

Learning to Be Overconfident

Simon Gervais

University of Pennsylvania

Terrance Odean

University of California, Davis

We develop a multiperiod market model describing both the process by which traders learn about their ability and how a bias in this learning can create overconfident traders. A trader in our model initially does not know his own ability. He infers this ability from his successes and failures. In assessing his ability the trader takes too much credit for his successes. This leads him to become overconfident. A trader's expected level of overconfidence increases in the early stages of his career. Then, with more experience, he comes to better recognize his own ability. The patterns in trading volume, expected profits, price volatility, and expected prices resulting from this endogenous overconfidence are analyzed.

It is a common feature of human existence that we constantly learn about our own abilities by observing the consequences of our actions. For most people there is an attribution bias to this learning: we tend to overestimate the degree to which we are responsible for our own successes [Wolosin, Sherman, and Till (1973), Langer and Roth (1975), Miller and Ross (1975)]. As Hastorf, Schneider, and Polifka (1970) write, "We are prone to attribute success to our own dispositions and failure to external forces."

In this article we develop a multiperiod market model describing both the process by which traders learn about their ability and how a bias in this learning can create overconfident traders. Traders in our model initially do not know their ability. They learn about their ability through experience. Traders who successfully forecast next period dividends improperly update their beliefs; they overweight the possibility that their success was due to superior ability. In so doing they become overconfident.

In our model, a trader's level of overconfidence changes dynamically with his successes and failures. A trader is not overconfident when he begins to trade. Ex ante, his expected overconfidence increases over his first several trading periods and then declines. Thus the greatest overconfidence in a

The authors would like to thank Franklin Allen, Roger Edelen, Gary Gorton, Hayne Leland, Andrew Lo, Ananth Madhavan, Maureen O'Hara, David Romer, Matthew Spiegel, Brett Trueman, Greg Willard, and two anonymous referees, as well as seminar participants at l'Ecole des Hautes Etudes Commerciales (Montréal), Rutgers University, the University of Texas at Austin, the 1998 meetings of the American Finance Association, and the UCLA conference on the "Market Efficiency Debate" for their comments and suggestions. All remaining errors are the authors' responsibility. Address correspondence to Simon Gervais, Finance Department, Wharton School, University of Pennsylvania, Steinberg Hall-Dietrich Hall, Suite 2300, Philadelphia, PA 19104-6367.

trader's life span comes early in his career. After this he tends to develop a progressively more realistic assessment of his abilities as he ages.

One criticism of models of nonrational behavior is that nonrational traders will underperform rational traders and eventually be driven to the margins of markets, if not out of them altogether [Alchian (1950), Friedman (1953), and more recently, Blume and Easley (1982, 1992), Luo (1998)]. This is, however, not always the case. De Long et al. (1990) present an overlapping generations model in which nonrational traders earn higher expected profits than rational traders by bearing a disproportionate amount of the risk that they themselves create. In our model, the most overconfident and nonrational traders are not the poorest traders. For any given level of learning bias and trading experience, it is successful traders, though not necessarily the most successful traders, who are the most overconfident. Overconfidence does not make traders wealthy, but the process of becoming wealthy *can* make traders overconfident.

A large literature demonstrates that people are usually overconfident and that, in particular, they are overconfident about the precision of their knowledge.¹ De Long et al. (1991), Kyle and Wang (1997), Wang (1997), Benos (1998), and Odean (1998) examine models with statically overconfident traders.² Daniel, Hirshleifer, and Subrahmanyam (1998) look at trader overconfidence in a dynamic model. Our article differs from theirs in that we concentrate on the dynamics by which self-serving attribution bias engenders overconfidence in traders, and not on the joint distribution of trader ability and the risky security's final payoff. Our approach also has the advantage of being analytically tractable.

In our model, overconfidence is determined endogenously and changes dynamically over a trader's life. This enables us to make predictions about when a trader is most likely to be overconfident (when he is inexperienced and successful) and how overconfidence will change during a trader's life (it will, on average, increase early in a trader's career and then gradually decrease). The model also has implications for changing market conditions. For example, most equity market participants have long positions and benefit from upward price movements. We would therefore expect aggregate overconfidence to be higher after market gains and lower after market losses. Since, as we show, greater overconfidence leads to greater trading volume, this suggests that trading volume will be greater after market gains and lower after market losses. Indeed, Statman and Thorley (1998) find that this is the case.

¹ See, for example, Alpert and Raiffa (1982) and Lichtenstein, Fischhoff, and Phillips (1982). Odean (1998) provides an overview of this literature.

² Rabin and Schrag (1999) develop a model of confirmatory bias, the tendency to interpret new information as confirming one's previous beliefs. Inasmuch as people tend to have positive self-images, confirming bias and self-serving attribution bias are related. Our article differs considerably from Rabin and Schrag's in that we analyze the effect of attribution bias on the overconfidence of traders in financial markets.

The rest of this article is organized as follows. In Section 1, we introduce a one-security multiperiod economy with one insider, one liquidity trader, and one market maker. In Section 2, we develop the conditions under which there is a unique linear equilibrium in our economy. This linear equilibrium is used in Section 3 to analyze the effects of the insider's learning bias on his overconfidence and profits, as well as on the market's trading volume, volatility, and price patterns. Section 4 discusses the empirical implications of the model. Section 5 concludes. All the proofs are contained in Appendix A.

1. The Economy

We study a multiperiod economy in which only one risky asset is traded among three market participants: an informed trader, a liquidity trader, and a market maker. At the end of period t , the risky asset pays off a dividend \hat{v}_t , unknown to all the market participants at the beginning of the period.³

At the beginning of each period t , the risk-neutral informed trader (also called the *insider*) observes a signal $\hat{\theta}_t$ which is correlated with \hat{v}_t . The signal $\hat{\theta}_t$ is given by $\hat{\theta}_t = \hat{\delta}_t \hat{v}_t + (1 - \hat{\delta}_t) \hat{\varepsilon}_t$, where $\hat{\varepsilon}_t$ has the same distribution as \hat{v}_t , but is independent from it. The variable $\hat{\delta}_t$ takes the values 0 or 1. Since $\hat{\varepsilon}_t$ is independent from \hat{v}_t , the insider's information will only be useful when $\hat{\delta}_t$ is equal to one. We assume that this will happen with probability \hat{a} , where \hat{a} is the insider's *ability*. We assume that nobody (including the insider himself) knows the insider's ability \hat{a} at the outset. Instead, we assume that a priori the insider's ability is high ($\hat{a} = H$) with probability ϕ_0 and low ($\hat{a} = L$) with probability $1 - \phi_0$, where $0 < L < H < 1$ and $0 < \phi_0 < 1$. Of course, since the security dividend \hat{v}_t is announced at the end of every period t , the insider will know at the end of every period whether his information for that period was real ($\hat{\delta}_t = 1$) or was just pure noise ($\hat{\delta}_t = 0$).⁴ For tractability reasons, we also assume that the market maker observes $\hat{\theta}_t$ at the end of period t , so that his information at the end of every period is the same as the insider's. This information will be useful to both the insider and the market maker in assessing the insider's ability.

Our model seeks to describe the behavior of an informed trader with a learning bias. In particular, we want to model the phenomenon that traders usually think too much of their ability when they have been successful at predicting the market in the past. In statistical terms, this will mean that traders update their ability beliefs too much when they are right. Before formally including this behavior into our model, we describe how a rational/unbiased insider would react to the information he gathers from past trading rounds.

Let \hat{s}_t denote the number of times that the insider's information was real in the first t periods, that is, $\hat{s}_t = \sum_{u=1}^t \hat{\delta}_u$. It can be shown, using Bayes' rule,

³ Throughout the whole article, we use a "hat" over a variable to denote the fact that it is a random variable.

⁴ This will be the case since $\hat{\varepsilon}_t = \hat{v}_t$ happens with zero probability with the continuous distributions that we will specify later.

that at the end of t periods, a rational insider's updated beliefs about his own ability will be given by

$$\phi_t(s) \equiv \Pr\{\hat{a} = H \mid \hat{s}_t = s\} = \frac{H^s(1-H)^{t-s}\phi_0}{H^s(1-H)^{t-s}\phi_0 + L^s(1-L)^{t-s}(1-\phi_0)}. \quad (1)$$

We denote this rational insider's updated expected ability by

$$\mu_t(s) \equiv E[\hat{a} \mid \hat{s}_t = s] = H\phi_t(s) + L[1 - \phi_t(s)]. \quad (2)$$

Since we do not assume any kind of irrational behavior on the part of the market maker, and since the market maker's information set is the same as the insider's at the *end* of every period, this will be the market maker's updated belief at the end of period t .

In modeling the self-serving attribution bias (which we simply refer to as the *learning bias* from now on), we assume that a trader who successfully forecasts a dividend weights this success too heavily when applying Bayes' rule to assess his own ability. In choosing our updating rule we seek to accurately model the behavior psychologists have observed, to create a simple, tractable model, and to choose an updating rule that departs gradually from rational updating. Such a rule can then be used to describe traders with different degrees of bias, including unbiased traders. Psychologists find that when people succeed, they are prone to believe that success was due to their personal abilities rather than to chance or outside factors; when they fail, they tend to attribute their failure to chance and outside factors rather than to their lack of ability. They also find that "self-enhancing attributions for success are more common than self-protective attributions for failure" [Fiske and Taylor (1991); also see Miller and Ross (1975)]. This observed behavior can be modeled by assuming that, when a trader applies Bayes' rule to update his belief about his ability, he overweights his successes, he underweights his failures, and he overweights successes more than he underweights failures. The model is simpler and the qualitative results unchanged if one simply assumes, as we do, that successes are weighted too heavily and failures are weighted correctly.

