions of risk aversion, if one adds the assumption that projects with more distant returns are measured less accurately. If that is true, rational managers will impose a greater conservative bias for long-term projects, and it's estimated return will have to pass a higher hurdle rate. Managers will also reject a greater percentage of genuinely good long-term projects, because a greater proportion of the measured high values will be due to bad projects. This increased conservatism towards long-term projects might be interpreted as irrational myopia.

Some readers may be troubled by the existence of innovative industries, because the regression model does not seem to apply there. In the computer industry, after all, the firm that does not innovate does not survive. It does not follow, however, that computer firms are not conservative in the sense of this article, turning down a multitude of projects with apparently positive present discounted values. Nothing in the model says that a firm with a conservative bias will not innovate; only that it will be careful when it innovates. If a single staffer is sent out, a conservative firm will most likely retain the status quo. But it might be optimal to investigate 200 new policies with the expectation that 10 will appear superior to the status quo, that only 3 will pass the threshold for acceptance, and that if none pass it, the best alternative is to exit the industry. A firm can be both extremely conservative in its decisionmaking and extremely likely to innovate. Whether the computer industry is conservative in this sense is an interesting empirical question.

The regression argument can explain the more specific empirical observations mentioned in the introduction. Hayes & Garvin (1982) found hurdle rates of 30% or more, compared to actual returns of 20%. This is exactly the finding that $y^* > \theta_0$ on average. Pruitt & Gitman (1987) found that 80% of responding financial officers thought that measurements were overoptimistic, 59% thought that decisionmakers scaled them back, 37% blamed intentional overestimation and 36% blamed inexperience. This is evidence that $E(\theta|y) < y$ in the opinion of decisionmakers, and that they attribute this to both agency problems and to staffer ignorance (not knowing f_θ). Ross (1986) found that the hurdle rate for large energy conservation projects was near the cost of capital, that for small projects the hurdle rate was higher, and that for small projects the decisionmakers made informal adjustments to the estimates. The distinction between large and small projects is interesting, and might be due to more precise staffers being assigned to larger projects, or to the staffer for the large project being the same person as the decisionmaker, and thus able to add the conservative bias.

Two additional studies are relevant. Beardsley & Mansfield (1978) looked at data on the forecasted and actual success of new products and processes developed between 1960 and 1965 by an anonymous multibillion-dollar company. Of the 57 new projects, six were within 10% of the forecast, 25 had pessimistic forecasts, and 26 had optimistic forecasts, with a tendency to pessimism for large and optimism for small projects. This is the same relative pattern observed by Ross.[12]

Little and Mirrlees (1991) report on results from an internal World Bank report by Pohl & Dabrarko (1989) which compared ex ante project appraisals to ex post estimates of the returns. The appraisals averaged 17 percent in 1968 and rose in a clear trend to 29 percent in 1980, possibly driven by the "McNamara effect" of an executive who favored increased lending. What is more relevant to the regression argument is that the average ex post estimate, showing no similar trend, ranged between 13 percent and 17 percent, averaging 16 percent, and that this was well above the cost of capital. This is consistent with the regression argument, since the ex post estimate is still, in the case of these public, LDC projects, an estimate with a possibly large error.[13]

4. A Miscellany of Applications

The regression argument applies to a variety of situations in which the value of a status quo is known precisely and innovation would not be made to a random alternative. This section suggests a few speculative applications.

The Fallacy of Sunk Cost. Suppose that the status quo is 100 dollars that might be invested safely at a return of 5% or spent on the new project, that $y^* = 150$, and that upon observing $y = 200$, the manager adopts the project. Suppose further that the manager observes the true value θ after 20 dollars is spent, and he discovers that it equals 100. The manager will continue with the project, paying the additional 80 to get the value of 100 (a return of 25%) . He has received bad news, but his information is now precise. An outside observer, however, would see that the measurement of the total return has dropped from 100% to 0%, and the return from continuing the project has sunk to 25%, which is below the original 50% threshold for acceptance, yet the manager, an apparent victim of the fallacy of sunk cost, refuses to abandon the project. If, in general, the information about a project's value becomes more accurate as more is spent on the project, the threshold for continuing the project will fall over time, giving rise to apparent ignorance of the fallacy of sunk cost.

