

NBER WORKING PAPER SERIES

A MODEL OF INVESTOR SENTIMENT

Nicholas Barberis
Andrei Shleifer
Robert W. Vishny

Working Paper 5926

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 1997

We are grateful to the NSF for financial support, and to Olivier Blanchard, Alon Brav, John Campbell, Ed Glaeser, J.B. Heaton, Danny Kahneman, David Laibson, Owen Lamont, and Dick Thaler for comments. This paper is part of NBER's research program in Asset Pricing. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1997 by Nicholas Barberis, Andrei Shleifer and Robert W. Vishny. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

A Model of Investor Sentiment
Nicholas Barberis, Andrei Shleifer and
Robert W. Vishny
NBER Working Paper No. 5926
February 1997
JEL No. 313
Asset Pricing

ABSTRACT

Recent empirical research in finance has uncovered two families of pervasive regularities: *underreaction* of stock prices to news such as earnings announcements; and *overreaction* of stock prices to a series of good or bad news. In this paper, we present a parsimonious model of investor sentiment—that is, of how investors form beliefs—that is consistent with the empirical findings. The model is based on psychological evidence and produces both underreaction and overreaction for a wide range of parameter values.

Nicholas Barberis
Graduate School of Business
The University of Chicago
1101 East 58th Street
Chicago, IL 60637

Andrei Shleifer
Department of Economics
Harvard University
315 Littauer Center
Cambridge, MA 02138
and NBER
shleifer@fas.harvard.edu

Robert W. Vishny
Graduate School of Business
The University of Chicago
1101 East 58th Street
Chicago, IL 60637
and NBER
vishny@gsb.uchicago.edu

1 Introduction

Recent empirical research in finance has identified two families of pervasive regularities: underreaction and overreaction. The underreaction evidence shows that over horizons of perhaps one to twelve months, security prices underreact to news¹. As a consequence, news is incorporated only slowly into prices, which tend to exhibit positive autocorrelations over these horizons. A related way to make this point is to say that current good (bad) news has power in predicting positive (negative) returns in the future. The overreaction evidence shows that over longer horizons of perhaps 3 to 5 years, security prices overreact to consistent patterns of news pointing in the same direction. That is, securities that have had a long record of good (bad) news tend to become overpriced (underpriced) and have low (high) average returns afterwards². Put differently, securities with strings of good (bad) performance, however measured, get extremely high (low) valuations, and these valuations on average return to the mean.

This evidence presents a challenge to the efficient markets theory because it suggests that in a variety of markets, sophisticated investors can earn superior returns by taking advantage of underreaction and overreaction without bearing extra risk. Recent attempts to explain the evidence from the efficient markets viewpoint have in fact failed (Fama and French (1996)). This evidence also presents a challenge to behavioral finance theory because early models do not successfully explain the facts³. The challenge is to explain how investors might form beliefs that lead to both underreaction and overreaction.

In this paper, we propose a parsimonious model of investor sentiment – of how investors form beliefs – that is consistent with the available statis-

¹Some of the papers in this area, discussed in more detail in Section 2, include Cutler, Poterba and Summers (1991), Thomas (1992), Jegadeesh and Titman (1993), and Chan, Jegadeesh, Lakonishok (1996).

²Some of the papers in this area, discussed in more detail in Section 2 include Cutler, Poterba and Summers (1991), DeBondt and Thaler (1985), Chopra, Lakonishok, and Ritter (1992), Fama and French (1992), Lakonishok, Shleifer and Vishny (1994) and La Porta (1996).

³The model of De Long et al (1990a) generates negative autocorrelation in returns, and that of De Long et al (1990b) generates positive autocorrelation. Cutler, Poterba and Summers (1991) combine elements of the two De Long et. al. models in an attempt to explain some of the autocorrelation evidence. These models focus exclusively on prices and hence do not confront the crucial earnings evidence discussed in Section 2.

tical evidence. The model is also consistent with experimental evidence on both the failures of individual judgment under uncertainty and the trading patterns of investors in experimental situations. In particular, our specification is consistent with the results of Kahneman and Tversky (1974) on the important behavioral heuristic known as representativeness: this is the tendency of experimental subjects to view events as typical or representative of some specific class and to ignore the laws of probability in the process. In the stock market, for example, investors might classify some stocks as growth stocks based on recent history, ignoring the likelihood that there are very few companies that just keep growing. Our model also relates to another phenomenon documented in psychology, namely conservatism, defined as slow updating of models in the face of new evidence (Edwards (1968)). The underreaction evidence in particular is consistent with conservatism.

Our model is that of one investor and one asset. This investor should be viewed as one whose beliefs reflect “consensus forecasts” even when different investors hold different expectations. The beliefs of this representative investor affect prices and returns.

We do not explain in this model why arbitrage fails to eliminate the mispricing. For the purposes of this paper, we rely on earlier work that has shown why deviations from efficient prices can persist (De Long et al (1990a), Shleifer and Vishny (1996)). According to this work, an important reason that arbitrage is limited is that movements in investor sentiment are in part unpredictable, and therefore arbitrageurs betting against mispricing run the risk that, at least in the short run, investor sentiment becomes more extreme and prices move even further away from fundamental value. As a consequence of such “noise trader risk”, arbitrage positions often lose money in the short run. When arbitrageurs are risk-averse, or when they manage other people’s money and run the risk of losing funds under management when performance is poor, the risk of deepening mispricing reduces the size of the positions they take. Hence arbitrage fails to eliminate the mispricing completely and investor sentiment affects security prices in equilibrium. In the model below, investor sentiment is indeed in part unpredictable, and therefore, if arbitrageurs were introduced into the model, arbitrage would be limited.

While these earlier papers argue that mispricing can persist, they say little about the nature of the mispricing that might be observed. To do the latter, we need a model of how people form expectations. This is the focus of the current paper. We do not model the process of arbitrage explicitly and

simply rely on the earlier work to assume that it is not fully effective.

In our model, the earnings of the asset follow a random walk. However, the investor does not know that. Rather, he believes that the behavior of a given firm's earnings moves between two "states" or "regimes". In the first state, earnings are mean reverting. In the second state, they trend, i.e. are likely to rise further after an increase. The transition probabilities between the two regimes, as well as the statistical properties of the earnings process in each one of them, are fixed in the investor's mind. In particular, in any given period, the firm's earnings are more likely to stay in a given regime than to switch. Each period, the investor observes earnings. He then uses this information to update his beliefs about which regime he is in. In his updating, the investor is Bayesian, although his model of the return generating process is inaccurate. Specifically, when a positive earnings surprise is followed by another positive surprise, the investor raises the likelihood that he is in the trending regime, whereas when a positive surprise is followed by a negative surprise, the investor raises the likelihood that he is in the mean-reverting regime. We solve this model and show that, for a plausible range of parameter values, it generates the empirical predictions observed in the data⁴.

Section 2 of the paper summarizes the evidence that we try to explain. Section 3 presents the model. Section 4 solves it and presents the implications for the data. Section 5 concludes.

2 The Evidence

In this section, we summarize the statistical evidence of underreaction and overreaction in security returns. Although most of the sophisticated research on financial markets has focused on the behavior of aggregate stock (and bond) price indices, we devote only minor attention to this evidence. The reason for this is that aggregate time series generally do not provide enough information to reject the hypothesis of efficient markets. As a consequence, most of the anomalous evidence that our model tries to explain has come from the cross-section of stock returns. Much of this evidence is from the

⁴Daniel, Hirshleifer, and Subrahmanyam (1996) also construct a model of investor sentiment aimed at reconciling the various empirical findings of underreaction and overreaction. They too use concepts from the psychology literature to support their framework, although the underpinnings of their model, namely investor overconfidence and self-attribution, are rather different from our own.

United States, although some recent research has found similar patterns in other markets. In addition, this section presents some direct evidence on expectations that our model draws on.

2.1 Statistical Evidence of Underreaction

Before presenting the empirical findings, we first explain what we mean by underreaction to news announcements. Suppose that in each time period, the investor hears news about a particular company. We denote the news he hears in period t as z_t . This news can be either good or bad, i.e. $z_t = G$, or $z_t = B$. By underreaction we mean that the average return on the company's stock in the period following an announcement of good news is *higher* than the average return in the period following bad news:

$$E(r_{t+1}|z_t = G) > E(r_{t+1}|z_t = B).$$

In other words, the stock underreacts to the good news, a mistake which is corrected in the following period, giving a higher return at that time. In this paper, the good news consists of an earnings announcement that is higher than expected, although as we will discuss below, there is considerable evidence of underreaction to other types of news as well.

Empirical analysis of aggregate time series has produced evidence which strongly suggests underreaction. Cutler, Poterba, and Summers (1991) examine autocorrelations in returns on various indexes over different horizons. They look at returns on stocks, bonds, and foreign exchange in different markets over the period 1960-1988, and generally, though not uniformly, find positive autocorrelations in index returns over horizons of between one month and one year. For example, the average one month autocorrelation in stock returns across the world is around .1 (which is similar in the United States), and that in bond returns is around .2 (and around 0 in the United States). Many of these autocorrelations are statistically significant. This autocorrelation evidence is consistent with the underreaction hypothesis, which states that stock prices incorporate information slowly, leading to trends in returns over short horizons.

Much better and more direct evidence for the underreaction hypothesis comes from the studies of the cross-section of stock returns in the United States, which look at the actual news events as well as the predictability of returns. Bernard (1992) surveys one class of such studies, which deals with the underreaction of stock prices to announcements of company earnings.

The finding of the studies surveyed by Bernard is roughly as follows. Suppose we sort stocks into groups (say deciles) based on how much of a surprise is contained in their earnings announcement. One naive way to measure an earnings surprise is to look at standard unexpected earnings (SUE), defined as the difference between a company's earnings in a given quarter and its earnings during the quarter a year before, scaled by the standard deviation of the company's earnings. Another way to measure an earnings surprise is by the stock price reaction to the earnings announcement. A general (and unsurprising) finding is that stocks with positive earnings surprises also earned relatively high returns in the period prior to the earnings announcement, as information about earnings was slowly incorporated into prices. A much more surprising finding is that stocks with higher earnings surprises also earn higher returns in the period after portfolio formation: the market underreacts to the earnings announcement in revising a company's stock price. For example, in the 60 day period after portfolio formation, stocks with the highest SUE earn a cumulative risk-adjusted return that is 4.2 percent higher than the return on stocks with the lowest SUE (see Bernard (1992)). Thus stale information, namely the SUE or the past earnings announcement return, has predictive power for future risk-adjusted returns. Or, put differently, information about earnings is only slowly incorporated into stock prices.

Bernard (1992) also summarizes some evidence on the actual properties of the time series of earnings, and provides an interpretation for his findings. The relevant series is changes in a company's earnings in a given quarter relative to the same calendar quarter in the previous year. Over the period 1974-1986 and using a sample of 2,626 firms, Bernard and Thomas (1989) find that these series tend to exhibit an autocorrelation of about .34 at a lag of one quarter, of .19 at two quarters, of .06 at three quarters, and -.24 at four quarters. That is, earnings changes exhibit a slight trend at one, two and three quarter horizons and a slight reversal after a year. In interpreting the evidence, Bernard (1992) conjectures that market participants do not recognise the positive autocorrelations in earnings changes, and in fact believe that earnings follow a random walk. This belief causes them to underreact to earnings announcements. Our stylized model in Section 3 uses a related idea for generating underreaction: we suppose that earnings follow a random walk but that investors typically assume that earnings are mean-reverting. The key idea that generates underreaction, and that Bernard's and our analysis share, is that investors typically (not always) believe that earnings are more stationary than they really are. As we show below, this

idea has firm foundations in psychology.