More precisely, we assume that when evaluating his own ability, the insider overweights his successes at predicting the security's dividend by a learning bias factor $\gamma \geq 1$, where $\gamma = 1$ represents a rational insider. For example, at the end of the first period, if the insider finds that $\hat{\theta}_1 = \hat{v}_1$, the insider will adjust his beliefs to

$$\bar{\phi}_1(1) \equiv \Pr_b\{\hat{a} = H \mid \hat{s}_1 = 1\} = \frac{\gamma H\phi_0}{\gamma H\phi_0 + L(1-\phi_0)}, \quad (3)$$

where the subscript to "Pr" denotes the fact that the probability is calculated by a biased insider. This updated probability is larger than that of a rational insider, that is, $\bar{\phi}_1(1) \geq \phi_1(1)$. Also, as can be seen from Equation (3),

$\bar{\phi}_1(1)$ will be higher the larger γ is, and $\bar{\phi}_1(1) \rightarrow 1$ as $\gamma \rightarrow \infty$; in other words, the learning bias dictates by how much the insider adjusts his beliefs toward being a high ability insider. Moreover, our model departs continuously from rationality in the sense that $\bar{\phi}_1(1) \rightarrow \phi_1(1)$ as $\gamma \rightarrow 1$. It is easily shown that, in this case,

$$\bar{\phi}_t(s) \equiv \Pr_b\{\hat{a} = H \mid \hat{s}_t = s\} = \frac{(\gamma H)^s (1-H)^{t-s} \phi_0}{(\gamma H)^s (1-H)^{t-s} \phi_0 + L^s (1-L)^{t-s} (1-\phi_0)}, \quad (4)$$

and the (biased) insider's updated expected ability is given by

$$\bar{\mu}_t(s) \equiv E_b[\hat{a} \mid \hat{s}_t = s] = H \bar{\phi}_t(s) + L[1 - \bar{\phi}_t(s)]. \quad (5)$$

Let $\hat{\pi}_t$ denote the insider's profits in period t . At the beginning of the period, the risk-neutral insider observes his signal $\hat{\theta}_t$; he then chooses his demand for the risky security in order to maximize his expected period t profits,⁵ conditional on both his signal and his ability beliefs $\bar{\mu}_{t-1}(\hat{s}_{t-1})$ at that time. We denote this demand by $\hat{x}_t = X_t(\hat{\theta}_t, \hat{s}_{t-1})$.

The other trader in the economy trades for liquidity purposes in every period. This *liquidity trader's* demand in period t is given exogenously by the random variable \hat{z}_t . Both orders, \hat{x}_t and \hat{z}_t , are sent to a market maker who fills the orders. As in Kyle (1985), we assume that the market maker is risk neutral and competitive, and will therefore set prices so as to make zero expected profits. So if we denote the total order flow coming to the market maker in period t by $\hat{\omega}_t = \hat{x}_t + \hat{z}_t$, the market maker will set the security's price equal to

$$\hat{p}_t = P_t(\hat{\omega}_t, \hat{s}_{t-1}) \equiv E[\hat{v}_t \mid \hat{\omega}_t, \hat{s}_{t-1}]$$

in period t . An equilibrium to our model is defined as a sequence of pairs of functions (X_t, P_t) , $t = 1, 2, \dots$, such that the insider's demand X_t in period t maximizes his expected profits (according to his own beliefs) for that period given that he faces a price curve P_t , while the market maker is expecting zero profit in that period.

As will become obvious later, the main results of this article are driven by the insider's updating dynamics. As such, the liquidity trader and the market maker play only a minimal role in this model. In fact, their role is essentially one of market clearing. The presence of the liquidity trader introduces noise that will prevent the "no trade equilibrium" described by Milgrom and Stokey (1982) from occurring. The competitive market maker assumption is simply made out of convenience; the presence of a risk-neutral rational trader would serve a similar purpose, as in Daniel, Hirshleifer, and Subrahmanyam (1998). The model can be extended to eliminate the presence

⁵ Since both the risky dividend and the insider's signal are announced at the end of every period, this objective is consistent with the insider maximizing long-term profits.

of the liquidity trader and the market maker, to introduce a monopolist specialist, and to increase the number of market participants in an overlapping generations setting. These alternative specifications do not affect the nature of the insider's updating and, therefore, do not change the main results of the article. Details are available from the authors upon request.

2. A Linear Equilibrium

In this section we show that, when \hat{v}_t , $\hat{\varepsilon}_t$, and \hat{z}_t are jointly and independently normal, there is a linear equilibrium to our economy. We assume that

$$\begin{bmatrix} \hat{v}_t \\ \hat{\varepsilon}_t \\ \hat{z}_t \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Omega \end{bmatrix} \right), \quad t = 1, 2, \dots, \quad (6)$$

and that each such vector is independent of all the others. Note that it is crucial that $\text{var}(\hat{v}_t) = \text{var}(\hat{\varepsilon}_t)$, since we do not want the size of $\hat{\theta}_t$ to reveal anything about the likelihood that $\hat{\delta}_t = 1$ until \hat{v}_t is announced. In other words, ability updating is only possible when both $\hat{\theta}_t$ and \hat{v}_t become observable. Also note that, as in Kyle (1985), this normality assumption implies that dividends and prices can take on negative values in this model; a positive mean for \hat{v}_t and $\hat{\varepsilon}_t$ would reduce the likelihood of such occurrences, but would not affect any of our results.

Let us conjecture that, in equilibrium, the function $X_t(\theta, s)$ is linear in θ , and that the function $P_t(\omega, s)$ is linear in ω :

$$X_t(\theta, s) = \beta_t(s) \theta, \quad (7a)$$

$$P_t(\omega, s) = \lambda_t(s) \omega. \quad (7b)$$

Our objective is to find $\beta_t(s)$ and $\lambda_t(s)$ which are consistent with this conjecture. We start with the following result.

Lemma 1. *Assume that a linear equilibrium exists in period t ; that is, assume that Equations (7a) and (7b) hold. Then, in period t , the insider's demand for the risky security is given by*

$$\hat{x}_t = \frac{\bar{\mu}_{t-1}(\hat{s}_{t-1})\hat{\theta}_t}{2\lambda_t(\hat{s}_{t-1})}, \quad (8)$$

and the market maker's price schedule is given by

$$\hat{p}_t = \frac{\mu_{t-1}(\hat{s}_{t-1})\beta_t(\hat{s}_{t-1})\Sigma}{\beta_t^2(\hat{s}_{t-1})\Sigma + \Omega} \hat{\omega}_t. \quad (9)$$

This lemma establishes that we can indeed write $\hat{x}_t = \beta_t(\hat{s}_{t-1})\hat{\theta}_t$ and $\hat{p}_t = \lambda_t(\hat{s}_{t-1})\hat{\omega}_t$ as long as

$$\beta_t(s) = \frac{\bar{\mu}_{t-1}(s)}{2\lambda_t(s)}, \quad \text{and} \quad (10)$$

$$\lambda_t(s) = \frac{\mu_{t-1}(s)\beta_t(s)\Sigma}{\beta_t^2(s)\Sigma + \Omega}. \quad (11)$$

However, the result relies on the assumption that a linear equilibrium exists. It turns out that this assumption is not always satisfied given the insider's learning bias. In fact, the following lemma derives the exact condition under which such an equilibrium will exist in a given period t .

Lemma 2. *In any given period t , there exists a linear equilibrium of the form conjectured in Equations (7a) and (7b) if and only if $\bar{\mu}_{t-1}(\hat{s}_{t-1}) \leq 2\mu_{t-1}(\hat{s}_{t-1})$.*

This condition states that, for an equilibrium to exist, the biased insider's beliefs about his ability cannot exceed those of the rational market maker by too much. It effectively ensures that the insider will on average be making profits, even though he may not be optimizing correctly due to his biased beliefs.⁶ Were this condition not satisfied, the market maker would know that the insider would take a position resulting on average in negative profits. Our condition that the competitive market maker quotes a price schedule earning him on average zero profits could then never be satisfied. This would result in a market breakdown. The following lemma derives a condition that is both necessary and sufficient to always avoid such outcomes.

Lemma 3. *A necessary and sufficient condition for $\bar{\mu}_{t-1}(\hat{s}_{t-1}) > 2\mu_{t-1}(\hat{s}_{t-1})$ to be avoided for any history up to any period t and any $\gamma > 1$ is that $H \leq 2L$.*

Note that, for a given *fixed* value of γ , $H \leq 2L$ is too strong a condition to avoid market breakdowns for any history; that is, the condition is then sufficient but not necessary. However, since market breakdowns are outside the scope of this article [see, e.g., Bhattacharya and Spiegel (1991)], we assume that $H \leq 2L$ is always satisfied in the rest of our analysis. Such an assumption allows us to vary γ throughout the article without having to worry about the existence of an equilibrium, but does not affect the qualitative aspects of our results. We finish this section with the following characterization of the equilibrium.

⁶ This condition does not violate the absence of arbitrage results established by Bossaerts (1999). Although the insider's learning bias could introduce profit opportunities for other informed traders (were they present in the economy), they do not lead to *riskless* arbitrage opportunities.

Proposition 1. *Assuming that $H \leq 2L$, there is always a unique linear equilibrium to the economy described in Section 1. In this equilibrium, the insider's demand and the market-maker's price schedule are given by Equations (7a) and (7b) with*

$$\beta_t(s) = \sqrt{\frac{\Omega}{\Sigma} \frac{\bar{\mu}_{t-1}(s)}{2\mu_{t-1}(s) - \bar{\mu}_{t-1}(s)}}, \quad \text{and} \quad (12a)$$

$$\lambda_t(s) = \frac{1}{2} \sqrt{\frac{\Sigma}{\Omega} \bar{\mu}_{t-1}(s) [2\mu_{t-1}(s) - \bar{\mu}_{t-1}(s)]}. \quad (12b)$$

3. Properties of the Model

In this section we analyze the effects of the insider's learning bias on the properties and dynamics of the economy in equilibrium. We introduce a measure of overconfidence analogous to that found in static models and show that the learning bias results in dynamically evolving insider overconfidence. We then look at the effect of this changing overconfidence on trading volume, trader profits, price volatility, as well as expected price patterns.

3.1 Convergence

If this financial market game were played to infinity, we would expect both the insider and the market maker to eventually learn the exact ability \hat{a} of the insider. This would be true for a rational insider ($\gamma = 1$). However, since our insider learns his ability with a personal bias, this result is not immediate; in fact, as the following proposition shows, it is not true for a highly biased insider.