Loss Aversion. Libby & Fishburn (1977) note that various studies find experimental subjects to be more averse to losses than to risk *per se.* The regression argument explains this if potential loss is measured more accurately than potential gain. If the experimental subject begins with a known 100 dollars, and is offered, in exchange, a fair gamble with a known bad outcome of 0 and an

[12] The inaccurate forecasts for large projects are puzzling. Classification bias could explain this (if a project with a small forecast is successful enough, it becomes a large project), or the unusual macroeconomic growth of the 1960's, which for most industries would have made rational forecasts look pessimistic.

[13] Little and Mirrlees suggest another reason why estimates would be overoptimistic on average: If the hurdle rate is 10 percent, and it becomes clear to the staff that the estimate will be well above 10 percent, then the staff will be less careful to avoid large positive errors. If the estimate seems to be close, on the other hand, the staff will measure more carefully. On average, this leads to overestimates, but harmless ones.

unknown good outcome estimated at 201, the regression argument applies, and the subject should behave conservatively. If, on the other hand, the subject starts with 0 dollars and faces a choice between a nonstochastic outcome estimated at 100 dollars and the risky gamble between 0 and 201, the effect of the regression argument is not clear, since it applies to both alternatives. But the first tradeoff, with the known 100, is more usual, because in business and personal decisions alike, the potential loss—the resources invested—is usually known with more accuracy than the potential gain.

The Ellsberg Paradox. Suppose that two urns are each filled with 100 black and white balls. Urn X has 50 white and 50 black balls, whereas urn Y has an unknown number of each color. Experimental subjects should be indifferent between betting on the draw of a black ball from urn X and betting on the draw of a black ball from urn Y, but they generally prefer urn X, which has the known probabilities.

Bordley & Hazen (forthcoming) explain this using the idea that players have pessimistic priors over unknown distributions. The subject of the experiment fears that the experimenter will fill urn Y with fewer black balls, and that even if the subject were offered a choice between white and black, the experimenter would somehow rig the game.

The regression argument elaborates on this slightly. Assume that the subjects know that most strangers who offer bets have stacked the odds in their favor. In the experiment with black and white balls, the experimenter may do his best to convince the subjects that he is offering a fair bet. But the rational subject should be aware that he can be fooled, particularly in a situation contrived by a person clever enough to have a doctorate in mathematics. The subject does not trust himself to have truly figured out all the angles of the experiment, and so he picks urn X, which is easier to understand. Not because of measurement error, but because of recognizedly bounded rationality, the subject regresses the odds of urn Y towards the unfavorable odds of the typical tricky bet.

Scientific Conservatism. Under the plausible assumption that the average randomly conceived explanation for a phenomenon is inferior to the current explanation, the regression argument predicts that scholars would be slow to accept a new idea, even if it seems correct to them. (In fact, if the regression argument is correct, scholars should be slow to accept the regression argument!) The way classical hypothesis testing is used may be an example of this. Classical hypothesis testing imposes a strong bias towards not rejecting the null hypothesis. That might be justified as a rule of thumb, if the regression argument is applicable because a randomly selected alternative hypothesis is likely to be further from the truth than the null hypothesis.

Marriage. If a woman requires that a man be 20% better than a randomly drawn man if she is to abandon the status quo of single life, the regression argument says that the man's measured value must be more than 20% better than average. She will hesitate to marry even if the potential husband seems to meet her standards. Of course, unless women are naive—which is perhaps plausible in this context—it is not possible for every woman's husband to be above average in the same characteristic. If tastes differ over characteristics, however, this becomes a matching problem, such as has been analyzed by various authors (see Mortensen [1988] for references), and every woman could expect her husband to be above average in the characteristics she values. Empirically, one could work backwards from the regression argument to check its assumptions in this application. If, for example, most wives lower their estimates of their husbands after marriage, one might conclude that they had attempted to marry above-average men. If, on the other hand, your wife is not disappointed in you, she probably wasn't very selective in her choice of a husband.