The earnings drift evidence surveyed by Bernard appears to be sufficiently strong that it is regarded as an empirical challenge even by normally unimpressed accountants. It is not the only evidence of underreaction, however. Other events are also followed by a drift in stock prices. For example, Ikenberry et al (1995) find that stock prices rise on the announcement of share repurchases, but then continue to drift in the same direction over the next several years. Michaely, Thaler, and Womack (1995) find similar evidence of drift following dividend initiations and omissions. Finally, Seyhun (1992) finds that investors may be able to earn abnormal returns by trading on public information about insider trades, suggesting that prices adjust only incompletely to the news of such trades. All these examples are evidence of underreaction similar to Bernard's evidence on the earnings drift.

Further evidence of underreaction comes from Jegadeesh and Titman (1993) who examine a cross-section of U.S. stock returns and find rather reliable evidence that over the six-month horizon, stock returns are positively autocorrelated. Similarly to the earnings drift evidence, they interpret their finding of the "momentum" in stock returns as pointing to underreaction to information, and slow incorporation of information into prices. Chan, Jegadeesh, and Lakonishok (1996) integrate the earnings drift evidence with the momentum evidence. They use three measures of earnings surprise: SUE, stock price reaction to the earnings announcement, and changes in analysts' forecasts of earnings. The authors find that all these measures, as well as the past return, help predict subsequent stock returns at horizons of 6 months and one year. That is, they find that stocks with a positive earnings surprise, as well as stocks with high past returns, tend to subsequently outperform stocks with a negative earnings surprise and poor returns. Like the other authors, Chan, Jegadeesh and Lakonishok conclude that investors underreact to news, and incorporate information into prices slowly.

2.2 Statistical Evidence of Overreaction

Analogous to the definition of underreaction at the start of the previous subsection, we now define overreaction as follows: the average return following not one, but a series of announcements of good news is *lower* than the average return following a series of bad news announcements. Using the same

notation as before,

$$E(r_{t+1}|z_t = G, z_{t-1} = G, \dots, z_{t-j} = G) < E(r_{t+1}|z_t = B, z_{t-1} = B, \dots, z_{t-j} = B),$$

where j is at least 1, and probably rather higher. The idea here is simply that after a series of announcements of good news, the investor may become overly optimistic that future news announcements will also be good, and hence overreact, sending the stock price to unduly high levels. Subsequent news announcements are likely to contradict his optimism, leading to lower returns.

Empirical studies of predictability of aggregate index returns over long horizons are extremely numerous. Some of the early papers include Fama and French (1988) and Poterba and Summers (1988); Cutler, Poterba and Summers (1991) examine some of this evidence for a variety of markets. The thrust of the evidence is that over horizons of 3 to 5 years, there is a (relatively slight) negative autocorrelation of stock returns in many markets. Moreover, over similar horizons, some measures of valuation of stocks, such as the dividend yield, have predictive power for returns in a similar direction, namely low dividend yield or high past return tend to predict a low subsequent return (see Campbell and Shiller (1988)).

As before, the more reliable evidence comes from the cross-section of stock returns. In an early important paper, DeBondt and Thaler (1985) discover from looking at the US data dating back to 1933 that portfolios of stocks that have extremely poor returns over the previous 5 years dramatically outperform portfolios of stocks with extremely high returns, even after making the standard risk adjustments. DeBondt and Thaler's findings were initially greeted with skepticism, but were nonetheless corroborated by later work (e.g., Chopra, Lakonishok, and Ritter (1992)). In the case of earnings, Zarowin (1989) finds that firms which have had a sequence of bad earnings realizations subsequently outperform firms with a sequence of good earnings. This evidence suggests that stocks with consistent records of good news, and hence extremely high past returns, are overvalued, and hence an investor can earn abnormal returns by betting against this overreaction to consistent patterns of news. Similarly, stocks with consistent records of bad news become undervalued and subsequently earn superior returns.

Subsequent work has changed the focus from past returns to other measures of valuation, such as the ratio of market value to book value of assets (DeBondt and Thaler (1987), Fama and French (1992)), market value to

cash flow (Lakonishok, Shleifer and Vishny (1994)) and other accounting measures. All this evidence points in the same direction. Stocks with very high valuations relative to their assets or earnings (glamour stocks), which tend to be stocks of companies with extremely high earnings growth over the previous several years, earn relatively low risk adjusted returns in the future, whereas stocks with low valuations (value stocks) earn relatively high returns. For example, Lakonishok et al (1994) find spreads of 8-10 percent a year between returns of the extreme value and glamour deciles. Again, this evidence points to overreaction to a prolonged record of extreme performance, whether good or bad: the prices of stocks with such extreme performance tend to be too extreme relative to what these stocks are worth, and to what the subsequent returns actually deliver. Recent research extends the overreaction evidence to other markets, including those in Europe and in Japan (Lakonishok, Shleifer and Vishny (1996)).

The economic interpretation of this evidence proved more controversial, since some authors, particularly Fama and French (1992, 1996) argue that glamour stocks are in fact less risky, and value stocks more risky, once risk is properly measured. Generally speaking, evidence has not been kind to this hypothesis. In a direct test of the overreaction hypothesis, La Porta (1996) sorts stocks on the basis of their long term growth rate forecasts made by professional analysts, and finds consistent evidence that analysts are excessively bullish about the stocks they are most optimistic about, and excessively bearish about the stocks they are pessimistic about. In particular, stocks with high growth forecasts earn much lower future returns than stocks with low growth forecasts. Moreover, on average, stocks with high growth forecasts earn negative returns when they subsequently announce earnings and stocks with low growth forecasts earn high returns. All this evidence points to overreaction not just by analysts but more importantly in prices as well: in the world of efficient markets, stocks with optimistic growth forecasts should not earn low returns.

Finally, La Porta et al (1996) find direct evidence of overreaction in glamour and value stocks defined using accounting variables. Specifically, glamour stocks earn negative returns on the days of their future earnings announcements, and value stocks earn positive returns. The market learns when earnings are announced that its valuations have been too extreme.

In sum, the cross-sectional overreaction evidence, like the cross-sectional underreaction evidence, appears to be a rather reliable regularity. These regularities taken in their entirety are inconsistent with the efficient markets

hypothesis. More important for this paper, the two regularities pose a challenge to behavioral finance to provide a model of how investors form beliefs that can account for the regularities.

2.3 Direct Evidence on Expectations

In a sequence of important papers, Frankel and Froot (1986, 1987, 1990) investigate the properties of actual forecasts of future depreciation of the dollar. They use consensus forecasts from three surveys of professional forecasters of the dollar. During the period they examine – the first half of the 1980s – the dollar appreciated significantly against other major currencies. Frankel and Froot find that, during most of this period, investors respond to current appreciation by expecting – at horizons of three to twelve months – future depreciation. If they actually traded on these expectations, the resulting price patterns would have been similar to those found in the underreaction studies of the stock market. However, Frankel and Froot also find that toward the end of their period in late 1984 and early 1985, the consensus forecast switched to expecting a near-term appreciation of the dollar precisely at the time of the dollar’s consistent and spectacular rise. In the subsequent period, the dollar collapsed. This evidence again is consistent with the stock market findings of overreaction, in which a consistent record of good performance leads to excessively optimistic forecasts that are later disappointed.

3 A Model of Investor Sentiment

3.1 Informal Description of the Model

The evidence summarized in the previous section suggests that investors can be in one of two frames of mind, depending on the circumstances. Typically, they underreact to news; however, after repeated bombardment by news of a similar kind, they may overreact to the actual content of the news. The model presented in this section attempts to capture this story. To begin, we briefly overview the model. In the two following subsections, we provide evidence from the psychology literature to support our model, and then present a mathematical formulation of it.

We consider a model with a representative, risk-neutral investor with discount rate δ . We can think of this investor as one whose beliefs reflect the

“consensus”, even if different investors have different beliefs. There is only one security; in this context, the equilibrium price of the security is clearly equal to the net present value of the future earnings that the representative investor expects. Moreover, in contrast to models with heterogeneous agents, there is no information in prices over and above the information already in earnings.

An important point to note is that given the assumptions of risk-neutrality and a constant discount rate, returns will be unpredictable if the investor knows the correct process followed by the earnings stream, a fact first established by Samuelson (1965). If our model is to generate the kind of predictability in returns documented in the empirical studies discussed in Section 2, it must be the case that the investor is using the wrong model to form expectations.

We suppose that the earnings stream follows a random walk. This assumption is not entirely accurate, as we discussed above, since earnings growth rates at one to three quarter horizons are slightly positively auto-correlated (Bernard and Thomas (1989)). We make our assumption for concreteness, and it is not at all essential for generating the results. What is essential is that investors sometimes believe that earnings are more stationary than they really are – the idea stressed by Bernard and captured within our model below. It is this relative misperception that is the key to under-reaction.

The investor in our model does not realize that earnings follow a random walk. He thinks that the world moves between two “states” or “regimes” and that there is a different model governing earnings in each regime. When the world is in regime 1, Model 1 determines earnings; in regime 2, it is Model 2 that determines them. Neither of the two models is a random walk: rather, under Model 1, earnings are mean-reverting; in Model 2, they trend. For simplicity, we specify these models as Markov processes: that is, in each model the change in earnings in period t depends only on the change in earnings in period $t - 1$. The only difference between the two models lies in the transition probabilities. Under Model 1, earnings shocks are likely to be reversed in the following period: hence a positive shock to earnings is more likely to be followed in the next period by a negative shock than by another positive shock. Under Model 2, however, the opposite is true. In this case, shocks are more likely to be followed by another shock of the same sign.

The investor also believes that there is an underlying regime switching process which determines which regime the world is in at any time. We

specify this underlying process as a Markov process as well, so that whether the current regime is Model 1 or Model 2 depends only on what the regime was last period. We focus attention on cases where regime switches are relatively rare. That is, if Model 1 determines the change in earnings in period t , it is likely that it determines earnings in period $t + 1$ also. The same applies to Model 2. With some small probability, though, the regime changes, and the other Model begins generating earnings. For reasons that will be apparent later, we often require the regime-switching probabilities to be such that the investor thinks that the world is in the mean-reverting regime of Model 1 more often than he believes it to be in the trending regime of Model 2.

The transition probabilities associated with Models 1 and 2 and with the underlying regime switching process are fixed in the investor's mind. In order to value the security, the investor needs to forecast future earnings. To do this, he uses the earnings stream he has observed to update his beliefs about which regime is generating earnings. Once this is done, he uses the regime-switching model to forecast future earnings. The investor updates in a Bayesian fashion even though his model of earnings is incorrect. For instance, if he observes two consecutive earnings shocks of the same sign, he will believe more strongly that he is in the trending earnings regime of Model 2. If the earnings shock this period is of the opposite sign to last period's earnings shock, he will put more weight on Model 1, the mean-reverting regime.

It is important to note that our model differs from more typical models of learning: in our framework, the investor never changes the model he is using to forecast earnings: that is, he uses the same regime-switching model, with the same regimes and transition probabilities, throughout. Even after observing a very long stream of earnings data, he does not change his model to something more like a random walk, the true earnings process. His only task is to figure out which of the two regimes of his model is currently generating earnings. That is the only sense in which he is learning from the data. At the end of Section 3, once we have presented a mathematical formulation of our model, we will return briefly to the issue of why it may be reasonable that the investor retain the wrong model even after observing a long data series.

We are now ready to provide some preliminary intuition for how investor behavior of the kind described above coupled with the true random walk process for earnings can generate some of the empirical phenomena discussed

in Section 2. In particular, we show how our framework can lead to both underreaction to earnings announcements, as well as to long-run overreaction.

Take overreaction first. The general view of overreaction described in Section 2 is that investors become too bullish after observing a long record of good news, with the result that low returns are subsequently realized. Similarly, they become too pessimistic after a consistent record of bad news; subsequent returns are higher. In the context of our model, a natural way of capturing this is to say that the average realized return following a string of positive shocks to earnings is lower than the average realized return following a string of negative shocks to earnings⁵.