Proposition 2. *When $\hat{a} = H$, the updated posteriors of the insider $\bar{\phi}_t(\hat{s}_t)$ will converge to 1 almost surely as $t \rightarrow \infty$. When $\hat{a} = L$, the updated posteriors of the insider $\bar{\phi}_t(\hat{s}_t)$ will converge as follows:*

$$\bar{\phi}_t(\hat{s}_t) \xrightarrow{a.s.} \begin{cases} 0, & \text{if } \gamma < \gamma^* \\ \phi_0, & \text{if } \gamma = \gamma^* \\ 1, & \text{if } \gamma > \gamma^*, \end{cases}$$

where $\gamma^* = \frac{L}{H} \left(\frac{1-L}{1-H} \right)^{(1-L)/L}$.

So both the insider and the market maker will eventually learn the insider's ability precisely when it is high, and the market maker will always learn the insider's ability when it is low, but the insider will only do so if his learning bias is not too large. One implication of this lemma is that a low-ability

insider whose learning bias is sufficiently extreme may never acknowledge his low ability, no matter how much experience he has.⁷

3.2 Patterns of overconfidence

As we showed in Section 3.1, the insider eventually learns his own ability, provided that his learning is not too biased (i.e., provided that $\gamma < \gamma^*$). So when the insider's ability is low, the insider eventually comes to his senses and recognizes the fact that he is a low-ability insider. However, it is always the case that the insider thinks too highly of himself relative to an otherwise identical unbiased insider. This section introduces a measure for this discrepancy; we call it the insider's *overconfidence*. The evolution of the insider's overconfidence throughout his life is central to our study, as this overconfidence measures by how much our model departs from a purely rational setup in any given period.

In our model, an insider is considered very overconfident at the end of a given period t if his updated expected ability at that time $[\bar{\mu}_t(\hat{s}_t)]$ is large compared to the updated expected ability that a rational insider would have reached with the same past history of successes and failures $[\mu_t(\hat{s}_t)]$. To measure the insider's overconfidence at the end of t periods, we therefore define the random variable

$$\hat{\kappa}_t \equiv K_t(\hat{s}_t) \equiv \frac{\bar{\mu}_t(\hat{s}_t)}{\mu_t(\hat{s}_t)}. \quad (13)$$

Of course, when the insider is rational ($\gamma = 1$), the numerator is exactly equal to the denominator of this expression, and $\hat{\kappa}_t = 1$ for all t . On the other hand, when the insider's learning is biased ($\gamma > 1$), we have $\bar{\mu}_t(\hat{s}_t) \geq \mu_t(\hat{s}_t)$, and $\hat{\kappa}_t \geq 1$ for all t . As the next proposition shows, the insider's overconfidence in period t is greater when the insider's learning bias is large. In other words, the insider's overconfidence is directly attributable to his learning bias.

Proposition 3. *The function $K_t(s)$ defined in Equation (13) is increasing in γ .*

Since the insider's overconfidence results from his learning bias when he is successful, it is tempting to conclude that the more successful an insider is, the more overconfident he will be. As we next show, this intuition is wrong.

First, since the insider updates his beliefs incorrectly only after successful predictions, it is always true that $\bar{\mu}_t(0) = \mu_t(0)$, and therefore $K_t(0) = 1$. However, as soon as the insider successfully predicts one risky dividend,

⁷ Bossaerts (1999) shows that prior beliefs that are correctly updated using Bayes' law converge to the right posterior beliefs, whether the priors are biased or not. In contrast, our result shows that correct priors updated incorrectly may not lead to the correct posteriors.

his learning bias makes him overconfident,⁸ and $\bar{\mu}_t(1) > \mu_t(1)$, so that $K_t(1) > 1$. So it is always true that the insider's first successful prediction makes him overconfident. However, it is not always the case that an additional successful prediction always makes the insider more overconfident.

To see this, suppose that we are at the end of the second period. The insider will then have been successful 0, 1, or 2 times. We already know that $K_2(1) > K_2(0) = 1$ for any value of the insider's learning bias parameter γ . Now, suppose that γ is large. This means that if the insider is successful in the first period, he will immediately (and perhaps falsely) jump to the conclusion that he is a high ability insider, that is, $\bar{\mu}_1(1)$ is close to H . Since this one successful period has already convinced the insider that his ability is high, the second period results will not have much of an effect on his beliefs, whether he is successful or not in that period, that is, $\bar{\mu}_2(2)$ is close to $\bar{\mu}_2(1)$. On the other hand, if the insider had been rational ($\gamma = 1$), he would have adjusted his expected ability beliefs more gradually. In particular, after a first-period success, a rational insider does not automatically conclude that his ability is high. Instead, he adjusts his posterior expected ability beliefs toward H , and uses the second-period result to further adjust these beliefs: upward toward H if he is successful, and downward toward L otherwise. As a result, $\mu_2(2)$ will be somewhat larger than $\mu_2(1)$. Therefore, since $\bar{\mu}_2(2) \approx \bar{\mu}_2(1)$ and $\mu_2(2) > \mu_2(1)$, we have

$$K_2(2) \equiv \frac{\bar{\mu}_2(2)}{\mu_2(2)} < \frac{\bar{\mu}_2(1)}{\mu_2(1)} \equiv K_2(1),$$

and we see that $K_2(s)$ decreases when s goes from 1 to 2. The following proposition describes this phenomenon in more details.

Proposition 4. *The function $K_t(s)$ defined in Equation (13) is increasing over $s \in \{0, \dots, s_t^*\}$ and decreasing over $s \in \{s_t^*, \dots, t\}$, for some $s_t^* \in \{1, \dots, t\}$.*

We now turn to how the insider's overconfidence is expected to behave over time. To do this we describe the ex ante expected overconfidence level of the insider, $E[\hat{\kappa}_t]$, through time. This is done in Figure 1 for different values of γ . When γ is relatively small ($\gamma < \gamma^*$), the insider will on average be overconfident at first but, over time, will converge to rational behavior. This can be explained as follows. Over a small number of trading periods, a trader's success rate may greatly exceed that predicted by his ability.

⁸ In our model, traders are not overconfident when they begin to trade. It is through making forecasts and trading that they become overconfident. This leads market participants to be, on average, overconfident. In real markets, selection bias may cause even beginning traders to be overconfident. Indeed, since not everyone trades, it is likely that people who rate their own trading abilities most highly will seek jobs as traders or will trade actively on their own account. Those with actual high ability and those with high overconfidence will rate their own ability highest. Thus, even at the entry level, we would expect to find overconfident traders.

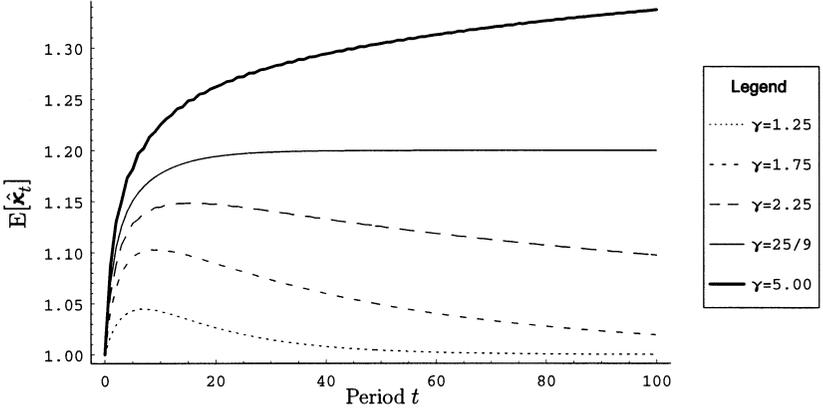


Figure 1
Expected evolution of overconfidence

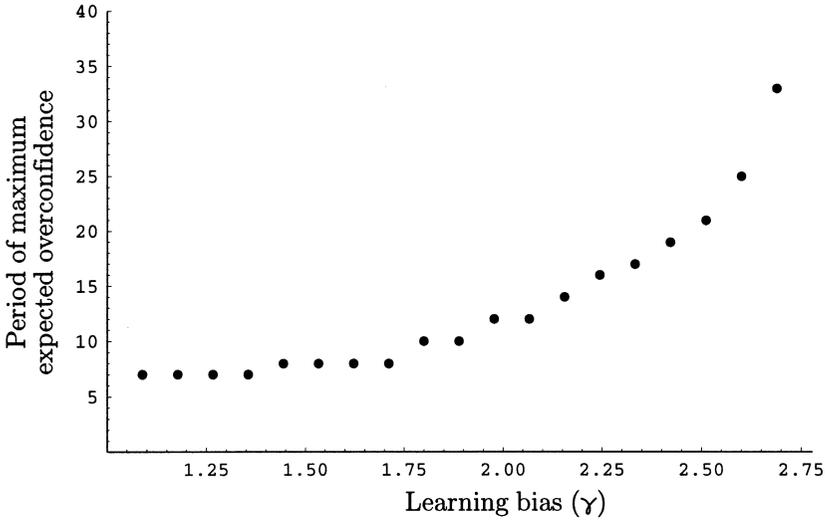
Ex ante expected patterns in the level of overconfidence of the insider over time. The level \hat{k}_t of insider overconfidence at the end of any period t is measured as a ratio of the biased insider's expected ability over that of an otherwise unbiased insider, $\hat{k}_t = \bar{\mu}_t(\hat{s}_t)/\mu_t(\hat{s}_t)$, where \hat{s}_t denotes the number of successful predictions by the insider in the first t periods. This figure plots the ex ante expected value of \hat{k}_t as a function of the time period using $H = 0.9$, $L = 0.5$, $\phi_0 = 0.5$, $\Sigma = \Omega = 1$. Each line was obtained with a different value of the insider's learning bias γ shown in the legend. Note that, with these parameter values, the insider eventually learns his ability with probability one (zero) if his bias parameter γ is smaller (greater) than $\gamma^* = 25/9 \approx 2.78$.