Despite this variety of applications, the regression argument is not capable of being twisted to explain every kind of conservative behavior. Below are listed a number of situations to which it does *not* apply.

Not the High Risk Premium on Equity. The regression argument might have the implication that the market value of an asset would be less than its measured value. But it cannot explain the equity premium puzzle of Mehra & Prescott (1985). If ex ante stock returns were very high relative to the riskless rate of return, the reason might be the regression argument rather than risk aversion. But what Mehra and Prescott point out is that *ex post* stock returns are peculiarly high, whereas the regression argument says that ex post there is no conservative bias.

Not Takeovers. Roll (1986) proposes, in his "hubris hypothesis," that the lack of statistically significantly positive returns to bidders in takeovers might be explained by naive optimism. There is an asymmetry because pessimistic bidders never attempt a takeover. Hubris is not formally modelled, but it might arise from lack of recognition of the regression argument. Naive decisionmaking is unusually foolish in this context, however: in a takeover, the bidder knows that other managers have looked at the same project, and rejected it, and that the target managers are likely to have information superior to his own. Miller (1977) follows a similar line of reasoning, but suggests that the winning bidder is not alone in his optimism, which is what forces up the takeover price. If one believes that bidders are irrational, the regression argument can be used to explain the precise nature of their mistakes, but the story needs irrationality as well as the regression argument.

Not Conservative Accounting Practices. Employees sometimes steal a company's goods, but they never surreptitiously add to them. Hence, when those goods are audited, the auditor should expect the amount of goods to be less than the nominal amount. Moreover, there is measurement error in the audit, particularly if it is using random sampling. One might conclude that if the audit shows the amount of goods to be close to the nominal amount, the decisionmaker should adjust his estimate downwards, applying what might be called a conservative bias. This is an application of regression towards the mean, but not the same application as in this paper. The choice is not between a status quo and a new project; the audit is a pure measurement problem. Moreover, if the first step of the audit reveals an unusually low measured value, the value should be adjusted *upwards* (which will happen in a full fifty percent of cases), which does not fit our notion of conservative accounting practices. Hence, the regression argument does not seem to explain such practices.

5. Concluding Remarks

Using the idea of regression towards the mean, this article offers a normative reason for conservatism in decisionmaking and a positive reason for perceptions of excess conservatism. The theory is simple— simpler than the complex decision trees and refinements of risk aversion that one finds in textbooks on decision theory and finance, and simpler than the agency arguments used to explain conservatism in the economics literature. The key assumptions are that the value of the status quo is known better than the value of innovations, and that random uninvestigated innovation is unprofitable. With the addition of mild technical assumptions, these assumptions imply that raw measurements of innovation values, even if made with unbiased error, will lead to overinnovation if taken literally as the basis for decisions. Managers should be conservative, and their conservatism should increase with the success of their policies.

Appendix: Proofs.

LEMMA 1: (i) $E(\theta|y)$ is increasing in y and (ii) $\overline{\theta} < E(\theta|y) < y$ for $y > \overline{\theta}$, $E(\theta|y) = y$ for $y = \overline{\theta}$, and $y < E(\theta|y) < \overline{\theta}$ for $y < \overline{\theta}$.