To see how this can be a feature of our model, consider what happens when our investor sees a series of positive earnings shocks. He naturally puts a high probability on the event that Model 2 has generated current earnings. Since he believes regime switches to be rare, this means that Model 2 is also likely to generate earnings in the next period. The investor therefore expects that the shock to earnings next period will again be positive. Earnings, however, follow a random walk: next period's earnings are equally likely to go up or down. If they go up, the return will not be large, as the investor is expecting exactly that, namely a rise in earnings. If they fall, however, the return will be large and negative as the investor is taken by surprise by the negative announcement. Therefore the average realized return after a string of positive shocks is negative; symmetrically, the average return after a string of negative earnings shocks is positive. The difference between the average returns in the two cases is negative, consistent with the empirically observed overreaction.

Now we turn to underreaction. Following our discussion in Section 2, we can think of underreaction as the fact that the average realized return following a positive shock to earnings is greater than the average realized return following a negative shock to earnings. This corresponds closely to the empirical findings of Bernard (1992), who notes that stocks with high standard unexpected earnings (SUE) earn a higher subsequent return than do stocks with low SUE's⁴.

Underreaction obtains in our model as long as on average, the investor places more weight on Model 1 than on Model 2. If this holds, then consider

⁵For the purposes of the discussion here, we will suppose that good news about earnings corresponds to a positive change in earnings; a more natural definition might be that good news corresponds to earnings that are *higher than expected*. However, these two measures are closely related in our model, as we will see in Section 3.3.

the realized return following a positive earnings shock: since on average, the investor believes Model 1, he on average believes that this positive earnings shock will be partly reversed in the next period. In practice, however, a positive shock is equally likely to be followed by a positive or by a negative shock. If the shock is negative, the realized return is not large, since this is the earnings realization that was expected by the investor. If the shock is positive, the realized return is large and positive, since this shock is quite unexpected to the investor. Similarly, the average realized return following a negative earnings shock is negative, and hence the difference in the average realized returns is indeed positive as required in our definition of underreaction.

The mechanism for expectation formation that we propose here is closely related to that used by Barsky and De Long (1993) in an attempt to explain Shiller's (1981) finding of excess volatility of the price-dividend ratio. They suppose that investors view the growth rate of dividends as a parameter that is not only unknown but also changing over time. The optimal estimate of the parameter will closely resemble a distributed lag on past one-period dividend growth rates, with declining weights. If dividends rise steadily over several periods, the investor's estimate of the current dividend growth rate will also rise, leading him to forecast higher dividends in the future as well. Analogously, in our model, a series of positive shocks to earnings will lead the investor to raise the probability that earnings changes are currently being generated by the trending regime 2, leading him to make more bullish predictions for future earnings.

3.2 Some Psychological Evidence

Our model is consistent with two important phenomena documented by psychologists: the *representativeness heuristic* and *conservatism*. In this subsection, we briefly describe this psychological evidence. We would not go so far as to say that our model is derived from these psychological foundations: only that the psychological evidence is supportive of our assumptions.

According to Kahneman and Tversky (1974), "people replace the laws of chance by heuristics, which sometimes yield reasonable estimates and quite often do not" (p.32). One of the best documented of such heuristics is representativeness. "A person who follows this heuristic evaluates the probability of an uncertain event, or a sample, by a degree to which it is (i) similar in its essential properties to the parent population, (ii) reflects the salient features of the process by which it is generated" (p.33). For example, if a detailed

description of an individual's personality matches up well with the subject's experiences with people of a particular profession, the subject will tend to significantly overestimate the actual probability that the given individual belongs to that profession. In overweighting the representative description, the subject underweights the statistical base rate evidence of the small fraction of the population belonging to that profession.

An important manifestation of the representativeness heuristic, discussed in detail by Kahneman and Tversky (1974) is that people see patterns in random sequences. They find that for someone to assign a high probability to the event that some data sample is generated by a random walk process, the sample must be considered representative of a random walk. Kahneman and Tversky note that "a representative sample is one in which the essential characteristics of the parent population are represented not only globally in the entire sample, but also locally in each of its parts" (p.36).

The reason this is important for our analysis is that *locally*, a sample drawn from a random walk may not look at all like a random walk. Suppose that we write a positive earnings change as H and a negative one as L; even if changes are random, we will expect to see sequences of the form HLLHLHL or LHLLLLL. These sequences do not reflect the properties of the parent population, because they contain either too many switches or else runs that are too long. The investor will not regard these sequences as drawn from a random walk, but rather as having been drawn from models whose population characteristics match the sample features closely: in the case of the two sequences given above, obvious examples of such models would be a mean-reverting model and a trending model respectively: in other words, Regimes 1 and 2 of our framework. Of course, representativeness does not imply that these are the *only* non-random patterns that people will see; however, if one believes that investors look for one of the two most obvious patterns – trending and mean-reverting — then our model fits the psychological evidence very closely.

The root of the problem facing investors is that there is very little theoretical guidance as to the process that earnings might follow. It is therefore natural that investors will look to observed earnings to try to infer what the process is. The bias implied by the representativeness heuristic is that investors will try to infer the process *too quickly*, on the basis of *too short a sample*, and hence will make erroneous forecasts. As Rabin (1996) puts it in his discussion of this heuristic, "When the underlying probability distribution generating observed sequences is uncertain, the fallacy leads people to

over-infer the probability distribution from short sequences.”

A different idea identified by several psychologists including Ward Edwards (1968), is conservatism. Conservatism states that once individuals have formed an impression, they are slow to change that impression in the face of new evidence. Edwards benchmarked a subject’s reaction to new evidence against that of the idealized rational Bayesian in experiments where the true normative value of a piece of evidence was well-defined. In his experiments, individuals updated their posteriors in the right direction, but *by too little* in magnitude relative to the rational Bayesian benchmark. This finding of conservatism is actually more pronounced, the more objectively useful is the new evidence. In Edwards’ own words:

“It turns out that opinion change is very orderly, and usually proportional to numbers calculated from the Bayes Theorem – but it is insufficient in amount. A conventional first approximation to the data would say that it takes anywhere from two to five observations to do one observation’s worth of work in inducing a subject to change his opinions” (p.359).

Regime 1 of our model is consistent with Edwards’ experiments. As individuals see a positive earnings shock or a piece of news that forecasts higher earnings, they rightly increase their forecasts for next period, but by too little. Under regime 1, investors too quickly dismiss a new earnings number as partly an aberration, i.e., as having a big temporary component, and are too slow to abandon their prior earnings estimates. Edwards would describe this behavior in Bayesian terms as a failure to properly aggregate the information in the new earnings number with their own prior information to form a new posterior earnings estimate.

In addition to being consistent with the evidence from behavioral psychology, our model is also consistent with some stock trading experiments conducted by Andreassen and Kraus (1992). They show subjects – who are university undergraduates untrained in finance – a time series of stock prices and ask them to trade at the prevailing price. After subjects trade, the next realization of price appears, and they can trade again. Trades do not affect prices: subjects trade with a time series rather than with each other. Stock prices are rescaled real stock prices taken from the financial press, and sometimes modified by the introduction of trends.

Andreassen and Kraus’ basic findings are as follows. Subjects generally “track prices”, i.e. sell when prices rise and buy when prices fall, even when the series they are offered is a random walk. This is the fairly universal mode of behavior, which is consistent with underreaction to news in markets.

However, when subjects are given a series of data with an ostensible trend, they reduce tracking, i.e. they trade less in response to price movements. It is not clear from Andreassen and Kraus' results whether subjects actually switch from bucking trends to chasing them as they do in our regime 2.

A recent paper by DeBondt (1993) nicely complements Andreassen and Kraus' findings. Using a combination of classroom experiments and investor surveys, DeBondt finds strong evidence that people extrapolate past trends. In one case, he asks subjects to forecast future stock price levels after showing them past stock prices over unnamed periods. He also analyses a sample of regular forecasts of the Dow Jones Index from a survey of members of the American Association of Individual Investors. In both cases, the forecast change in price level is higher following a series of previous price increases than following price decreases, suggesting that investors may indeed chase trends once they appear.

3.3 A Formal Model

We now present a mathematical model of the investor behavior described above, and check that the intuition can be formally justified. Suppose that earnings at time t are $N_t = N_{t-1} + y_t$, where y_t is the shock to earnings at time t , which can take one of two values, $+y$, or $-y$. The investor believes that the value of y_t is determined by one of two models, Model 1 or Model 2, depending on the "state" or "regime" of the economy. Models 1 and 2 have the same structure: they are both Markov processes, in the sense that the value taken by y_t depends only on the value taken by y_{t-1} . The essential difference between the two processes lies in the transition probabilities. To be precise, the transition matrices for the two models are:

| Model 1 | $y_{t+1} = y$ | $y_{t+1} = -y$ | Model 2 | $y_{t+1} = y$ | $y_{t+1} = -y$ |
|------------|---------------|----------------|------------|---------------|----------------|
| $y_t = y$ | π_H | $1 - \pi_H$ | $y_t = y$ | π_L | $1 - \pi_L$ |
| $y_t = -y$ | $1 - \pi_H$ | π_H | $y_t = -y$ | $1 - \pi_L$ | π_L |

The key is that π_H is small - we shall think of π_H being in the range $0 < \pi_H < 0.5$ - and of π_L as being large, $0.5 < \pi_L < 1$. In other words, under Model 1 a positive shock is likely to be reversed; under Model 2 a positive shock is more likely to be followed by another positive shock.

The investor is convinced that he knows the parameters π_H and π_L ; he is also sure that he is right about the underlying process controlling the

switching from one regime to another, or equivalently from Model 1 to Model 2. It too is Markov, so that the state of the world today depends only on the state of the world in the previous period. The transition matrix is:

$$\begin{array}{c|cc} & s_{t+1} = 1 & s_{t+1} = 2 \\ \hline s_t = 1 & 1 - \lambda_1 & \lambda_1 \\ s_t = 2 & \lambda_2 & 1 - \lambda_2 \end{array}$$

The state of the world at time t is written s_t . If $s_t = 1$, we are in the first regime and the earnings shock in period t , y_t is generated by Model 1; similarly if $s_t = 2$, we are in the second regime and the earnings shock is generated by Model 2. The parameters λ_1 and λ_2 determine the probabilities of transition from one state to another. We focus particularly on small λ_1 and λ_2 , which means that transitions from one state to another occur rarely. Certainly we shall assume $\lambda_1 + \lambda_2 < 1$. Typically, we will think of λ_1 being smaller than λ_2 . Since the unconditional probability of being in state 1 is $\frac{\lambda_2}{\lambda_1 + \lambda_2}$, this implies that the investor thinks of Model 1 as being on average more likely than Model 2. Our results do not depend, however, on λ_1 being smaller than λ_2 . The effects that we will document can also obtain if $\lambda_1 > \lambda_2$.