Very successful traders will overestimate the likelihood that success is due to ability rather than luck. But over many trading periods a trader's success rate is likely to be close to that predicted by his ability. Only those traders with extreme learning bias (or with very unlikely success patterns) will fail to recognize their true ability. Indeed, as γ increases, the insider tends to put more and more weight on his past successes, and so takes a little more time to find his correct ability. However, if γ is too large ($\gamma > \gamma^*$), it is possible that the insider puts so much weight on his past successes in the stock market that he never recognizes his correct ability. It can be shown that $E[\hat{k}_t]$ then converges to $\phi_0 + (1 - \phi_0)\frac{H}{L}$.

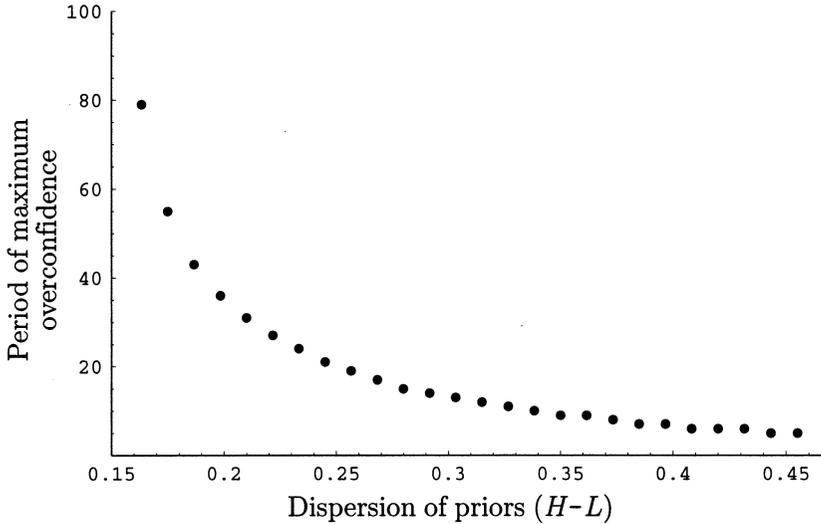
Thus our model predicts that more inexperienced traders will be more overconfident than experienced traders. Less experienced traders are more likely to have success records which are unrepresentative of their abilities. For some, this will lead to overconfidence. By the law of large numbers, older traders are likely to have success records which are more representative of their abilities; they will, on average, have more realistic self-assessments.

Given Figure 1, it is natural to ask what factors determine the point at which a trader's overconfidence is likely to peak. All other things being equal, the greater a trader's learning bias, γ , the longer it is likely to take for his overconfidence level to a peak. Figure 2a illustrates this effect.

In addition to the degree of learning bias, how quickly a trader's overconfidence peaks (and how quickly he ultimately learns his true ability) depends



(a) Period of maximum overconfidence as a function of bias.



(b) Period of maximum overconfidence as a function of priors dispersion.

Figure 2

Period of maximum overconfidence

The level $\hat{\kappa}_t$ of insider overconfidence at the end of any period t is measured as a ratio of the biased insider's expected ability over that of an otherwise unbiased insider, $\hat{\kappa}_t = \bar{\mu}_t(\hat{s}_t)/\mu_t(\hat{s}_t)$, where \hat{s}_t denotes the number of successful predictions by the insider in the first t periods. These figures plot the period of maximum expected overconfidence, $\text{argmax}_t E[\hat{\kappa}_t]$, as a function of (a) the insider's learning bias γ ; (b) the dispersion $H - L$ of the insider's prior ability beliefs. Figure (a) was obtained with $H = 0.9$, $L = 0.5$, and $\phi_0 = 0.5$. Figure (b) was obtained with $\phi_0 = 0.5$, $\gamma = 1.1$, and keeping the insider's ex ante expected ability $\phi_0 H + (1 - \phi_0)L$ constant at 0.7.

on the frequency, speed, and clarity of the feedback he receives. A trader who receives frequent, immediate, and clear feedback will, on average, peak in overconfidence early and quickly realize his true ability. One who receives infrequent, delayed, and ambiguous feedback will peak in overconfidence later and more slowly realize his true ability. In general, financial markets are difficult environments for learning, as feedback is often ambiguous and comes well after a decision was made. We would expect those who trade most frequently, such as professional traders, and those who keep careful records rather than relying on memory, to learn most quickly. In our model a trader receives immediate feedback each period. The greater the difference between the high (H) and the low (L) ability levels, the clearer the feedback. Figure 2b illustrates how long it takes on average for the insider's overconfidence to peak as a function of $H - L$.

3.3 Effects of the insider's learning bias

Section 3.2 shows that the insider's learning bias has a dynamic impact on his beliefs about his ability and on the way he interprets future private information. This in turn affects the future trading process. In this and the next section, we describe how this trading process, as measured by trading volume, trader profits, and price volatility, is affected. Let $\hat{\psi}_t$ denote the trading volume in period t . Since this trading volume comes from both the insider and the liquidity trader, it is given by $\hat{\psi}_t \equiv \frac{1}{2} (|\hat{x}_t| + |\hat{z}_t|)$.

Lemma 4. *Conditional on the insider having been successful s times in the first t periods (i.e., conditional on $\hat{s}_t = s$), the expected volume, the expected insider profits, and the price variance in period $t+1$ are respectively given by*

$$\mathbb{E} \left[\hat{\psi}_{t+1} \mid \hat{s}_t = s \right] = \frac{1}{\sqrt{2\pi}} \left[\sqrt{\Sigma} \beta_{t+1}(s) + \sqrt{\Omega} \right], \quad (14)$$

$$\mathbb{E} \left[\hat{\pi}_{t+1} \mid \hat{s}_t = s \right] = \frac{1}{2} \sqrt{\Sigma \Omega} \sqrt{\bar{\mu}_t(s) [2\mu_t(s) - \bar{\mu}_t(s)]}, \quad \text{and} \quad (15)$$

$$\text{Var} \left[\hat{p}_{t+1} \mid \hat{s}_t = s \right] = \frac{\Sigma}{2} \bar{\mu}_t(s) \mu_t(s). \quad (16)$$

Since the insider's learning bias is unaffected by his success rate and vice versa, our model allows us to analyze the effects of the learning bias on the insider's behavior and on the properties of the economy in two different ways. First, given a fixed past history of the insider's successes and failures, we can vary the size of his learning bias to get an idea of the impact of that bias. This is the focus of the current section. Second, we can fix the insider's learning bias and determine the effects of different trading histories on the insider and the economy in general. We will turn to this in Section 3.4.

Recall from Equation (7a) in Section 2 that the insider will multiply his period $t+1$ signal, $\hat{\theta}_{t+1}$, by $\beta_{t+1}(s)$ to obtain his demand for the risky asset in

that period. In other words, $\beta_{t+1}(s)$ represents the insider's trading intensity in period $t + 1$ after having been successful s times in the first t periods. As Equation (14) shows, greater average insider intensity leads to larger expected volume.

Moreover, as noted in Section 3.2, a biased insider, who has had at least one success, is always overconfident. In other words, the insider thinks that his signal $\hat{\theta}_{t+1}$ in period $t + 1$ is more informative than it really is. This leads him to use his information more aggressively than he should and results in higher expected trading volume in the risky security. As the next proposition demonstrates, the greater the learning bias the greater this trading.

Proposition 5. *Given that $\hat{s}_t = s$, the expected volume in period $t + 1$ is increasing in the insider's learning bias parameter γ .*

Notice that we can rewrite $\beta_{t+1}(s)$ given in Equation (12a) as

$$\beta_{t+1}(s) = \sqrt{\frac{\Omega}{\Sigma} \left[\frac{2}{K_t(s)} - 1 \right]^{-1}}. \quad (17)$$

This tells us that the trading intensity $\beta_{t+1}(s)$ of the insider in period $t + 1$ is a monotonically increasing function of the insider's overconfidence $K_t(s)$ after t periods. Since we showed in Proposition 3 that the insider's overconfidence in any period is increasing in γ , it is natural to find that expected volume in a particular period will also be increasing in γ .

We know that the biased insider trades too aggressively on his information; in other words, the insider's learning bias makes him act suboptimally. It is therefore not surprising that the insider's expected profits in any given period are decreasing in his learning bias parameter γ .

Proposition 6. *Given that $\hat{s}_t = s$, the expected insider profits in period $t + 1$ are decreasing in the insider's learning bias parameter γ .*

The more overconfident the insider, the more he trades in response to any given signal. This increases his expected trading relative to that of the liquidity trader. Therefore, the signal-to-noise ratio in total order flow increases and the market maker is able to make better inferences about the insider's signal. The market maker is then able to set prices that vary more in response to $\hat{\theta}_t$ and are closer to the expected dividend conditional on the insider's signal ($E[\hat{v}_t \mid \hat{\theta}_t]$) and further from its unconditional expectation (zero). This increases price volatility.

Proposition 7. *Given that $\hat{s}_t = s$, the expected price volatility in period $t + 1$, as measured by the price's variance, is increasing in the insider's learning bias parameter γ .*

3.4 Effects of the insider's past performance

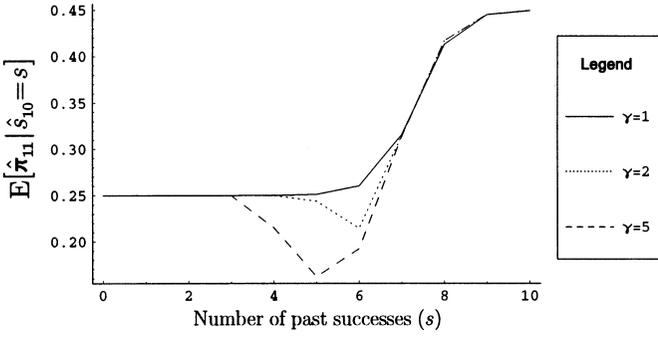
The effects described in Section 3.3 are static in the sense that they do not depend on the insider's past performance. Given any success history, the next period's trading volume and price volatility are expected to be larger and the next period's insider profits are expected to be lower when γ is large. These results are analogous to the results documented by Odean (1998), who shows that trader overconfidence has these effects in a one-period economy. Since, as documented in Section 3.2, the insider's learning bias eventually makes him overconfident, our results are natural extensions of that study's static results. However, we also know from Section 3.2 that the insider's overconfidence changes dynamically with his past performance. In this section we look at how, for an insider with a particular learning bias, the economy is affected by that insider's past performance.