Proof: What must be shown is that the posterior mean is (i) increasing in the observed datum, and (ii) lies in the open interval between the prior mean and the observed datum. This posterior mean, the conditional expectation of θ, is

$$E(\theta|y) = \int \theta f(\theta|y) d\theta. \tag{10}$$

Using Bayes Rule:

$$f(\theta|y) = \frac{f(y|\theta)f(\theta)}{f(y)}. \tag{11}$$

Substituting for $f(\theta|y)$ from (11) into (10) gives

$$E(\theta|y) = \int \theta \left(\frac{f(y|\theta)f(\theta)}{f(y)}\right) d\theta = \left(\frac{1}{f(y)}\right) \int \theta f(y|\theta) f(\theta) d\theta. \tag{12}$$

(i) The first part of (12) can be written as

$$E(\theta|y) = \int \theta f(\theta) \left(\frac{f(y|\theta)}{f(y)}\right) d\theta. \tag{13}$$

Since $\overline{\theta} = \int \theta f(\theta) d\theta$, the question is what effect multiplying each θ in the integral by the weights $\frac{f(y|\theta)}{f(y)}$ has. For given y, this weight is greatest for $\theta = y$ and is equal for $\theta = y + \delta$ and $\theta = y - \delta$ given the assumption that f_u is symmetric and weakly unimodal. Hence, if y increases, the weights on all the values of θ less than y decrease and those on the greater values increase, so $E(\theta|y)$ will increase in y.

(ii) Fix y at some level greater than $\overline{\theta}$. It will be shown that $E(\theta|y) < y$, and the other results will follow easily. Because $f(y|\theta)$ is symmetric around $\theta = y$,

$$\int \theta f(y|\theta) d\theta = y. \tag{14}$$

It is sufficient to show that multiplying θ by $f(y|\theta)f(\theta)$ in the integral in (12), instead of by $f(y|\theta)$ as in (14), puts relatively greater weight on smaller values of θ. (Dividing by $f(y)$ just makes the density integrate to one.) The symmetry of f_u implies that $f(y|(\theta = y - \delta)) = f(y|(\theta = y + \delta))$ for any δ, so what needs to be shown is that

$$f_\theta(y - \delta) \geq f_\theta(y + \delta), \tag{15}$$

with strict inequality for at least one value of θ such that $f(y|\theta) > 0$ (otherwise, $f(\theta)f(y|\theta) = 0$, and the strictness of (15) would be irrelevant). Given (15), multiplication by $f(\theta)$ in expression (12) weights the values of θ less than y more heavily than those greater than y.

Inequality (15) is true for $\delta < y - \overline{\theta}$ because $f'_\theta \leq 0$ in that range, by the assumption of weak unimodality of f_θ. Inequality (15) is true for $\delta > y - \overline{\theta}$ because $y - \delta$ is nearer to $\overline{\theta}$ than is $y + \delta$, and the symmetry and weak unimodality of f_θ together imply that $f_\theta(\theta)$ is weakly greater for values of θ nearer $\overline{\theta}$.

Inequality (15) is strict for at least one θ such that $f(y|\theta) > 0$, either because $f'_\theta < 0$ for some $\delta < y - \overline{\theta}$ (if f_θ is sloping) , or because there is some δ such that $f_\theta(y - \delta) > f_\theta(y + \delta) = 0$ (if f_θ is flat). This is true for flat f_θ under the assumption that f_u has a larger support than f_θ. The support of $f(y|\theta)$ overlaps with at least the most positive half of the support of f_θ, so $f(y|\theta) > 0$ for at least one δ such that inequality (15) is strictly satisfied. Hence, equation (12) does have heavier weight than (14) for some θ such that $f(y|\theta) > 0$.

Given that equation (14) shows that the unweighted function integrates to y, the heavy weights on low values of θ in equation (12) imply that the weighted function integrates to less than y. Hence, $E(\theta|y) < y$ if $y > \overline{\theta}$.