In order to value the security, the investor needs to forecast earnings into the future. Since the model he is using dictates that earnings at any time are generated by one of two regimes, the investor sees his task as trying to understand which of the two regimes is currently governing earnings. Since he observes earnings each period, he uses that information to make as good a guess as possible about which regime he is in. In particular, at time t , having observed the earnings shock y_t , he calculates π_t , the probability that y_t was generated by Model 1, using the new data to update his estimate from the previous period, π_{t-1} . Formally, $\pi_t = \Pr(s_t = 1 | y_t, y_{t-1}, \pi_{t-1})$. We suppose that the updating follows Bayes Rule, so that:

$$\pi_{t+1} = \frac{((1 - \lambda_1)\pi_t + \lambda_2(1 - \pi_t)) \Pr(y_{t+1} | s_{t+1} = 1, y_t)}{((1 - \lambda_1)\pi_t + \lambda_2(1 - \pi_t)) \Pr(y_{t+1} | s_{t+1} = 1, y_t) + (\lambda_1\pi_t + (1 - \lambda_2)(1 - \pi_t)) \Pr(y_{t+1} | s_t = 2, y_t)}$$

In particular if the shock to earnings in period $t + 1$, y_{t+1} is the same as the shock in period t , y_t , the investor updates π_{t+1} from π_t using:

$$\pi_{t+1} = \frac{((1 - \lambda_1)\pi_t + \lambda_2(1 - \pi_t))\pi_H}{((1 - \lambda_1)\pi_t + \lambda_2(1 - \pi_t))\pi_H + (\lambda_1\pi_t + (1 - \lambda_2)(1 - \pi_t))\pi_L}$$

and it can be shown that in this case, $\pi_{t+1} < \pi_t$. In other words, the investor puts more weight on Model 2 if he sees two consecutive shocks of the same sign. Similarly, if the shock in period $t + 1$ has the opposite sign to that in period t ,

$$\pi_{t+1} = \frac{((1 - \lambda_1)\pi_t + \lambda_2(1 - \pi_t))(1 - \pi_H)}{((1 - \lambda_1)\pi_t + \lambda_2(1 - \pi_t))(1 - \pi_H) + (\lambda_1\pi_t + (1 - \lambda_2)(1 - \pi_t))(1 - \pi_L)}$$

To aid intuition with the workings of the model described here, we now present a simple illustrative example. Suppose that in period 0, the shock to earnings y_0 is positive and the probability assigned to Model 1 by the investor, i.e. π_0 is 0.5. The following Table gives a randomly generated earnings stream for the next 20 periods and presents the investor's belief π_t that the time t shock to earnings is generated by Model 1. The particular parameter values chosen here are: $\pi_H = \frac{1}{3} < \frac{3}{4} = \pi_L$, and $\lambda_1 = 0.1 < 0.3 = \lambda_2$. Note that the earnings stream below is generated using the true process for earnings, namely a random walk:

| t | y_t | π_t | t | y_t | π_t |
|-----|-------|---------|-----|-------|---------|
| 0 | y | .50 | | | |
| 1 | $-y$ | .80 | 11 | y | .74 |
| 2 | y | .90 | 12 | y | .56 |
| 3 | $-y$ | .93 | 13 | y | .44 |
| 4 | y | .94 | 14 | y | .36 |
| 5 | y | .74 | 15 | $-y$ | .74 |
| 6 | $-y$ | .89 | 16 | y | .89 |
| 7 | $-y$ | .69 | 17 | y | .69 |
| 8 | y | .87 | 18 | $-y$ | .87 |
| 9 | $-y$ | .92 | 19 | y | .92 |
| 10 | y | .94 | 20 | y | .72 |

Notice that in periods 0 through 4 positive shocks to earnings alternate with negative shocks. Since Model 1 stipulates that earnings shocks are likely to be reversed in the following period, we observe an increase in π_t , the probability that Model 1 is generating the earnings shock at time t , rising to a high of 0.94 in period 4. From periods 10 to 14, we observe five successive positive shocks; since this is behavior typical of that specified by Model 2, π_t falls through period 14 to a low of 0.36. One feature that is evident in the above example, is that π_t rises if the earnings shock in period t has the

opposite sign from that in period $t - 1$ and falls if the shock in period t has the same sign as that in period $t - 1$.

To end this section, we return briefly to an issue raised in Section 3.1. One apparent difficulty with the framework we use is that while the psychological evidence may suggest the expectations formation process we propose, after a sufficiently long data series the investor will learn that the right model for earnings is a random walk, and not a regime switching model.

As more data is received, a Bayesian observer will place more and more weight on the true random walk model. However, this rational learning does not occur very quickly. Suppose that the investor has prior views assigning probability 0.5 to the random walk model for earnings and probability 0.5 to a switching model with the same parameters as those used in the above Table, i.e., $\lambda_1 = 0.1$, $\lambda_2 = 0.3$, $\pi_H = \frac{1}{3}$, and $\pi_L = \frac{3}{4}$. Suppose that we show the investor 5 years of quarterly earnings data, i.e., 20 observations drawn from a random walk, and then calculate the log Bayes posterior odds ratio in favor of the random walk,

$$\log \frac{\Pr(\text{model} = \text{random walk} | \text{data})}{\Pr(\text{model} = \text{regime-switching} | \text{data})}.$$

We repeat this experiment 1000 times, in other words using 1000 different random sequences of 20 earnings changes. We find that even after 5 years of data, a sizeable 34% of the time investors *preferred* the regime switching model, i.e., produced log posterior odds that were negative. Moreover, in many cases, the log posterior ratio lies only slightly above zero, indicating a close tie between the models. To be specific, 80% of the time, investors consider the switching model at least half as likely as the random walk model. In other words, 80% of the time, investors attach a posterior probability of at least $\frac{1}{3}$ to the switching model. If investors give even this much weight to the switching model in forming their expectations, the effects we document in the next section will still figure prominently. Furthermore, the regime-switching model we use to capture the psychological evidence is somewhat stylized for simplicity: a more flexible model may fare even better against the random walk alternative⁶.

⁶For instance, a more general model consistent with the psychological evidence would be one of the form $N_t = \rho N_{t-1} + \varepsilon_t$, where N_t are earnings in period t , and where the investor believes that the growth rate of earnings ρ is not only unknown but changing over time, so that it needs to be re-estimated each period. Such a model would be considerably harder to reject in favor of the random walk.

The above experiment assumed that the investor uses Bayesian updating to decide between the random walk and regime switching models. In practice, however, people may take far longer to learn the true model than is suggested by the experiment. Evidence again comes from the psychology literature, which is teeming with studies documenting the substantial differences between people's actual learning behavior, and that dictated by Bayes' rule. A very recent contribution to this literature, Rabin and Schrag (1996), is particularly relevant here. Those authors start with a widely reported psychological phenomenon known as *confirmatory bias*: suppose that there are two hypotheses, and that the individual currently favors one over the other. Experiments have repeatedly shown that when subsequent evidence relevant to these hypotheses appears, the individual will sometimes misinterpret evidence that contradicts his favored hypothesis as evidence that in fact supports it. Rabin and Schrag show that this behavior can have important consequences: for instance, the individual may, with positive probability eventually come to believe with near certainty in the false hypothesis even after an infinite amount of information.

The implications of this study for our work are clear: if the work of Kahneman, Tversky and Edwards suggests that the investor may prefer a regime-switching model to a random walk model, the work of Rabin and Schrag shows that it may not be at all easy for the investor to learn the correct model, even after a large amount of data.

4 Model Solution and Empirical Implications

We now analyze the implications of our model for prices. Since our model has a representative agent, the price of the security is simply the value of the security as perceived by the investor. In other words:

$$P_t = E_t \left\{ \frac{N_{t+1}}{1 + \delta} + \frac{N_{t+2}}{(1 + \delta)^2} + \dots \right\}$$

Note that the expectations in this expression are the expectations of the investor who does not realize that the true process for earnings is a random walk. Indeed, if the investor did realize this, the series above would be simple enough to evaluate: since under a random walk, $E_t(N_{t+j}) = N_t$, price equals $\frac{N_t}{\delta}$. In our model, price deviates from this true value because the investor does not use the random walk model to forecast earnings, but rather some

combination of Models 1 and 2, neither of which is a random walk. The following Proposition, proved in the Appendix, summarizes the behavior of prices in this context, and shows that they depend on the state variables in a particularly simple way.

PROPOSITION 1: If the investor believes that earnings are generated by the regime-switching model described in Section 3, then prices satisfy:

$$P_t = \frac{N_t}{\delta} + y_t(p_1 - p_2\pi_t)$$

where p_1 and p_2 are constants which depend on π_H , π_L , λ_1 , and λ_2 . The full expressions for p_1 and p_2 are given in the Appendix⁷. \diamond

The formula for P_t has a very simple interpretation. The first term, $\frac{N_t}{\delta}$, is the price that would obtain if the investor used the true random walk process to forecast earnings. The second term, $y_t(p_1 - p_2\pi_t)$, gives the deviation of price from this fundamental value. Later in this section we look at the range of values of π_H , π_L , λ_1 , and λ_2 which allow the price function in Proposition 1 to exhibit both underreaction and overreaction to earnings news. In fact Proposition 2 below gives sufficient conditions on p_1 and p_2 to ensure that this is the case. For the next few paragraphs, in the run-up to Proposition 2, we will forsake mathematical rigor in order to build intuition for those sufficient conditions.

First, note that if the price function P_t is to exhibit underreaction to earnings news on average, then p_1 cannot be too large in relation to p_2 . For suppose the latest earnings shock y_t was a positive one. Underreaction means that on average, the stock price will not react sufficiently to this shock, leaving the price below fundamental value. This means that on average, $y(p_1 - p_2\pi_t)$, the deviation from fundamental value must be negative. If π_{avg} denotes an average value of π , this implies that we must have $p_1 < p_2\pi_{avg}$. This is the sense in which p_1 cannot be too large in relation to p_2 .

On the other hand, if P_t is also to display overreaction to sequences of similar earnings news, then p_1 cannot be too small in relation to p_2 . Suppose that the investor has just observed a series of good earnings shocks. Then overreaction would require that price now be above fundamental value.

⁷It is difficult to prove general results about p_1 and p_2 , although numerical computations show that over most of the range of values of π_L , π_H , λ_1 and λ_2 we are interested in, p_1 and p_2 are both positive.

Moreover, we know that after a series of shocks of the same sign, π_t is normally low, indicating a low weight on Model 1 and a high weight on Model 2. If we write π_{low} to represent a typical low value of π_t , overreaction then requires that $y(p_1 - p_2\pi_{low})$ be positive, or that $p_1 > p_2\pi_{low}$. This is the sense in which p_1 cannot be too small in relation to p_2 . Putting the two conditions together, we obtain:

$$p_2\pi_{low} < p_1 < p_2\pi_{avg}$$

In Proposition 2, we provide sufficient conditions on p_1 and p_2 for both underreaction and overreaction to obtain, and their form is very similar to what we have just obtained. In fact, the argument in Proposition 2 is essentially the one we have just made in the preceding paragraphs, although some effort is required to make the reasoning rigorous.

Before stating the Proposition, we repeat the definitions of overreaction and underreaction that were presented in Section 2. Overreaction can be thought of as meaning that the expected return following a sufficiently large number of positive shocks should be lower than the expected return following the same number of successive negative shocks. In other words, there exists some large number J , such that for all $j \geq J$,

$$E_t(P_{t+1} - P_t | y_t = y_{t-1} = \dots = y_{t-j} = y) - E_t(P_{t+1} - P_t | y_t = y_{t-1} = \dots = y_{t-j} = -y) < 0$$

Underreaction means that the expected return following a positive shock should exceed the expected return following a negative shock. In other words,

$$E_t(P_{t+1} - P_t | y_t = +y) - E_t(P_{t+1} - P_t | y_t = -y) > 0$$

Proposition 2 below provides sufficient conditions on π_H, π_L, λ_1 and λ_2 for these two inequalities to hold⁸.

PROPOSITION 2: If the underlying parameters π_H, π_L, λ_1 and λ_2 satisfy:

$$\underline{k}p_2 < p_1 < \bar{k}p_2$$

⁸For the purposes of Proposition 2, we have made two simplifications in our mathematical formulation of under- and overreaction. First, we examine the absolute price change $P_{t+1} - P_t$ rather than the actual return. Second, the good news is presumed here to be the event $y_t = +y$, i.e., a positive change in earnings, rather than better than expected earnings. Since in our model, the expected change in earnings $E_t(y_{t+1})$, always lies between $-y$ and $+y$, a positive earnings change will in fact be a positive surprise. Therefore the results are largely the same in the two cases. In the simulations in Section 4, we calculate returns in the usual way, and condition on earnings surprises as well as raw earnings changes.

then the price function in Proposition 1 exhibits both underreaction and overreaction to earnings. \underline{k} and \bar{k} are positive constants that depend on π_H , π_L , λ_1 , and λ_2 - the full expressions are given in the Appendix. \diamond

Note that the form of the conditions conforms exactly to the intuition presented earlier: p_1 must be high relative to p_2 if price is going to rise above fundamental value after long series of good earnings news; and yet p_1 must be low relative to p_2 so that on average, price underreacts to earnings announcements.