The monotonic relationship between $\beta_{t+1}(s)$ and $K_t(s)$ described in Equation (17) also helps us characterize the conditional expected volume in a particular period $t + 1$. For example, it is natural to find that the expected one-period volume given s insider successes in the first t periods has the same shape as $K_t(s)$ as a function of s .

Proposition 8. *The expected volume in period $t + 1$, conditional on the insider having been successful s times in the first t periods (i.e., given $\hat{s}_t = s$), is increasing over $s \in \{0, \dots, s_t^\circ\}$ and decreasing over $s \in \{s_t^\circ, \dots, t\}$, for some $s_t^\circ \in \{1, \dots, t\}$.*

We next show that a learning bias can cause a successful insider's expected future profits to be smaller than a less successful insider's. This is because two forces affect an insider's expected future profits: his overconfidence and his expected ability. To disentangle these two forces, let us describe the insider's expected profits in period $t + 1$ after he has been successful s times in the first t periods. We know from Section 3.2 that the insider's overconfidence at the end of t periods is at a minimum of 1 when $s = 0$. We also know from Proposition 4 that overconfidence increases with the number s of past successful dividend predictions (up to s_t^*). This means that the insider's decision in period $t + 1$ will be more and more distorted as s increases. At the same time, as s increases, it becomes increasingly likely that the insider's ability is high, though not as likely as the insider thinks. A biased insider who becomes sufficiently overconfident may act so suboptimally that he more than offsets the potential increase in expected profits coming from his probable higher ability. As the insider's overconfidence comes back down ($s > \hat{s}_t^*$), successes decrease the insider's overconfidence while increasing his expected ability. Thus both forces lead to additional expected future profits.

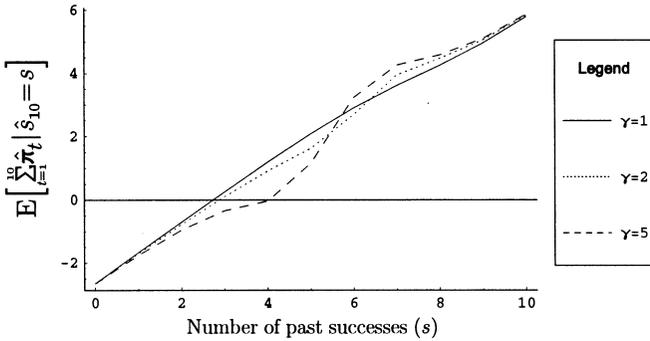
Figures 3a and b illustrate how the insider's overconfidence and expected ability counterbalance each other. In Figure 3a, we look at the insider's expected profits in period 11 as a function of the number of successes he has had in the first 10 periods. It is clear from that figure that an unbiased



(a) Expected insider profits.



(b) Overconfidence.



(c) Expected past profits.

Figure 3
Impact of insider's past successes

These graphs show (a) the expected insider profits in period 11, (b) the insider's overconfidence, $K_{10}(s) = (\hat{\mu}_{10}(s)/\mu_{10}(s))$, at the beginning of that period, and (c) the expected profits of the insider in the first 10 periods, as functions of the number s of successful predictions by the insider in the first 10 periods. Every graph uses $H = 0.9$, $L = 0.5$, $\phi_0 = 0.5$, $\Sigma = \Omega = 1$. Each line was obtained with a different value of the insider's learning bias γ shown in the legend.

insider with an additional past success always has higher expected profits in period $t+1$, since his expected ability is higher. However, when the insider's learning is biased, it is possible that his overconfidence (which we plot in Figure 3b) prevents him from benefiting from the boost in expected ability implied by an additional success. In fact, for this example, we can see that an insider with $\gamma = 2$ or $\gamma = 5$ who has had six successes in the first 10 periods does worse than an insider who has yet to predict one dividend correctly! This simple numerical example can in fact be generalized as follows.

Proposition 9. *Given that $\hat{s}_t = s$, the expected insider profits in period $t+1$ are increasing over $s \in \{0, \dots, s'_t\}$ and $s \in \{s''_t, \dots, t\}$, but are decreasing over $s \in \{s'_t, \dots, s''_t\}$ for some $(s'_t, s''_t) \in \{1, \dots, t\}^2$ such that $s'_t \leq s''_t$.*

Since the insider, in this model, can only be perfectly right or completely wrong in any given period, the correct measure of his past performance at the beginning of period $t+1$ is the number of his past successes (\hat{s}_t). In reality, traders can be right or wrong to different extents, and so the measure that is typically used to measure their performance is past profits. It is easily shown that, in our model, expected past profits are monotonically increasing in the number of past successes. Figure 3c illustrates this for the above numerical example. Therefore, future expected insider profits, as a function of past profits, are first increasing, then decreasing, and then increasing again.

We finish this section by looking at the volatility of prices conditional on the insider's past performance. Expected volatility is not affected by the insider's successes in the same way as are overconfidence and volume. Although expected overconfidence and expected trading volume can both be nonmonotonic in the number of past insider successes, expected volatility is always increased by one more insider success. More precisely, large posteriors by the biased insider $[\bar{\mu}_t(s)]$ and the rational market maker $[\mu_t(s)]$ both contribute to more expected volatility: the former by his unwarranted aggressiveness, and the latter by his steeper price schedule.

Proposition 10. *At the end of period t , the conditional expected volatility in period $t+1$ is increasing in the number of past successful predictions by the insider in the first t periods.*

3.5 Price patterns

Daniel, Hirshleifer, and Subrahmanyam (1998) argue that a trader's overconfidence in his *long-lived* private information can result in positive (negative) price autocorrelation in the short (long) run. The information in our model is short-lived in the sense that each signal obtained by the insider is advantageous to him for one, and only one, period. As a result, consecutive market-clearing prices are independent in our model, and this implies that the returns (i.e., the price changes) are spuriously negatively autocorrelated through a

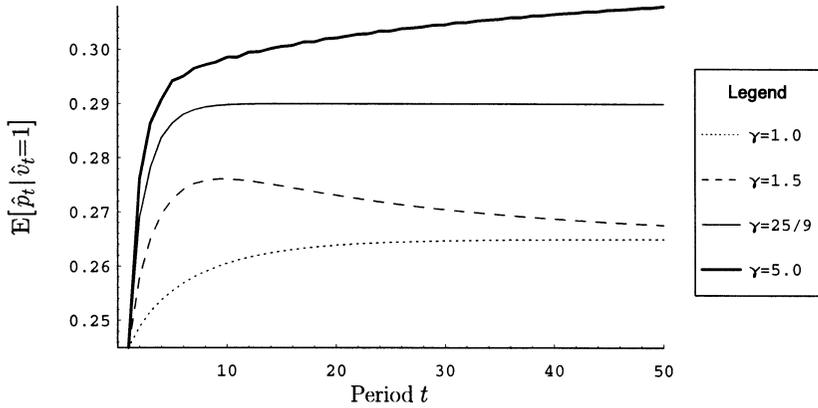


Figure 4
Expected price patterns

Expected price in period t conditional on a dividend of \$1 ($\hat{v}_t = 1$) being paid at the end of the period. The figure was obtained with $H = 0.9$, $L = 0.5$, $\phi_0 = 0.5$, and $\Sigma = \Omega = 1$. Each line was drawn with a different learning bias parameter γ , shown in the legends. Note that, with these parameter values, the insider eventually learns his ability with probability one (zero) if his bias parameter γ is smaller (larger) than $\gamma^* = 25/9 \approx 2.78$. As a result, the expected price only converges to its “unbiased” value of $(1 \cdot H^2/2)\phi_0 + (1 \cdot L^2/2)(1 - \phi_0) = 0.265$ if $\gamma < \gamma^*$; otherwise, it converges to a higher “biased” value.

bid-ask bounce, an effect originally documented by Roll (1984): $\text{cov}(\hat{p}_{t+1} - \hat{p}_t, \hat{p}_{t+2} - \hat{p}_{t+1}) = -\text{var}(\hat{p}_{t+1})$. However, as we next show, we do find, like Daniel, Hirshleifer, and Subrahmanyam (1998), that the expected evolution of prices for a given dividend size can be hump-shaped.

In an economy where the insider’s ability is known ex ante (i.e., $H = L = \mu$), the expected price in any period t given a subsequent announcement of $\hat{v}_t = v$ is constant at $E(\hat{p}_t | \hat{v}_t = v) = \frac{v\mu^2}{2}$. This is not true when the insider learns about his own ability through time (when $H > L$). First, when the insider learns his ability without a bias ($\gamma = 1$), he learns whether he should use his information more or less aggressively over time. Unconditionally, the expected price starts at $\frac{v\mu^2}{2}$ and converges to its long-run value of $\frac{vH^2}{2}\phi_0 + \frac{vL^2}{2}(1 - \phi_0)$. This is illustrated by the dotted line in Figure 4.

When the insider learns with a bias that is not too extreme ($\gamma < \gamma^*$), the insider may first push up the prices too high before learning the correct way to interpret his information. However, if the insider is so biased as to refuse to acknowledge his low ability despite persistent poor performance ($\gamma \geq \gamma^*$), the expected price will monotonically increase to its long-run biased level.

4. Discussion

Our model predicts that overconfident traders will increase their trading volume and thereby lower their expected profits. To the extent that trading is motivated by overconfidence, higher trading will correlate with lower profits. Barber and Odean (2000) find that this is true for individual investors.

While this evidence supports our model, our model does more than simply posit that investors are overconfident. We also describe a dynamic by which overconfidence may wax and wane, both on an individual level and in the aggregate (though the latter is not modeled formally). In times when aggregate success is greater than usual, overconfidence will be higher. In our model, success is measured by how well a trader forecasts dividends. This formulation allows us to present closed form solutions. In many markets, returns will be a trader's metric of success. Traders who attribute returns from general market increases to their own acumen will become overconfident and therefore trade more actively. Therefore, we would predict that periods of market increases will tend to be followed by periods of increased aggregate trading. Statman and Thorley (1998) find this is so for monthly horizons. Taking a longer view, overconfidence and its principal side effect, increased trading, are likely to rise late in a bull market and to fall late in a bear market. A bull market may also attract more investment capital, in part, because investors grow more confident in their personal investment abilities. This increase in investment capital could cause price pressures that send market prices even higher.