Parallel arguments show that $E(\theta|y) = y$ for $y = \overline{\theta}$, and $E(\theta|y) > y$ for $y < \overline{\theta}$. Given that $E(\theta|y = overline\theta) = \overline{\theta}$ and that $E(\theta|y)$ is increasing in y from part (i), it follows that $E(\theta|y) < \overline{\theta}$ for $y < \overline{\theta}$ and $E(\theta|y) > \overline{\theta}$ for $y > \overline{\theta}$. This completes the description of the intervals in which the posterior means lie. □

PROPOSITION 2: Progress instills conservatism. The leading firm
(a) has a greater conservative bias,
(b) has a higher threshold for adoption,
(c) does less research,
(d) is less likely to advance its policy,
(e) is less likely to advance the industry's best practice.
All of these points except for (a) remain true even if research discovers a policy's value with perfect accuracy.

Proof: Part (a) says that $y^* - \theta_0$ increases in θ_0. This is equivalent to $y^* - E(\theta|y^*)$ increasing in θ_0. To see this, refer to the proof of Lemma 1, starting with inequality (15). Assume that $y > \theta_0$. The proof showed that if $f_\theta(y - \delta) > f_\theta(y + \delta)$ for some θ such that $f(y|\theta) > 0$, then $y - E(\theta|y) > 0$. For the present proof, the size of $y - E(\theta|y)$ matters. The size of $y - E(\theta|y)$ increases with the difference $f_\theta(y - \delta) - f_\theta(y + \delta)$ and with the range of θ for which the difference is positive. If y is greater, then since f_θ is unimodal, either or both of the difference and range increase. Hence, $y - E(\theta|y)$ increases with y, and since y^* increases with θ_0, part (a) is proved.

Part (b) says that y^* rises with θ_0. y^* is chosen so that $E(\theta|y^*) = \theta_0$, and since Lemma 1 established that $E(\theta|y)$ increases in y, it follows that for larger θ_0, y^* is larger.

Parts (c), (d), and (e) pertain to optimal choice of n. Define $y_m(n) \equiv Max\{y_i\}_{i=1,...n}$. The boss will adopt the innovation if and only if $y_m(n) > y^*$; otherwise, the investigation was useless, *ex post*.

The advantage of searching is that $Ey_m(n)$ increases, so that $E(\theta|y_m(n))$ increases also. Values of $y_m(n)$ below y^* are useless, and assigning zero weight to these, the expected value of a research level of n is

$$V(n) = \theta_0 F(y^*)^n + \int_{y^*}^{\infty} E(\theta|y) n f(y) F(y)^{n-1} dy - cn. \tag{16}$$

The reasoning behind this is as follows. If $y_m(n) < y^*$, then all new policies are rejected, and the firm's value remains the status quo value of θ_0. The probability of any one new policy being worse than the status quo is $F(y^*)$, so the probability of all of them being worse is $F(y^*)^n$ and the first term of (16) is $\theta_0 F(y^*)^n$. The second term represents the values and probabilities of accepted policies. Only values greater than y^* are relevant, giving the integration bounds of y^* and ∞. Consider a measurement y, with expected true value $E(\theta|y)$, probability density $f(y)$ for each of the n measurements, and probability mass $nf(y)$ over all of them. With probability $F(y)$, each of the $n-1$ other measurements is less than y, so with probability $F(y)^{n-1}$ they are all less. The third term arises from the research cost cn.

It is convenient to rewrite $V(n)$ as follows:

$$V(n) = \theta_0 + \int_{-\infty}^{\infty} \int_{y^*}^{\infty} \left[1 - F(y)^n\right] f(\theta|y) dy d\theta - cn. \tag{17}$$

This expression equals equation (16) because it can be integrated by parts to get

$$\begin{aligned} V(n) &= \theta_0 + E(\theta_y) y \left[1 - F(y)\right]|_{y^*}^{\infty} - \textstyle\int_{y^*}^{\infty} \left[-nF(y)^{n-1} f(y) E(\theta|y)\right] dy - cn \\ &= \theta_0 + 0 - E(\theta|y^*) + E(\theta|y^*) F(y^*)^n + \textstyle\int_{y^*}^{\infty} \left[nF(y)^{n-1} f(y) y\right] dy - cn \\ &= \theta_0 F(y^*)^n + \textstyle\int_{y^*}^{\infty} E(\theta|y) n f(y) F(y)^{n-1} dy - cn, \end{aligned} \tag{18}$$

which is (16).