We now examine the range of values of the fundamental parameters π_L , π_H , λ_1 , and λ_2 for which the sufficient conditions for both underreaction and overreaction are satisfied. Since the conditions in Proposition 2 are somewhat involved, we evaluate them numerically for a large range of values of the four underlying parameters. Figure 1 illustrates one such exercise. We start by fixing $\lambda_1 = 0.1$ and $\lambda_2 = 0.3$. These numbers are small to ensure that regime switches do not occur very often; and $\lambda_2 > \lambda_1$, to represent the investor's belief that Model 1 is a better model for earnings than is Model 2, in the sense that the world is in the Model 1 regime more often than in the Model 2 regime.

Now that λ_1 and λ_2 have been initially fixed, we want to know for what range of values of π_H and π_L the conditions for underreaction and overreaction hold. By the nature of π_H and π_L , it makes sense to restrict their values to the ranges $0 < \pi_H < 0.5$, and $0.5 < \pi_L < 1$. Next, we evaluate the conditions in Proposition 2 for pairs of (π_H, π_L) where π_H ranges from 0 to 0.5 at intervals of 0.01 and π_L ranges from 0.5 to 1, again at intervals of 0.01.

Figure 1 marks with a "+" all the pairs for which the sufficient conditions hold. We see that underreaction and overreaction hold for a wide range of values. On the other hand, it is not a trivial result: there are many parameter values for which at least one of the two phenomena does not hold.

The figure shows that the sufficient conditions do not hold if both π_H and π_L are near the high end of their feasible ranges, or if both π_H and π_L are near the low end of their ranges. The reason for this is the following: suppose both π_H and π_L are high. This means that whatever the regime, the investor believes that shocks are relatively likely to be followed by another shock of the same sign. The consequence of this is that overreaction certainly obtains, although underreaction may not: following a positive shock, the investor on average expects another positive shock; since the true process is a random walk, returns are on average negative. Hence the average return following

a positive shock is lower than that following a negative shock, which is a characterization of overreaction.

On the other hand, if π_H and π_L are both at the low end, the investor believes that shocks are relatively likely to be reversed, regardless of the regime: this leads to underreaction, but overreaction may not hold.

To confirm this intuition, we plot in Figure 2 separately the ranges of (π_H, π_L) pairs for which underreaction and overreaction hold. The intersection of the two regions is the original one shown in Figure 1. Figure 2 confirms the intuition that if π_H and π_L are on the higher side, overreaction obtains, but underreaction may not.

Figures 3 and 4 perform the same exercise with different initial values for λ_1 and λ_2 . In all cases, there are non-trivial ranges of (π_H, π_L) pairs for which the sufficient conditions hold.

4.1 Some simulation experiments

One way to test the model we have developed is to try to replicate the empirical findings of the papers discussed in Section 2 using long time series of prices and earnings simulated from our model. That is, we pick values for π_H, π_L, λ_1 , and λ_2 , and generate a long series of 10,000 price observations using the formula for prices derived earlier:

$$P_t = \frac{N_t}{\delta} + y_t(p_1 - p_2\pi_t),$$

and calculate average returns conditional on previous realizations of earnings or returns, paralleling the empirical work on actual data⁹.

One natural exercise would be to calculate the difference between average returns in the period after a positive earnings shock r_+ and average returns in the period after a negative earnings shock, r_- . This should be positive, to match our definition of underreaction in Section 2, and to replicate the empirical findings of studies such as that of Bernard and Thomas (1992).

Moreover, the literature on overreaction that was reviewed in Section 2 suggested that investors become overly optimistic following a consistent pattern of good news, and too pessimistic after a similar pattern of bad news.

⁹Developing analytical formulae for the population values of the statistics used in the empirical papers is intractable in our model. By specifying parameters and simulating long time series, we hope to approximate these population values with reasonable accuracy.

Consistent with this, and with our definition of overreaction, we would expect that the average return in the period following a long series of consecutive positive shocks should be lower than the average return following a similarly long series of negative shocks. This motivates us to calculate the difference between average returns following two positive shocks and two negative shocks, $r_{++} - r_{--}$, following three consecutive shocks of the same sign, $r_{+++} - r_{---}$, and finally, following four consecutive shocks, $r_{++++} - r_{----}$.

What we hope to see, and what we indeed find is that these quantities decline as we condition on a progressively longer string of earnings shocks of the same sign, indicating a transition from underreaction to overreaction.

The values chosen for the underlying variables are:

$$\begin{aligned}\lambda_1 &= 0.1, \lambda_2 = 0.3 \\ \pi_H &= \frac{1}{3}, \pi_L = \frac{3}{4}\end{aligned}$$

and on the basis of a simulated sample of 10,000 data points, the results for the difference in average returns in the case of sorting by earnings are:

| Earnings Sort | |
|-----------------------|--------|
| $r_+ - r_-$ | .0025 |
| $r_{++} - r_{--}$ | .0011 |
| $r_{+++} - r_{---}$ | -.0003 |
| $r_{++++} - r_{----}$ | -.0016 |

The average return following a positive earnings shock is greater than the average return following a negative shock, indicating underreaction in the manner of Bernard and Thomas. As we increase the number of shocks of the same sign that we condition on, however, the difference in average returns decreases through zero and turns negative once we condition on three or more shocks of the same sign, indicating overreaction¹⁰.

Some of the other studies in Section 2 calculated average returns conditional not on previous earnings realizations, but on previous realizations of returns. For instance, the positive autocorrelation found by Cutler, Poterba, and Summers (1991) requires that average returns in periods following high

¹⁰As mentioned in Section 3, a more natural definition of a positive earnings shock may be an earnings realization that is higher *than expected*, rather than merely a positive earnings change. In our model, however, these two variables are highly correlated, so that the results are very similar in the two cases.

returns should be higher than average returns in periods following low returns, so that the difference in average returns is positive. Once we condition on cumulative returns over longer periods, though, studies of overreaction including that of Poterba and Summers (1988) suggest that this difference in average returns will turn negative.

In an attempt to replicate these studies, we use past returns as a sorting mechanism rather than earnings. In other words, we take the simulated series of 10,000 price observations and calculate non-overlapping cumulative returns over say three periods. We then rank these cumulative returns, and calculate the average return in the period following those three period intervals with cumulative returns in the top decile, τ_{HHH} , minus the average return in the period following those three period intervals with cumulative returns in the bottom decile, τ_{LLL} . We calculate this for cumulative returns over intervals of one, two, three and four periods, i.e. $\tau_H - \tau_L$, $\tau_{HH} - \tau_{LL}$, $\tau_{HHH} - \tau_{LLL}$, and $\tau_{HHHH} - \tau_{LLLL}$.

Again, what we hope to find, and indeed do find, is that these numbers decrease, with $\tau_H - \tau_L$ positive (positive correlation, underreaction in the manner of Cutler, Poterba, and Summers (1991)), and $\tau_{HHHH} - \tau_{LLLL}$ negative (overreaction, negative correlation as found in Poterba and Summers (1988)). Using the same set of values for the underlying parameters, the results are:

| Returns Sort | |
|-----------------------------|--------|
| $\tau_H - \tau_L$ | .0024 |
| $\tau_{HH} - \tau_{LL}$ | .0013 |
| $\tau_{HHH} - \tau_{LLL}$ | -.0012 |
| $\tau_{HHHH} - \tau_{LLLL}$ | -.0022 |

The average return following one-period returns in the top decile is higher than the average return following one-period returns in the bottom decile. Once we start sorting on cumulative returns over two or more periods, however, the difference in average returns falls through zero and turns negative.

Finally, we can also use our simulated data to try to replicate one more widely reported empirical finding, namely the predictive power of earnings-price ratios for returns. We calculate the difference between average returns in periods following higher than average earnings-price ratios and average returns in periods following lower than average earnings-price ratios. We repeat this calculation for non-overlapping multi-period returns. In each

case, we find the difference in average returns to be positive, replicating the predictive power of the earnings/price ratio:

| E/P sort | |
|---|--------|
| $r_{t+1} \Big \frac{N_t}{P_t} > \frac{\bar{N}}{\bar{P}} - r_{t+1} \Big \frac{N_t}{P_t} < \frac{\bar{N}}{\bar{P}}$ | 0.0016 |
| $r_{t+1,t+2} \Big \frac{N_t}{P_t} > \frac{\bar{N}}{\bar{P}} - r_{t+1,t+2} \Big \frac{N_t}{P_t} < \frac{\bar{N}}{\bar{P}}$ | 0.0015 |
| $r_{t+1,t+3} \Big \frac{N_t}{P_t} > \frac{\bar{N}}{\bar{P}} - r_{t+1,t+3} \Big \frac{N_t}{P_t} < \frac{\bar{N}}{\bar{P}}$ | 0.0013 |
| $r_{t+1,t+4} \Big \frac{N_t}{P_t} > \frac{\bar{N}}{\bar{P}} - r_{t+1,t+4} \Big \frac{N_t}{P_t} < \frac{\bar{N}}{\bar{P}}$ | 0.0015 |

Here, $r_{t+1,t+4}$ refers to the cumulative return $r_{t+1} + r_{t+2} + r_{t+3} + r_{t+4}$.

5 Conclusion

In this paper, we have presented a parsimonious model of investor sentiment, or of how investors form expectations of future earnings. The model we propose is motivated to a large extent by research on the heuristics that people use in assessing the probabilities of outcomes. We show that the interaction of the investor's belief formation process with the true model for earnings can generate two distinct empirical regularities, namely underreaction to news, and overreaction to consistent good or bad news. The result obtains for a wide range of parameter values in the underlying model.

6 Appendix

PROPOSITION 1: If the investor believes that earnings are generated by the regime-switching model described earlier, then prices satisfy:

$$P_t = \frac{N_t}{\delta} + y_t(p_1 - p_2\pi_t)$$

where p_1 and p_2 are given by the following expressions:

$$\begin{aligned} p_1 &= \frac{1}{\delta}(\gamma'_0(1+\delta)[I(1+\delta) - Q]^{-1}Q\gamma_1) \\ p_2 &= -\frac{1}{\delta}(\gamma'_0(1+\delta)[I(1+\delta) - Q]^{-1}Q\gamma_2) \end{aligned}$$

where:

$$\begin{aligned} \gamma'_0 &= (1, -1, 1, -1) \\ \gamma'_1 &= (0, 0, 1, 0) \\ \gamma'_2 &= (1, 0, -1, 0) \\ Q &= \begin{pmatrix} (1-\lambda_1)\pi_H & (1-\lambda_1)(1-\pi_H) & \lambda_2\pi_H & \lambda_2(1-\pi_H) \\ (1-\lambda_1)(1-\pi_H) & (1-\lambda_1)\pi_H & \lambda_2(1-\pi_H) & \lambda_2\pi_H \\ \lambda_1\pi_H & \lambda_1(1-\pi_L) & (1-\lambda_2)\pi_L & (1-\lambda_2)(1-\pi_L) \\ \lambda_1(1-\pi_L) & \lambda_1\pi_L & (1-\lambda_2)(1-\pi_L) & (1-\lambda_2)\pi_L \end{pmatrix} \end{aligned}$$

PROOF OF PROPOSITION 1:

The price will simply equal the value as gauged by the uninformed investors which we can calculate from the present value formula:

$$P_t = E_t\left\{\frac{N_{t+1}}{1+\delta} + \frac{N_{t+2}}{(1+\delta)^2} + \dots\right\}$$

Since:

$$\begin{aligned} E_t(N_{t+1}) &= N_t + E_t(y_{t+1}) \\ E_t(N_{t+2}) &= N_t + E_t(y_{t+1}) + E_t(y_{t+2}), \text{ and so on} \end{aligned}$$

we have:

$$P_t = \frac{1}{\delta}\left\{N_t + E_t(y_{t+1}) + \frac{E_t(y_{t+2})}{1+\delta} + \frac{E_t(y_{t+3})}{(1+\delta)^2} + \dots\right\}$$

So the key is calculate $E_t(y_{t+j})$. To get this, define:

$$\pi^{t+j} = (\pi_1^{t+j}, \pi_2^{t+j}, \pi_3^{t+j}, \pi_4^{t+j})$$

where:

$$\begin{aligned}
\pi_1^{t+j} &= \Pr(s_{t+j} = 1, y_{t+j} = y_t | \Phi_t) \\
\pi_2^{t+j} &= \Pr(s_{t+j} = 1, y_{t+j} = -y_t | \Phi_t) \\
\pi_3^{t+j} &= \Pr(s_{t+j} = 2, y_{t+j} = y_t | \Phi_t) \\
\pi_4^{t+j} &= \Pr(s_{t+j} = 2, y_{t+j} = -y_t | \Phi_t)
\end{aligned}$$

where Φ_t is the investor's information set at time t , consisting of the observed earnings series (y_0, y_1, \dots, y_t) which can be summarized as (y_t, π_t) .