In our model, investors are most overconfident early in their careers. With more experience, self-assessment becomes more realistic and overconfidence more subdued. Barber and Odean (1998) find that, after controlling for gender, marital status, children, and income, younger investors trade more actively than older investors while earning lower returns relative to a buy-and-hold portfolio. These results are consistent with our prediction that overconfidence diminishes with greater experience.

5. Conclusion

We go through life learning about ourselves as well as the world around us. We assess our own abilities not so much through introspection as by observing our successes and failures. Most of us tend to take too much credit for our own successes. This leads to overconfidence. It is in this way that overconfidence develops in our model. When a trader is successful, he attributes too much of his success to his own ability and revises his beliefs about his ability upward too much. In our model overconfidence is dynamic, changing with successes and failures. Average levels of overconfidence are greatest in those who have been trading for a short time. With more experience, people develop better self-assessments.

Since it is through success that traders become overconfident, successful traders, though not necessarily the most successful traders, are most overconfident. These traders are also, as a result of success, wealthy. Overconfidence does not make traders wealthy, but the process of becoming wealthy can make them overconfident. Due to their wealth, overconfident traders are in no immediate danger of being driven out of the marketplace. As they age,

less able, overconfident traders will lose both wealth and confidence. Some may even cease to trade. However, in markets where inexperienced traders continuously enter and old traders die, there will always be overconfident traders. Furthermore, these traders will tend to control more wealth than their less confident peers. Thus, overconfident traders can play an important long-term role in financial markets.

As shown in our model, an overconfident trader trades too aggressively, and this increases expected trading volume. Volatility is increasing in a trader's number of past successes. Both volume and volatility increase with the degree of a trader's learning bias. Overconfident traders behave suboptimally, thereby lowering their expected profits. A more successful trader is likely to have more information gathering ability but he may not use his information well. Thus, the expected future profits of a more successful trader may actually be lower than those of a less successful trader. Successful traders do tend to be good, but not as good as they think they are.

The principal goal of this article is to demonstrate that a simple and prevalent bias in evaluating one's own performance is sufficient to create markets in which investors are, on average, overconfident. Unlike models such as De Long et al. (1990), in which biased traders survive by earning greater profits, our model describes a market in which overconfident traders realize, on average, lower profits. Though overconfidence does not lead to greater profits, greater profits do lead to overconfidence. A particular trader's overconfidence will not flourish indefinitely; time and experience gradually rid him of it. However, in a market in which new traders are born every minute, overconfidence will flourish.

Appendix A

Proof of Lemma 1. Assume that $P_t(\hat{\omega}_t, \hat{s}_{t-1}) = \lambda_t(\hat{s}_{t-1}) \hat{\omega}_t$. This means that the insider's expected period t profits, when sending a market order of \hat{x}_t to the market maker, are given by

$$\begin{aligned} E_b[\hat{\pi}_t \mid \hat{\theta}_t, \hat{s}_{t-1}, \hat{x}_t] &= E_b \left\{ \hat{x}_t [\hat{v}_t - P_t(\hat{\omega}_t, \hat{s}_{t-1})] \mid \hat{\theta}_t, \hat{s}_{t-1}, \hat{x}_t \right\} \\ &= \hat{x}_t \left[E_b(\hat{v}_t \mid \hat{\theta}_t, \hat{s}_{t-1}) - \lambda_t(\hat{s}_{t-1}) \hat{x}_t \right]. \end{aligned} \quad (18)$$

Differentiating this last expression with respect to \hat{x}_t and setting the result equal to zero yields

$$\hat{x}_t = \frac{E_b(\hat{v}_t \mid \hat{\theta}_t, \hat{s}_{t-1})}{2\lambda_t(\hat{s}_{t-1})}. \quad (19)$$

Also, a simple use of iterated expectations and the projection theorem for normal variables shows that $E_b(\hat{v}_t \mid \hat{\theta}_t, \hat{s}_{t-1}) = \bar{\mu}_{t-1}(\hat{s}_{t-1})\hat{\theta}_t$, as in Equation (8).

Next, assume that $X_t(\hat{\theta}_t, \hat{s}_{t-1}) = \beta_t(\hat{s}_{t-1})\hat{\theta}_t$. As discussed in Section 1, the market maker's price is a function of the information he gathers from the order flow. More precisely,

$$\begin{aligned}\hat{p}_t &= \mathbb{E}[\hat{v}_t \mid \hat{\omega}_t, \hat{s}_{t-1}] = \mathbb{E}[\mathbb{E}(\hat{v}_t \mid \hat{\omega}_t, \hat{s}_{t-1}, \hat{\delta}_t) \mid \hat{\omega}_t, \hat{s}_{t-1}] \\ &= \mathbb{E}\{\hat{\delta}_t \mathbb{E}[\hat{v}_t \mid \hat{\omega}_t = \beta_t(\hat{s}_{t-1})\hat{v}_t + \hat{z}_t, \hat{s}_{t-1}] + (1 - \hat{\delta}_t) \cdot 0 \mid \hat{\omega}_t, \hat{s}_{t-1}\}.\end{aligned}\quad (20)$$

Use of the projection theorem for normal variables shows that

$$\mathbb{E}[\hat{v}_t \mid \hat{\omega}_t = \beta_t(\hat{s}_{t-1})\hat{v}_t + \hat{z}_t, \hat{s}_{t-1}] = \frac{\beta_t(\hat{s}_{t-1})\Sigma}{\beta_t^2(\hat{s}_{t-1})\Sigma + \Omega}\hat{\omega}_t,$$

so that we can rewrite Equation (20) as

$$\hat{p}_t = \mathbb{E}\left[\hat{\delta}_t \frac{\beta_t(\hat{s}_{t-1})\Sigma}{\beta_t^2(\hat{s}_{t-1})\Sigma + \Omega}\hat{\omega}_t \mid \hat{\omega}_t, \hat{s}_{t-1}\right] = \frac{\mu_{t-1}(\hat{s}_{t-1})\beta_t(\hat{s}_{t-1})\Sigma}{\beta_t^2(\hat{s}_{t-1})\Sigma + \Omega}\hat{\omega}_t.\quad (21)$$

■

Proof of Lemma 2. To see this, recall from the proof of Lemma 1 that the insider chooses a demand $\hat{x}_t = \frac{\bar{\mu}_{t-1}(\hat{s}_{t-1})\hat{\theta}_t}{2\lambda_t(\hat{s}_{t-1})}$ in order to maximize

$$\mathbb{E}_b[\hat{\pi}_t \mid \hat{\theta}_t, \hat{s}_{t-1}, \hat{x}_t] = \hat{x}_t \left[\bar{\mu}_{t-1}(\hat{s}_{t-1})\hat{\theta}_t - \lambda_t(\hat{s}_{t-1})\hat{x}_t \right].\quad (22)$$

Of course, this demand \hat{x}_t is not the same as that of a rational but otherwise identical insider, who instead would be maximizing unbiased expected profits:

$$\mathbb{E}[\hat{\pi}_t \mid \hat{\theta}_t, \hat{s}_{t-1}, \hat{x}_t] = \hat{x}_t \left[\mu_{t-1}(\hat{s}_{t-1})\hat{\theta}_t - \lambda_t(\hat{s}_{t-1})\hat{x}_t \right].\quad (23)$$

Since the market maker is rational in this model, he knows that the (biased) insider's correct expected profits are given by Equation (23), using the suboptimal demand \hat{x}_t :

$$\begin{aligned}\mathbb{E}[\hat{\pi}_t \mid \hat{\theta}_t, \hat{s}_{t-1}, \hat{x}_t] &= \frac{\bar{\mu}_{t-1}(\hat{s}_{t-1})\hat{\theta}_t}{2\lambda_t(\hat{s}_{t-1})} \left[\mu_{t-1}(\hat{s}_{t-1})\hat{\theta}_t - \lambda_t(\hat{s}_{t-1}) \frac{\bar{\mu}_{t-1}(\hat{s}_{t-1})\hat{\theta}_t}{2\lambda_t(\hat{s}_{t-1})} \right] \\ &= \frac{\bar{\mu}_{t-1}(\hat{s}_{t-1})\hat{\theta}_t^2}{4\lambda_t(\hat{s}_{t-1})} \left[2\mu_{t-1}(\hat{s}_{t-1}) - \bar{\mu}_{t-1}(\hat{s}_{t-1}) \right].\end{aligned}\quad (24)$$

Suppose first that the market maker quotes a price schedule with a positive slope. On average, he then expects to profit from the liquidity trader. To perform his market clearing duties competitively, it must therefore be the case that the market maker loses that same amount to the insider on average; that is, it must be the case that Equation (24) is positive. However, when $2\mu_{t-1}(\hat{s}_{t-1}) < \bar{\mu}_{t-1}(\hat{s}_{t-1})$, this is not the case, so that an equilibrium with a positively sloped price schedule is impossible.

What happens if the market maker quotes a price schedule with a negative slope? From Equation (22), we see that the insider's problem degenerates, as he would then choose an infinite demand. This would not only make his biased expected profits infinite (and positive), but would also make his unbiased expected profits in Equation (23) infinite (and negative). More than that, any negatively sloped price schedule implies that the market maker will also lose against the liquidity trader. It is therefore impossible for the market maker to perform his duties competitively with any negatively sloped price schedule. ■

Proof of Lemma 3. The sufficiency part of the argument is obvious as $H \leq 2L$ implies that

$$2\mu_{t-1}(\hat{s}_{t-1}) \geq 2L \geq H \geq \bar{\mu}_{t-1}(\hat{s}_{t-1}).$$

To show necessity, we show that if $2L < H$, then $2\mu_t(s) < \bar{\mu}_t(s)$ for some integers s and t such that $0 < s \leq t$, and some $\gamma > 1$. So suppose that $2L < H$. For any $\epsilon > 0$, it is possible to find integers s and t such that $0 < s \leq t$ and $\mu_t(s) \leq L + \epsilon$.