Using (17), it can be seen that there are diminishing returns to research. n is a discrete variable, but one may differentiate with respect to it and ignore all but discrete values. This gives an expression for the marginal value of search:

$$\frac{\partial V}{\partial n} = - \int_{-\infty}^{\infty} \int_{y^*}^{\infty} F(y)^n log(F(y)) f(\theta|y) dy d\theta - c \tag{19}$$

Expression (19) is positive for small c, because F is less than one. But search has diminishing returns:

$$\frac{\partial^2 V}{\partial n^2} = - \int_{-\infty}^{\infty} \int_{y^*}^{\infty} F(y)^n log^2(F(y)) f(\theta|y) dy d\theta < 0. \tag{20}$$

Also, the marginal benefit of search falls with y^*, because

$$\frac{\partial^2 V}{\partial n \partial y^*} = \int_{-\infty}^{\infty} F(y^*)^n log(F(y^*)) d\theta < 0. \tag{21}$$

These inequalities can be used to prove part (c), which says that n^*, the optimal level of research, falls in θ_0. Using the implicit function theorem, $\frac{dn}{dy^*} < 0$. As the status quo improves (for if

θ_0 increases, so does y^*), n^* does not increase, and may decrease. (The inequality is not strict, because n is a discrete variable and for small changes in y^*, n^* might not change.)

Part (d) says that $Prob(y_m > y^*(\theta_0))$ falls with θ_0. The leading firm does less research and has a higher threshold, both of which reduce the probability that $y_m > y^*$, so this is clearly true.

Part (e) says that $Prob(y_m > \alpha)$ falls with θ_0. Let the leading firm's optimal acceptance threshold be y^*_α. Although both firms adopt policies with $y > y^*_\alpha$, the lagging firm also adopts some other policies. These other policies are most likely inferior, but with positive probability their true value is greater than α. Hence, the lagging firm has a higher probability than the leading firm of choosing a policy with true value better than the best existing policy in the industry.

If research is perfectly accurate, $u = 0$ and $y = \theta$. This implies that $y^* = \theta_0$. Part (b) obviously is still true, because the leading firm has a higher θ_0. None of the proof of parts (c), (d), and (e) used the fact that $y \neq \theta$, so perfectly accurate research is a special case covered by those parts of the proposition. Only part (a) is invalid, which it is because with accurate research the conservative bias equals zero for both firms. $\square$

References.

Beardsley, George & Edwin Mansfield (1978), "A Note on the Accuracy of Industrial Forecasts of the Profitability of New Products and Processes," *Journal of Business*, 51: 127-135.

Bordley, Robert & Gordon Hazen (forthcoming) "SSB and Weighted Linear Utility as Expected Utility with Suspicion,"*Management Science*.

Brown, Keith (1974), "A Note on the Apparent Bias of Net Revenue Estimates for Capital Investment Projects" *Journal of Finance*, September 1974, 29, 4: 1215-16.

Brown, Keith (1978), "The Rate of Return of Selected Investment Projects," *Journal of Finance*, September 1978, 33,4: 1250-1253.

Casella, George (1985), "An Introduction to Empirical Bayes Data Analysis," *American Statistician*, May 1985, 39: 83-87.

Day, Richard (1987), "The General Theory of Disequilibrium and Economic Evolution," in D. Batten, J. Casti, and B. Johansson, eds., *Economic Evolution and Structural Adjustment*, Berlin: Springer-Verlag, 1987.

Harrison, J. & J. March (1984), "Decision Making and Postdecision Surprises," *Administrative Science Quarterly*, March 1984, 26-42.