Note that

$$\begin{aligned}
\Pr(y_{t+j} = y_t | \Phi_t) &= \pi_1^{t+j} + \pi_3^{t+j} = \bar{\gamma}' \pi^{t+j} \\
\bar{\gamma} &= (1, 0, 1, 0)
\end{aligned}$$

The key insight is the following:

$$\pi^{t+j} = Q \pi^{t+j-1}$$

where Q is just the transpose of the transition matrix for the states (s_{t+j}, y_{t+j}) , i.e.

$$Q' = \begin{array}{ccccc}
& \begin{matrix} (1) \\ (2) \\ (3) \\ (4) \end{matrix} & \begin{matrix} (1) \\ (2) \\ (3) \\ (4) \end{matrix} & \begin{matrix} (2) \\ (1) \\ (1) \\ (2) \end{matrix} & \begin{matrix} (3) \\ (1) \\ (1) \\ (2) \end{matrix} & \begin{matrix} (4) \\ (1) \\ (1) \\ (2) \end{matrix} \\
\begin{matrix} (1) \\ (2) \\ (3) \\ (4) \end{matrix} & & & & & \\
(1) & (1 - \lambda_1)\pi_H & (1 - \lambda_1)(1 - \pi_H) & \lambda_1\pi_L & \lambda_1(1 - \pi_L) & \\
(2) & (1 - \lambda_1)(1 - \pi_H) & (1 - \lambda_1)\pi_H & \lambda_1(1 - \pi_L) & \lambda_1\pi_L & \\
(3) & \lambda_2\pi_H & \lambda_2(1 - \pi_H) & (1 - \lambda_2)\pi_L & (1 - \lambda_2)(1 - \pi_L) & \\
(4) & \lambda_2(1 - \pi_H) & \lambda_2\pi_H & (1 - \lambda_2)(1 - \pi_L) & (1 - \lambda_2)\pi_L &
\end{array}$$

Therefore:

$$\pi^{t+j} = Q^j \pi^t = Q^j \begin{pmatrix} \pi_t \\ 0 \\ 1 - \pi_t \\ 0 \end{pmatrix}$$

(Note the distinction between π_t and π^t). Hence,

$$\begin{aligned}
\Pr(y_{t+j} = y_t | \Phi_t) &= \bar{\gamma}' Q^j \pi^t \\
&\text{and} \\
E_t(y_{t+j} | \Phi_t) &= y(\bar{\gamma}' Q^j \pi^t) + (-y)(\underline{\gamma}' Q^j \pi^t) \\
\underline{\gamma}' &= (0, 1, 0, 1)
\end{aligned}$$

Substituting this into the original formula for price gives:

$$\begin{aligned}
p_0 &= \frac{1}{\delta} \\
p_1 &= \frac{1}{\delta} (\gamma'_0(1+\delta)[I(1+\delta) - Q]^{-1} Q \gamma_1) \\
p_2 &= -\frac{1}{\delta} (\gamma'_0(1+\delta)[I(1+\delta) - Q]^{-1} Q \gamma_2) \\
\gamma'_0 &= (1, -1, 1, -1) \\
\gamma'_1 &= (0, 0, 1, 0) \\
\gamma'_2 &= (1, 0, -1, 0)
\end{aligned}$$

PROPOSITION 2: If the underlying parameters π_H, π_L, λ_1 and λ_2 satisfy:

$$kp_2 < p_1 < \bar{k}p_2$$

where:

$$\begin{aligned}
\underline{k} &= \pi + \frac{1}{2} \overline{\Delta}(\pi) \\
\bar{k} &= \pi^e + \frac{1}{2} (c_1 + c_2 \pi_*) \\
c_1 &= \frac{\overline{\Delta}(\pi) \bar{\pi} - \underline{\Delta}(\bar{\pi}) \pi}{\bar{\pi} - \pi} \\
c_2 &= \frac{\underline{\Delta}(\bar{\pi}) - \overline{\Delta}(\pi)}{\bar{\pi} - \pi} \\
\pi_* &= \begin{cases} \bar{\pi}^e & \text{if } c_2 < 0 \\ \pi^e & \text{if } c_2 \geq 0 \end{cases}
\end{aligned}$$

then both $f(\pi) < 0$ and $E_\pi(f(\pi)) > 0$ hold, i.e. the conditions for both underreaction and overreaction are satisfied. (Functions and variables which have not appeared in the analysis before will be defined in the proof).

PROOF OF PROPOSITION 2:

Before we enter the main argument of the proof, we present a short discussion of the behavior of π_t , the probability assigned by the investor at time t of being in Regime 1. Suppose that the earnings shock at time $t + 1$ is of the opposite sign to the shock in period t . Let the function $\overline{\Delta}(\pi_t)$ denote the increase in the probability assigned to being in Regime 1, i.e.,:

$$\begin{aligned}
\overline{\Delta}(\pi) &= \pi_{t+1} - \pi_t |_{y_{t+1} = -y_t, \pi_t = \pi} \\
&= \frac{((1 - \lambda_1)\pi + \lambda_2(1 - \pi))(1 - \pi_H)}{((1 - \lambda_1)\pi + \lambda_2(1 - \pi))(1 - \pi_H) + ((\lambda_1\pi + (1 - \lambda_2)(1 - \pi))(1 - \pi_L))} - \pi
\end{aligned}$$

Similarly, the function $\underline{\Delta}(\pi)$ measures the size of the fall in π_t if the period $t + 1$ earnings shock should be the same sign as that in period t .

$$\begin{aligned}\underline{\Delta}(\pi) &= \pi_t - \pi_{t+1} |_{y_{t+1}=y_t, \pi_t=\pi} \\ &= \pi - \frac{((1 - \lambda_1)\pi + \lambda_2(1 - \pi))\pi_H}{((1 - \lambda_1)\pi + \lambda_2(1 - \pi))\pi_H + ((\lambda_1\pi + (1 - \lambda_2)(1 - \pi))\pi_L)}\end{aligned}$$

By checking the sign of the second derivative, it is easy to see that both $\overline{\Delta}(\pi)$ and $\underline{\Delta}(\pi)$ are concave. More important though, is the sign of these functions over the interval $[0, 1]$. Under the conditions $\pi_H < \pi_L$, and $\lambda_1 + \lambda_2 < 1$, it is not hard to show that $\overline{\Delta}(\pi) \geq 0$ over an interval $[0, \overline{\pi}]$, and that $\underline{\Delta}(\pi) \geq 0$ over $[\underline{\pi}, 1]$, where $\underline{\pi}$ and $\overline{\pi}$ satisfy $0 < \underline{\pi} < \overline{\pi} < 1$.

The implication of this is that over the range $[\underline{\pi}, \overline{\pi}]$, the following is true: if the time t earnings shock has the same sign as the time $t + 1$ earnings shock, then $\pi_{t+1} < \pi_t$, or the probability assigned to Regime 2 rises. If the shocks are of different signs, however, then $\pi_{t+1} > \pi_t$, and Regime 1 will be seen as more likely.

Note that if $\pi_t \in [\underline{\pi}, \overline{\pi}]$, then $\pi_\tau \in [\underline{\pi}, \overline{\pi}]$ for $\forall \tau > t$. In other words, the investor's belief will always remain within this interval. If the investor sees a very long series of earnings shocks, all of which have the same sign, π_t will fall every period, tending towards a limit of $\underline{\pi}$. From the updating formulas, this means that $\underline{\pi}$ satisfies:

$$\underline{\pi} = \frac{((1 - \lambda_1)\underline{\pi} + \lambda_2(1 - \underline{\pi}))\pi_H}{((1 - \lambda_1)\underline{\pi} + \lambda_2(1 - \underline{\pi}))\pi_H + (\lambda_1\underline{\pi} + (1 - \lambda_2)(1 - \underline{\pi}))\pi_L}$$

Similarly, suppose that positive shocks alternate with negative ones for a long period of time. In this situation, π_t will rise every period, tending to the upper limit $\overline{\pi}$, which satisfies:

$$\overline{\pi} = \frac{((1 - \lambda_1)\overline{\pi} + \lambda_2(1 - \overline{\pi}))(1 - \pi_H)}{((1 - \lambda_1)\overline{\pi} + \lambda_2(1 - \overline{\pi}))(1 - \pi_H) + (\lambda_1\overline{\pi} + (1 - \lambda_2)(1 - \overline{\pi}))(1 - \pi_L)}$$

In the case of the parameters used for the Table in Section 3, $\underline{\pi} = 0.28$, and $\overline{\pi} = 0.95$.

There is no loss of generality in restricting the support of π_t to the interval $[\underline{\pi}, \overline{\pi}]$. Certainly, an investor may have prior beliefs that lie outside this interval, but with probability one, π_t will eventually belong to this interval, and will then stay within the interval forever.

We are now ready to begin the main argument of the proof. Underreaction means that the expected return following a positive shock should exceed the expected return following a negative shock. In other words,

$$E_t(P_{t+1} - P_t | y_t = +y) - E_t(P_{t+1} - P_t | y_t = -y) > 0$$

Overreaction means that the expected return following a large number of positive shocks is smaller than the expected return following a large number of negative shocks. In other words, there exists some large number J , such that for all $j \geq J$,

$$E_t(P_{t+1} - P_t | y_t = y_{t-1} = \dots = y_{t-j} = y) - E_t(P_{t+1} - P_t | y_t = y_{t-1} = \dots = y_{t-j} = -y) < 0$$

Proposition 2 provides sufficient conditions on p_1 and p_2 such that these two inequalities hold. A useful function for the purposes of our analysis is:

$$f(\pi) = E_t(P_{t+1} - P_t | y_t = +y, \pi_t = \pi) - E_t(P_{t+1} - P_t | y_t = -y, \pi_t = \pi)$$

The function $f(\pi)$ is the difference between the expected return following a positive shock and that following a negative shock, where we also condition on π_t equalling a specific value π . It is simple enough to write down an explicit expression for this function. Since:

$$P_{t+1} - P_t = p_0 y_{t+1} + (y_{t+1} - y_t)(p_1 - p_2 \pi_t) - y_t p_2 (\pi_{t+1} - \pi_t) - (y_{t+1} - y_t) p_2 (\pi_{t+1} - \pi_t),$$

we find:

$$\begin{aligned} E_t(P_{t+1} - P_t | y_{t+1} = +y, \pi_t = \pi) &= +y, \pi_t = \pi) \\ &= \frac{1}{2}(p_0 y + y p_2 \underline{\Delta}(\pi)) + \frac{1}{2}(-p_0 y - 2y(p_1 - p_2 \pi) - y p_2 \overline{\Delta}(\pi) + 2y p_2 \overline{\Delta}(\pi)) \\ &= y(p_2 \pi - p_1) + \frac{1}{2} y p_2 (\overline{\Delta}(\pi) + \underline{\Delta}(\pi)) \end{aligned}$$