Since $\bar{\mu}_t(s)$ increases to H as γ increases to infinity, it is also possible to find $\gamma > 1$ such that $\bar{\mu}_t(s) \geq H - \epsilon$. By choosing ϵ to be strictly smaller than $\frac{H-2L}{3}$, we have $2\mu_t(s) \leq 2(L + \epsilon) < H - \epsilon \leq \bar{\mu}_t(s)$. This completes the proof. ■

Proof of Proposition 1. By using Equation (11) in Equation (10) and rearranging, we obtain

$$2\mu_{t-1}(s)\Sigma\beta_t^2(s) = \bar{\mu}_{t-1}(s)\Sigma\beta_t^2(s) + \bar{\mu}_{t-1}(s)\Omega,$$

which is quadratic in $\beta_t(s)$. By Lemma 3, our assumption that $H \leq 2L$ ensures that $2\mu_{t-1}(s) \geq \bar{\mu}_{t-1}(s)$, so we can solve for $\beta_t(s)$ and obtain Equation (12a), as desired (the other root is rejected, since it represents a minimum, not a maximum). Finally, using Equation (12a) for $\beta_t(s)$ in Equation (11) yields Equation (12b). ■

Proof of Proposition 2. When $\hat{a} = H$, we expect the insider to correctly predict the one-period dividend a fraction H of the time. So, as t tends to ∞ , we expect his updated posteriors $\bar{\phi}_t(s)$ to behave like

$$\frac{1}{1 + \left(\frac{L}{\gamma H}\right)^{Ht} \left(\frac{1-L}{1-H}\right)^{t-Ht} \frac{1-\phi_0}{\phi_0}} = \frac{1}{1 + \left[\left(\frac{L}{\gamma H}\right)^H \left(\frac{1-L}{1-H}\right)^{1-H}\right]^t \frac{1-\phi_0}{\phi_0}}.$$

This last quantity will converge to 1 as $t \rightarrow \infty$ if $\left(\frac{L}{\gamma H}\right)^H \left(\frac{1-L}{1-H}\right)^{1-H} < 1$, which is easily shown to be the case.

When $\hat{a} = L$, we expect the insider to correctly guess the one-period dividend a fraction L of the times as $t \rightarrow \infty$. So as we play the game more and more often (as t tends to ∞), we expect his updated posteriors $\bar{\phi}_t(s)$ to behave like

$$\frac{1}{1 + \left(\frac{L}{\gamma H}\right)^{Lt} \left(\frac{1-L}{1-H}\right)^{t-Lt} \frac{1-\phi_0}{\phi_0}} = \frac{1}{1 + \left[\left(\frac{L}{\gamma H}\right)^L \left(\frac{1-L}{1-H}\right)^{1-L}\right]^t \frac{1-\phi_0}{\phi_0}}.$$

So $\bar{\phi}_t(s)$ will converge to 0, ϕ_0 , or 1 according to whether the expression in square brackets is greater than, equal to, or smaller than 1. It can be shown that these three situations will occur when $\gamma < \gamma^*$, $\gamma = \gamma^*$, and $\gamma > \gamma^*$ respectively, where $\gamma^* = \frac{L}{H} \left(\frac{1-L}{1-H}\right)^{(1-L)/L}$. ■

Proof of Proposition 3. Since the denominator of $K_t(s)$ in Equation (13) is not a function of γ , $\frac{\partial K_t(s)}{\partial \gamma}$ will have the same sign as $\frac{\partial \bar{\mu}_t(s)}{\partial \gamma}$. We show in Result B2 of Appendix B that $\frac{\partial \bar{\mu}_t(s)}{\partial \gamma} > 0$. ■

Proof of Proposition 4. In our model, the number of successes in the first t periods is obviously an integer in $\{0, 1, \dots, t\}$, but the function $K_t(s)$ is well defined for any $s \in [0, t]$. However, we will have the desired result if we can show that $K_t(s)$ is increasing for $s \in [0, s_0]$ and decreasing for $s \in [s_0, t]$ for some $s_0 \in [0, t]$.

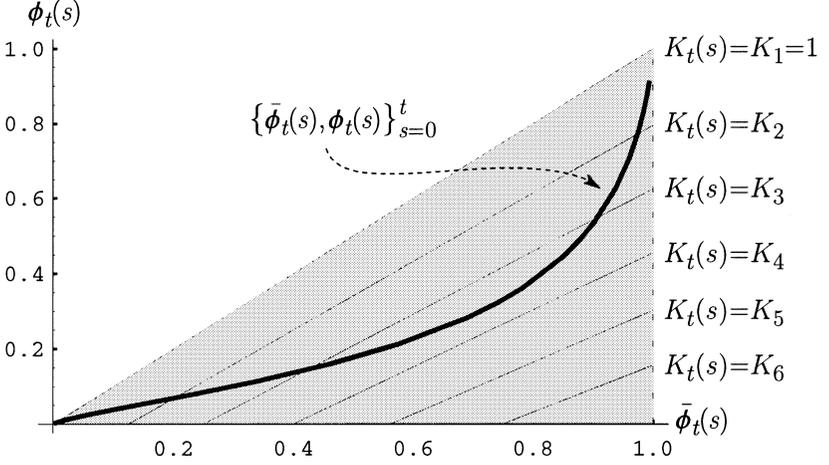


Figure 5
Relationship between biased and unbiased beliefs

This figure shows $\phi_t(s)$ as a function of $\bar{\phi}_t(s)$. For any $s \in [0, t]$, we have $\bar{\phi}_t(s) \geq \phi_t(s)$, so that all the points $\{\bar{\phi}_t(s), \phi_t(s)\}_{s=0}^t$ must lie in the gray area. The thin solid lines represent the linear “iso-confidence” curves $K_t(s) = K_i$, $i = 1, \dots, 6$ for $1 = K_1 < K_2 < \dots < K_6$. The thick solid line represents the parametric curve $\{\phi_t(s), \bar{\phi}_t(s)\}_{s=0}^t$, where $\phi_t(s)$ and $\bar{\phi}_t(s)$ are given in Equations (1) and (4), respectively. As a function of $\bar{\phi}_t(s)$, $\phi_t(s)$ is first concave and then convex.

To show this, recall that $K_t(s) = \frac{\bar{\mu}_t(s)}{\mu_t(s)} = \frac{L+(H-L)\bar{\phi}_t(s)}{L+(K-L)\phi_t(s)}$. If we define an “iso-confidence” curve by $K_t(s) = K_i$ for some constant $K_i \geq 1$, each of these curves can then be written as a straight line in a $\bar{\phi}_t(s)$ - $\phi_t(s)$ diagram. More precisely, each iso-confidence curve can be expressed as

$$\phi_t(s) = \frac{1}{K_i} \left[\bar{\phi}_t(s) - \frac{(K_i - 1)L}{H - L} \right].$$

These lines are shown as thin solid lines for $1 = K_1 < K_2 < \dots < K_6$ in Figure 5.

From Result B1 in Appendix B, we know that the parametric curve $\{\bar{\phi}_t(s), \phi_t(s)\}_{s=0}^t$ starts at $(0, 0)$ and is increasing. This curve is shown as a thick solid line in Figure 5. Since the iso-confidence curves are linear in this $\bar{\phi}_t(s)$ - $\phi_t(s)$ diagram, we only need to show that $\phi_t(s)$ is first concave and then convex as a function of $\bar{\phi}_t(s)$. Indeed, it will then be the case that each iso-confidence curve is crossed at most twice or, equivalently, that $K_t(s)$ is increasing and then decreasing as a function of s . This can be shown using standard calculus and Result B1 in Appendix B. ■

Proof of Lemma 4. First, a standard result for normal variables is that, if $\hat{y} \sim N(0, \sigma^2)$, then $E|\hat{y}| = \sqrt{\frac{2\sigma^2}{\pi}}$. We can therefore calculate

$$\begin{aligned} E \left[\hat{\psi}_{t+1} \mid \hat{s}_t = s \right] &= \frac{1}{2} E \left[|\beta_{t+1}(s) \hat{\theta}_{t+1}| + |\hat{z}_{t+1}| \mid \hat{s}_t = s \right] \\ &= \frac{1}{2} \left[\beta_{t+1}(s) \sqrt{\frac{2\Sigma}{\pi}} + \sqrt{\frac{2\Omega}{\pi}} \right] \\ &= \frac{1}{\sqrt{2\pi}} \left[\beta_{t+1}(s) \sqrt{\Sigma} + \sqrt{\Omega} \right], \end{aligned}$$

and this last expression is equal to Equation (14). The expression for expected profits is derived as follows:

$$\begin{aligned}
 E[\hat{\pi}_{t+1} | \hat{s}_t = s] &= E\left\{E\left[\hat{\pi}_{t+1} | \hat{s}_t, \hat{\theta}_{t+1}, \hat{x}_{t+1}\right] \middle| \hat{s}_t = s\right\} \\
 &\stackrel{(24)}{=} E\left\{E\left[\frac{\bar{\mu}_t(\hat{s}_t)\hat{\theta}_{t+1}^2}{4\lambda_{t+1}(\hat{s}_t)}\left[2\mu_t(\hat{s}_t) - \bar{\mu}_t(\hat{s}_t)\right] \middle| \hat{s}_t, \hat{\theta}_{t+1}, \hat{x}_{t+1}\right] \middle| \hat{s}_t = s\right\} \\
 &\stackrel{(12b)}{=} \frac{1}{2}\sqrt{\frac{\Sigma}{\Omega}}\bar{\mu}_{t-1}(s)\left[2\mu_{t-1}(s) - \bar{\mu}_{t-1}(s)\right]E\left(\hat{\theta}_{t+1}^2 | \hat{s}_t = s\right) \\
 &= \frac{1}{2}\sqrt{\Sigma\Omega}\sqrt{\bar{\mu}_t(s)}\left[2\mu_t(s) - \bar{\mu}_t(s)\right].
 \end{aligned}$$