Hayes, Robert & David Garvin (1982), "Managing as if Tomorrow Mattered," *Harvard Business Review*, May-June 1982, Vol 60, pp. 71-79

Heiner, Ronald (1983), "The Origin of Predictable Behavior," *American Economic Review*, 83: 560-595.

Hirshleifer, David & Anjan Thakor (1991), "Managerial Conservatism, Project Choice, and Debt," UCLA AGSM Finance Working Paper #14-89, January 1991.

Holmstrom, Bengt & Joan Ricart (1986), "Managerial Incentives and Capital Management," *Quarterly Journal of Economics*, November 1986, 835-860.

Kuran, Timur (1988), "The Tenacious Past: Theories of Personal and Collective Conservatism." *Journal of Economic Behavior and Organization*, 10: 143-171.

Lambert, R. (1986), "Executive Effort and Selection of Risky Projects," *Rand Journal of Economics*, Spring 1986, 17: 77-88.

Libby, Marc & Peter Fishburn (1977), "Behavioral Models of Risk Taking in Business Decisions: A Survey and Evaluation," *Journal of Accounting Research*, 15: 272-292.

Little, Ian & James Mirrlees (1991), "Project Appraisal and Planning Twenty Years On," *Proceedings of the World Bank Annual Conference on Development Economics, 1990*, Washington: World Bank, 1991.

Mehra, Rajnish & Edward Prescott (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15: 145-61.

Miller, Edward (1977), "Risk, Uncertainty, and Divergence of Opinion," *Journal of Finance*, September 1977, pp. 1151-67.

Miller, Edward (1978), "Uncertainty Induced Bias in Capital Budgeting," *Financial Management*, Autumn 1978, pp. 12-18.

Miller, Edward (1987), "The Competitive Market Assumption and Capital Budgeting Criteria," *Financial Management*, Winter 1987, pp. 22-28.

Mortensen, Dale (1988), "Matching: Finding a Partner for Life or Otherwise," *American Journal of Sociology*, 94 (supplement): S215-S240.

Narayanan, M. (1985), "Managerial Incentives for Short-Term Results," *Journal of Finance*, December 1985, 1465-1484.

Pohl, Gerhard & Mihaljek Dabrarko (1989), "Project Evaluation in Practice: Uncertainty at the World Bank," Economic Advisory Staff, World Bank.

Pruitt, Stephen & Lawrence Gitman (1987), "Capital Budgeting Forecast Bias: Evidence from the Fortune 500" *Financial Management*, Spring 1987, 46-51.

Reinganum, Jennifer (1989) "The Timing of Innovation: Research, Development and Diffusion" In Schmalensee, Richard & Robert Willig, eds. *The Handbook of Industrial Organization.* New York: North-Holland, 1989.

Roll, Richard (1986), "The Hubris Hypothesis of Corporate Takeovers," *Journal of Business*, 59: 197-216.

Ross, Marc (1986), "Capital Budgeting Practices of Twelve Large Manufacturers" *Financial Management*, Winter 1986, 15,4: 15-22.

Sah, Raaj & Joseph Stiglitz (1988) "Committees, Hierarchies, and Polyarchies," *Economic Journal* (June 1988) 98: 451-70.

Schnall, Lawrence & Gary Sundem (1980), "Capital Budgeting Methods and Risk: A Further Analysis," *Financial Management*, (Spring 1980) 9: 7-11.

Shleifer, Andrei & Robert Vishny (1989), "Equilibrium Short Horizons of Investors and Firms," mimeo, University of Chicago Graduate School of Business, November 1989.

Smidt, S. (1979), "A Bayesian Analysis of Post-Audits," *Journal of Finance*, June 1979, 34,3: 675-688.

Stein, Jeremy (1988), "Takeover Threats and Managerial Myopia," *Journal of Political Economy*, February 1988, 96:61-80.

Weitzman, Martin (1979), "Optimal Search for the Best Alternative," *Econometrica*, May 1979, 47: 641-54.