Further it is easily checked that:

$$E_t(P_{t+1} - P_t | y_{t+1} = +y, \pi_t = \pi) = -E_t(P_{t+1} - P_t | y_{t+1} = -y, \pi_t = \pi)$$

and hence that:

$$f(\pi) = 2y(p_2 \pi - p_1) + y p_2 (\overline{\Delta}(\pi) + \underline{\Delta}(\pi)).$$

First, we show that a sufficient condition for overreaction is:

$$f(\underline{\pi}) < 0.$$

For suppose that this condition holds. Then this implies:

$$E_t(P_{t+1} - P_t | y_t = +y, \pi_t = \underline{\pi}) < E_t(P_{t+1} - P_t | y_t = -y, \pi_t = \underline{\pi})$$

Now as $j \rightarrow \infty$,

$$E_t(P_{t+1} - P_t | y_t = y_{t-1} = \dots = y_{t-j} = y) \rightarrow E_t(P_{t+1} - P_t | y_t = +y, \pi_t = \underline{\pi})$$

and

$$E_t(P_{t+1} - P_t | y_t = y_{t-1} = \dots = y_{t-j} = -y) \rightarrow E_t(P_{t+1} - P_t | y_t = -y, \pi_t = \underline{\pi}).$$

Therefore, $\forall j \geq J$ sufficiently large, it must be true that:

$$E_t(P_{t+1} - P_t | y_t = y_{t-1} = \dots = y_{t-j} = y) < E_t(P_{t+1} - P_t | y_t = y_{t-1} = \dots = y_{t-j} = -y)$$

which is nothing other than our original definition of overreaction.

Rewriting the condition $f(\underline{\pi}) < 0$ as:

$$2y(p_2\underline{\pi} - p_1) + yp_2(\overline{\Delta}(\underline{\pi}) + \underline{\Delta}(\underline{\pi})) < 0$$

we obtain:

$$p_1 > p_2\left(\underline{\pi} + \frac{\overline{\Delta}(\underline{\pi})}{2}\right) \quad (1)$$

which is one of the sufficient conditions given in the Proposition.

We now turn to a sufficient condition for underreaction. The definition of underreaction can also be succinctly stated in terms of $f(\pi)$ as:

$$E_\pi(f(\pi)) > 0,$$

where E_π denotes an expectation taken over the unconditional distribution of π . Rewriting this we obtain,

$$2yp_2E(\pi) - 2yp_1 + yp_2E_\pi(\overline{\Delta}(\pi) + \underline{\Delta}(\pi)) > 0$$

and hence,

$$p_1 < p_2\left(E(\pi) + \frac{E_\pi(\overline{\Delta}(\pi) + \underline{\Delta}(\pi))}{2}\right) \quad (2)$$

Unfortunately we are not yet finished because we do not have closed form formulas for the expectations in this expression. To provide sufficient conditions, we need to bound these quantities. Note that if $p_2 < 0$, it is impossible for both (1) and (2) to be satisfied simultaneously because:

$$\underline{\pi} + \frac{\overline{\Delta}(\underline{\pi})}{2} < E(\pi) + \frac{E_\pi(\overline{\Delta}(\pi) + \underline{\Delta}(\pi))}{2}.$$

Therefore, from now on we will suppose that $p_2 \geq 0$.

The first step in bounding the expression $E(\pi) + \frac{1}{2}E_\pi(\overline{\Delta}(\pi) + \underline{\Delta}(\pi))$ is to bound $E(\pi)$. To do this, note that:

$$\begin{aligned} E(\pi_t) &= E(\pi_{t+1}) = E_{\pi_t}(E(\pi_{t+1} | \pi_t)) \\ &= E_{\pi_t}\left(\frac{1}{2}(\pi_t + \overline{\Delta}(\pi_t)) + \frac{1}{2}(\pi_t - \underline{\Delta}(\pi_t))\right) \\ &= E_\pi(g(\pi)) \end{aligned}$$

Consider the function $g(\pi)$ defined on $[\underline{\pi}, \overline{\pi}]$. The idea is to bound this function above and below over this interval by straight lines, parallel to the line passing through the endpoints of $g(\pi)$, namely $(\underline{\pi}, g(\underline{\pi}))$, and $(\overline{\pi}, g(\overline{\pi}))$. In other words, suppose that we bound $g(\pi)$ above by $\overline{g}(\pi) = a + b\pi$. The slope of this line,

$$b = \frac{g(\overline{\pi}) - g(\underline{\pi})}{\overline{\pi} - \underline{\pi}} = \frac{(\overline{\pi} - \underline{\pi}) - \frac{1}{2}(\underline{\Delta}(\overline{\pi}) + \overline{\Delta}(\underline{\pi}))}{\overline{\pi} - \underline{\pi}} < 1$$

and a will be such that:

$$\inf_{\pi \in [\underline{\pi}, \bar{\pi}]} (a + b\pi - g(\pi)) = 0.$$

Given that

$$E_{\pi}(g(\pi) - \pi) = 0$$

we must have:

$$E_{\pi}(\bar{g}(\pi) - \pi) \geq 0$$

or

$$\begin{aligned} E(a + b\pi - \pi) &\geq 0 \\ E(\pi) &\leq \frac{a}{1-b} \end{aligned}$$

since $b < 1$. This gives us an upper bound on $E(\pi)$, which we will call $\bar{\pi}^e$. A similar argument produces a lower bound $\underline{\pi}^e$.

The final step before completing the argument is to note that since $\bar{\Delta}(\pi)$ and $\underline{\Delta}(\pi)$ are both concave, $\bar{\Delta}(\pi) + \underline{\Delta}(\pi)$ is also concave, so that

$$\begin{aligned} (\bar{\Delta} + \underline{\Delta})(\pi) &> \left(\frac{\pi - \underline{\pi}}{\bar{\pi} - \underline{\pi}} \right) \underline{\Delta}(\bar{\pi}) + \left(\frac{\bar{\pi} - \pi}{\bar{\pi} - \underline{\pi}} \right) \bar{\Delta}(\underline{\pi}) \\ &= c_1 + c_2\pi \end{aligned}$$

where:

$$\begin{aligned} c_1 &= \frac{\bar{\Delta}(\underline{\pi})\bar{\pi} - \underline{\Delta}(\bar{\pi})\underline{\pi}}{\bar{\pi} - \underline{\pi}} \\ c_2 &= \frac{\underline{\Delta}(\bar{\pi}) - \bar{\Delta}(\underline{\pi})}{\bar{\pi} - \underline{\pi}}. \end{aligned}$$

Therefore,

$$\begin{aligned} E(\pi) + \frac{1}{2}E(\bar{\Delta}(\pi) + \underline{\Delta}(\pi)) &\geq \bar{\pi}^e + \frac{1}{2}E(c_1 + c_2\pi) \\ &\geq \bar{\pi}^e + \frac{1}{2}(c_1 + c_2\pi_{\star}) \end{aligned}$$

where:

$$\pi_{\star} = \begin{cases} \bar{\pi}^e & \text{if } c_2 < 0 \\ \underline{\pi}^e & \text{if } c_2 \geq 0 \end{cases}$$

This completes the proof of the Proposition.

References

- [1] Andreassen, P., and S. Kraus (1990): "Judgmental Extrapolation and the Salience of Change", *Journal of Forecasting*, vol. 9.
- [2] Barsky, R., and J.B. De Long (1993): "Why Does the Stock Market Fluctuate?", *Quarterly Journal of Economics*, vol. 108.
- [3] Bernard, V., and J. Thomas (1989): "Post-Earnings Announcement Drift: Delayed Price Response or Risk Premium?", *Journal of Accounting Research*.
- [4] Bernard, V., and J. Thomas (1990): "Evidence that Stock Prices do not Fully Reflect the Implications of Current Earnings for Future Earnings", *Journal of Accounting and Economics*.
- [5] Bernard, V. (1992): "Stock Price Reactions to Earnings Announcements", in R. Thaler ed., *Advances in Behavioral Finance*, Russel Sage Foundation, New York.
- [6] Campbell, J.Y., and R. Shiller (1988): "Stock Prices, Earnings, and Expected Dividends", *Journal of Finance*, vol. 43.
- [7] Chan, L., N. Jegadeesh, and J. Lakonishok (1996): "Momentum Strategies", NBER Working Paper 5375, forthcoming *Journal of Finance*.
- [8] Chopra, N., J. Lakonishok, and J. Ritter (1992): "Measuring Abnormal Performance: Do Stocks Overreact?", *Journal of Financial Economics*, vol. 31.
- [9] Cutler, D., J. Poterba, and L. Summers (1992): "Speculative Dynamics", *Review of Economic Studies*, vol. 58, no. 3.
- [10] Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1996): "A Theory of Overconfidence, Self-Attribution, and Security Market Under- and Over-reactions", Working Paper, Kellogg School of Management.
- [11] De Bondt, W. (1993): "Betting on Trends: Intuitive Forecasts of Financial Risk and Return", *International Journal of Forecasting*, vol. 9, no. 3.

- [12] De Bondt, W., and R. Thaler (1985): "Does the Stock Market Overreact?", *Journal of Finance*, vol. 40, no. 3.
- [13] De Bondt, W., and R. Thaler (1987): "Further Evidence of Investor Overreaction and Stock Market Seasonality", *Journal of Finance*, vol. 42, no.3.
- [14] De Long, J.B., A. Shleifer, L. Summers, and R. Waldmann (1990a): "Noise Trader Risk in Financial Markets", *Journal of Political Economy*, vol. 98, no.4.
- [15] De Long, J.B., A. Shleifer, L. Summers, and R. Waldmann (1990b): "Positive Feedback Investment Strategies and Destabilizing Rational Speculation", *Journal of Finance*, vol. 45, no.2.
- [16] Edwards, W. (1968): "Conservatism in Human Information Processing", in B. Kleinmütz ed., *Formal Representation of Human Judgment*, New York, John Wiley and Sons.
- [17] Fama, E., and K. French (1988): "Permanent and Temporary Components of Stock Market Prices", *Journal of Political Economy*, no. 96.
- [18] Fama, E., and K. French (1992): "Cross-sectional Variation in Expected Stock Returns", *Journal of Finance*, no. 47.
- [19] Fama, E., and K. French (1996): "Multifactor Explanations of Asset Pricing Anomalies", *Journal of Finance*, no. 51.
- [20] Frankel, J., and K. Froot (1987): "Using Survey Data to Test some Standard Propositions regarding Exchange Rate Expectations", *American Economic Review*, vol. 77, no.1.
- [21] Frankel, J., and K. Froot (1989): "Forward Discount Bias: is it an Exchange Risk Premium?", *Quarterly Journal of Economics*, vol. 104, no.1
- [22] Frankel, J., and K. Froot (1990): "Chartists, Fundamentalists, and Trading in Foreign Exchange Markets", *American Economic Review*, vol. 80, no.2.