Finally, we have

$$\begin{aligned}
 \text{var}(\hat{p}_{t+1} | \hat{s}_t = s) &= E\left(\hat{p}_{t+1}^2 | \hat{s}_t = s\right) \\
 &= E\left[\lambda_{t+1}^2(\hat{s}_t)\hat{\omega}_{t+1}^2 | \hat{s}_t = s\right] \\
 &= E\left[\lambda_{t+1}^2(\hat{s}_t)(\hat{x}_{t+1} + \hat{z}_{t+1})^2 | \hat{s}_t = s\right] \\
 &= E\left\{\lambda_{t+1}^2(\hat{s}_t)\left[\beta_{t+1}(\hat{s}_t)\hat{\theta}_{t+1} + \hat{z}_{t+1}\right]^2 \middle| \hat{s}_t = s\right\} \\
 &= \lambda_{t+1}^2(s)\left[\beta_{t+1}^2(s)E(\hat{\theta}_{t+1}^2 | \hat{s}_t = s)\right. \\
 &\quad \left.+ 2\beta_{t+1}(s)E(\hat{\theta}_{t+1}\hat{z}_{t+1} | \hat{s}_t = s) + E(\hat{z}_{t+1}^2 | \hat{s}_t = s)\right] \\
 &= \lambda_{t+1}^2(s)\left[\beta_{t+1}^2(s)\Sigma + \Omega\right] = \frac{\Sigma}{2}\bar{\mu}_t(s)\mu_t(s),
 \end{aligned}$$

where the last equality is obtained by using Equations (12a) and (12b) in Proposition 1. This completes the proof. ■

Proof of Proposition 5. Given the expression for the conditional expected volume in Equation (14), it is sufficient to prove that $\frac{\partial \beta_t(s)}{\partial \gamma} > 0$. Straightforward differentiation of the expression for $\beta_t(s)$ in Equation (12a) results in

$$\frac{\partial \beta_t(s)}{\partial \gamma} = \sqrt{\frac{\Omega}{\Sigma}} \frac{2\mu_{t-1}(s) - \bar{\mu}_{t-1}(s)}{\bar{\mu}_{t-1}(s)} \frac{\mu_{t-1}(s)}{[2\mu_{t-1}(s) - \bar{\mu}_{t-1}(s)]^2} \frac{\partial \bar{\mu}_{t-1}(s)}{\partial \gamma},$$

which in turn shows that it is sufficient to show that $\frac{\partial \bar{\mu}_{t-1}(s)}{\partial \gamma} > 0$. This is shown to be true in Result B2 of Appendix B. ■

Proof of Proposition 6. To show the desired result, we only need to show that $\bar{\mu}_{t-1}(s)[2\mu_{t-1}(s) - \bar{\mu}_{t-1}(s)]$ is decreasing in γ . This is straightforward to show since

$$\frac{\partial}{\partial \gamma} \left\{ \bar{\mu}_{t-1}(s) [2\mu_{t-1}(s) - \bar{\mu}_{t-1}(s)] \right\} = -2[\bar{\mu}_{t-1}(s) - \mu_{t-1}(s)] \frac{\partial}{\partial \gamma} \bar{\mu}_{t-1}(s),$$

and $\frac{\partial}{\partial \gamma} \bar{\mu}_{t-1}(s)$ is shown to be positive in Result B2 of Appendix B. ■

Proof of Proposition 7. The result easily follows from the fact that $\frac{\partial}{\partial \gamma} \bar{\mu}_{t-1}(s) > 0$, which is shown to be true in Result B2 of Appendix B. ■

Proof of Proposition 8. As shown in Lemma 4, the expected volume in period $t + 1$ is proportional to the expected trading intensity $\beta_{t+1}(s)$ of the insider in that period. Since $\beta_{t+1}(s)$ is monotonically increasing in $K_t(s)$ [see Equation (17)], Proposition 4 immediately implies this result. ■

Proof of Proposition 9. This result is shown in essentially the same way that Proposition 4 was shown earlier, except that the “iso-profit” curves are now quadratic. ■

Proof of Proposition 10. In view of Equation (16), this amounts to showing that the product $\bar{\mu}_t(s)\mu_t(s)$ is increasing in s . However, since both these quantities are increasing in s (see Result B2 in Appendix B), the result follows easily. ■

Appendix B

This appendix contains results that are used in the proofs of some propositions in Appendix A.

Result B1. *The partial derivatives of $\bar{\phi}_t(s)$ in Equation (4) with respect to γ and s are respectively equal to*

$$\frac{\partial \bar{\phi}_t(s)}{\partial \gamma} = \frac{s}{\gamma} \bar{\phi}_t(s) [1 - \bar{\phi}_t(s)] \geq 0, \text{ and} \quad (25)$$

$$\frac{\partial \bar{\phi}_t(s)}{\partial s} = \bar{\phi}_t(s) [1 - \bar{\phi}_t(s)] \log \left(\frac{\gamma H}{L} \frac{1-L}{1-H} \right) \geq 0. \quad (26)$$

Proof. Straightforward differentiation of $\bar{\phi}_t(s)$ yields $\frac{\partial \bar{\phi}_t(s)}{\partial \gamma} = \frac{s}{\gamma} \bar{\phi}_t(s) [1 - \bar{\phi}_t(s)]$, which is obviously greater than or equal to zero. Now, it is easy to show that

$$\frac{\partial \bar{\phi}_t(s)}{\partial s} = \bar{\phi}_t(s) [1 - \bar{\phi}_t(s)] \log \left(\frac{\gamma H}{L} \frac{1-L}{1-H} \right).$$

Since $\gamma H > L$ and $1-L > 1-H$, this last quantity is obviously greater than or equal to zero. ■

Result B2. *The partial derivatives of $\bar{\mu}_t(s)$ in Equation (5) with respect to γ and s are respectively equal to*

$$\frac{\partial \bar{\mu}_t(s)}{\partial \gamma} = (H-L) \frac{s}{\gamma} \bar{\phi}_t(s) [1 - \bar{\phi}_t(s)] \geq 0, \text{ and} \quad (27)$$

$$\frac{\partial \bar{\mu}_t(s)}{\partial s} = (H-L) \bar{\phi}_t(s) [1 - \bar{\phi}_t(s)] \log \left(\frac{\gamma H}{L} \frac{1-L}{1-H} \right) \geq 0. \quad (28)$$

Proof. Since we have $\bar{\mu}_t(s) = H\bar{\phi}_t(s) + L[1 - \bar{\phi}_t(s)] = L + (H-L)\bar{\phi}_t(s)$ and $H > L$, this result follows immediately from Result B1 above. ■

References

- Alchian, A. A., 1950, “Uncertainty, Evolution, and Economic Theory,” *Journal of Political Economy*, 58, 211–221.
- Alpert, M., and H. Raiffa, 1982, “A Progress Report on the Training of Probability Assessors,” in D. Kahneman, P. Slovic, and A. Tversky (eds.), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, UK, pp. 294–305.

- Barber, B., and T. Odean, 1998, "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment," working paper, University of California, Davis; forthcoming in *Quarterly Journal of Economics*.
- Barber, B., and T. Odean, 2000, "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance*, 55, 773–806.
- Benos, A., 1998, "Aggressiveness and Survival of Overconfident Traders," *Journal of Financial Markets*, 1, 353–383.
- Bhattacharya, U., and M. Spiegel, 1991, "Insiders, Outsiders, and Market Breakdowns," *Review of Financial Studies*, 4, 255–282.
- Blume, L., and D. Easley, 1982, "Learning To Be Rational," *Journal of Economic Theory*, 26, 340–351.
- Blume, L., and D. Easley, 1992, "Evolution and Market Behavior," *Journal of Economic Theory*, 58, 9–40.
- Bossaerts, P., 1999, "The Dynamics of Limited-Liability Securities Prices in Fallible Markets," working paper, California Institute of Technology.
- Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, "Investor Psychology and Security Market Under- and Overreactions," *Journal of Finance*, 53, 1839–1885.
- De Long, J., A. Shleifer, L. Summers, and R. Waldmann, 1990, "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98, 703–738.
- De Long, J., A. Shleifer, L. Summers, and R. Waldmann, 1991, "The Survival of Noise Traders in Financial Markets," *Journal of Business*, 64, 1–19.
- Fiske, S., and S. Taylor, 1991, *Social Cognition* (2nd ed.), McGraw-Hill, New York.
- Friedman, M., 1953, *Essays in Positive Economics*, University of Chicago Press, Chicago.
- Hastorf, A. H., D. J. Schneider, and J. Polefka, 1970, *Person Perception*, Addison-Wesley, Reading, Mass.
- Kyle, A., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.
- Kyle, A., and F. A. Wang, 1997, "Speculation Duopoly With Agreement to Disagree: Can Overconfidence Survive the Market Test?" *Journal of Finance*, 52, 2073–2090.
- Langer, E., and J. Roth, 1975, "Heads I Win, Tails It's Chance: The Illusion of Control as a Function of the Sequence of Outcomes in a Purely Chance Task," *Journal of Personality and Social Psychology*, 32, 951–955.
- Lichtenstein, S., B. Fischhoff, and L. Phillips, 1982, "Calibration of Probabilities: The State of the Art to 1980," in D. Kahneman, P. Slovic, and A. Tversky (eds.), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, UK, pp. 306–334.
- Luo, G. Y., 1998, "Market Efficiency and Natural Selection in a Commodity Futures Market," *Review of Financial Studies*, 11, 647–674.
- Milgrom, P., and N. Stokey, 1982, "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26, 17–27.
- Miller, D., and M. Ross, 1975, "Self-serving Biases in Attribution of Causality: Fact or Fiction?" *Psychological Bulletin*, 82, 213–225.
- Odean, T., 1998, "Volume, Volatility, Price, and Profit When All Traders Are Above Average," *Journal of Finance*, 53, 1887–1934.
- Rabin, M., and J. Schrag, 1999, "First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics*, 114, 37–82.
- Roll, R., 1984, "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *Journal of Finance*, 39, 1127–1139.

Statman, M., and S. Thorley, 1998, "Overconfidence and Trading Volume," working paper, Santa Clara University.

Wang, A., 1997, "Overconfidence, Delegated Fund Management, and Survival," working paper, Columbia University.

Wolosin, R. J., S. J. Sherman, and A. Till, 1973, "Effects of Cooperation and Competition on Responsibility Attribution After Success and Failure," *Journal of Experimental Social Psychology*, 9, 220–235.