- [23] Ikenberry, D., J. Lakonishok, and T. Vermaelen (1995): "Market Underreaction to Open Market Share Repurchases", *Journal of Financial Economics*, vol. 39.
- [24] Jegadeesh, N., and S. Titman (1993): "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency", *Journal of Finance*, vol. 48, no. 1.
- [25] Kahneman, D., P. Slovic, and A. Tversky eds. (1974): *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge University Press.
- [26] Lakonishok, J., A. Shleifer, and R. Vishny (1994): "Contrarian Investment, Extrapolation, and Risk", *Journal of Finance*, vol, 49, no. 5.
- [27] Lakonishok, J., A. Shleifer, and R. Vishny (1996): "International Evidence on Glamour and Value Stocks", Work in Progress, Harvard University.
- [28] La Porta, R. (1996): "Expectations and the Cross-Section of Returns", *Journal of Finance*, vol. 51, no.5.
- [29] La Porta, R., J. Lakonishok, A. Shleifer, and R. Vishny (1996): "Good News for Value Stocks: Further Evidence on Market Efficiency", NBER Working Paper, no. 5311, forthcoming *Journal of Finance*.
- [30] Michaely, R., R. Thaler, and K. Womack (1995): "Price Reactions to Dividend Initiations and Omissions: Overreaction or Drift", *Journal of Finance*, vol. 50.
- [31] Poterba, J., and L. Summers (1988): "Mean-reversion in Stock Prices: Evidence and Implications", *Journal of Financial Economics*, no. 22.
- [32] Rabin, M. (1996): "Psychology and Economics", Manuscript, University of California - Berkeley.
- [33] Rabin, M., and J. Schrag (1996): "First Impressions Matter: A Model of Confirmatory Bias", Manuscript, University of California - Berkeley.
- [34] Samuelson, P. (1965): "Proof that Properly Anticipated Prices Fluctuate Randomly", *Industrial Management Review*, no.6.

- [35] Seyhun, A. (1992): “Why Does Aggregate Insider Trading Predict Future Stock Returns?”, *Quarterly Journal of Economics*, vol. 107.
- [36] Shleifer, A., and R. Vishny (1996): “The Limits of Arbitrage”, NBER Working Paper, no. 5167, forthcoming *Journal of Finance*.
- [37] Zarowin, P. (1989): “Does the Stock Market Overreact to Corporate Earnings Information?”, *Journal of Finance*, vol. 44.

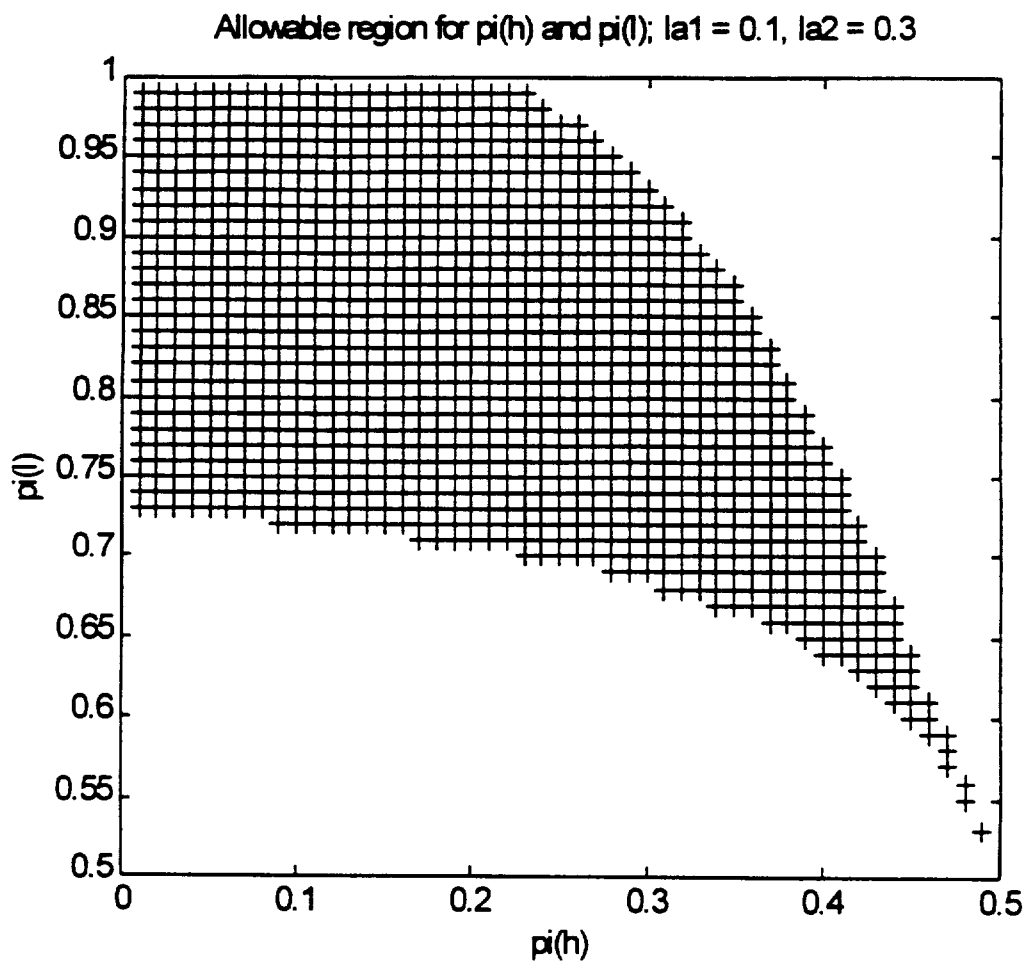
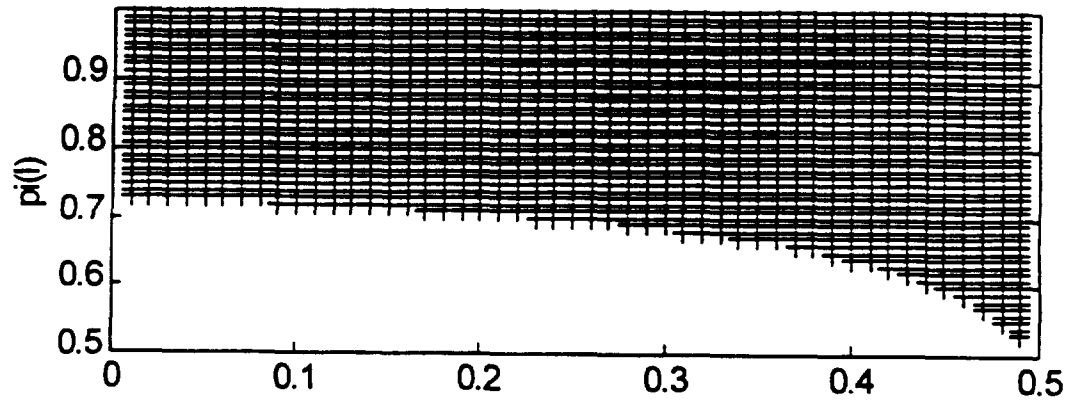


Figure 1: Shaded area marks the $(\pi_i(h), \pi_i(l))$ pairs which satisfy the sufficient conditions for both underreaction and overreaction

Allowable region for OVERREACTION; $la_1 = 0.1, la_2 = 0.3$



Allowable region for UNDERREACTION; $la_1 = 0.1, la_2 = 0.3$

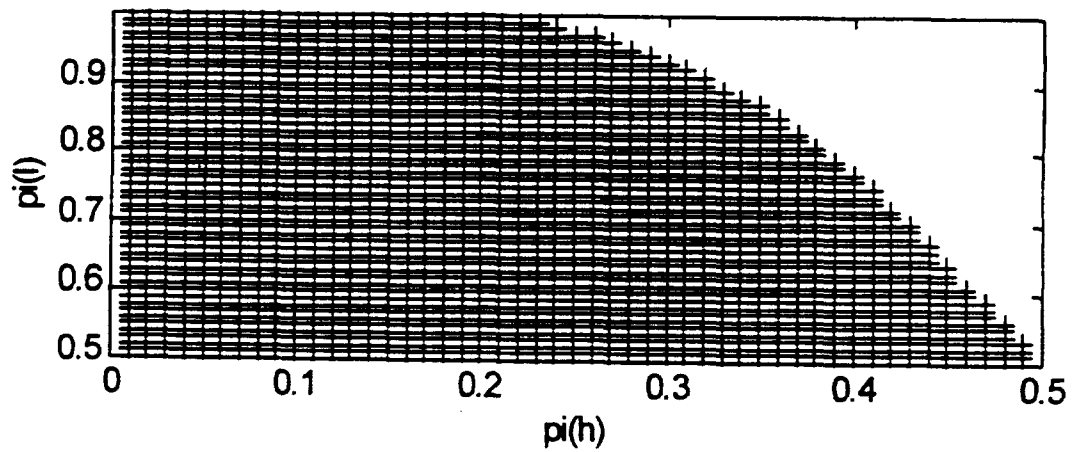


Figure 2: $(pi(h), pi(l))$ pairs which satisfy the sufficient conditions for under-reaction and overreaction plotted out separately

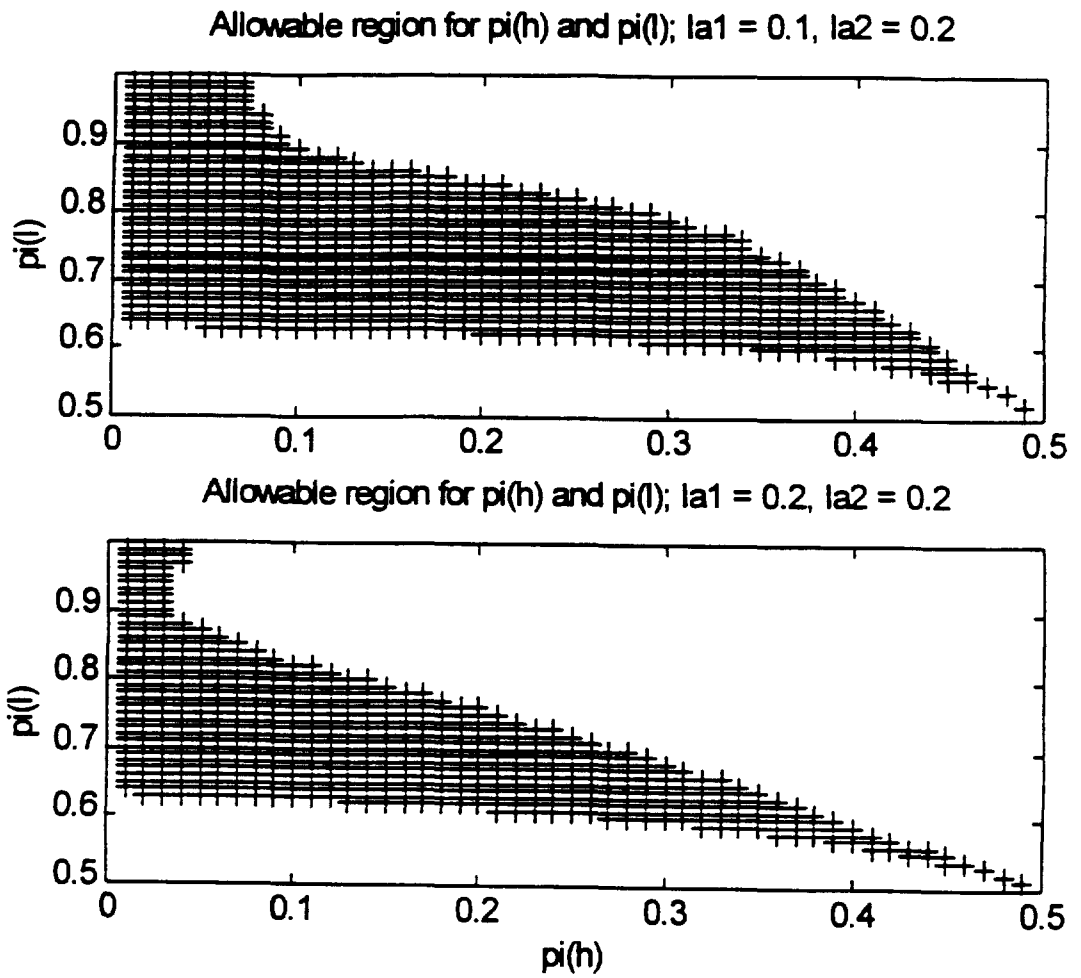


Figure 3: Shaded area shows the $(\pi(h), \pi(l))$ pairs which satisfy the sufficient conditions for both underreaction and overreaction

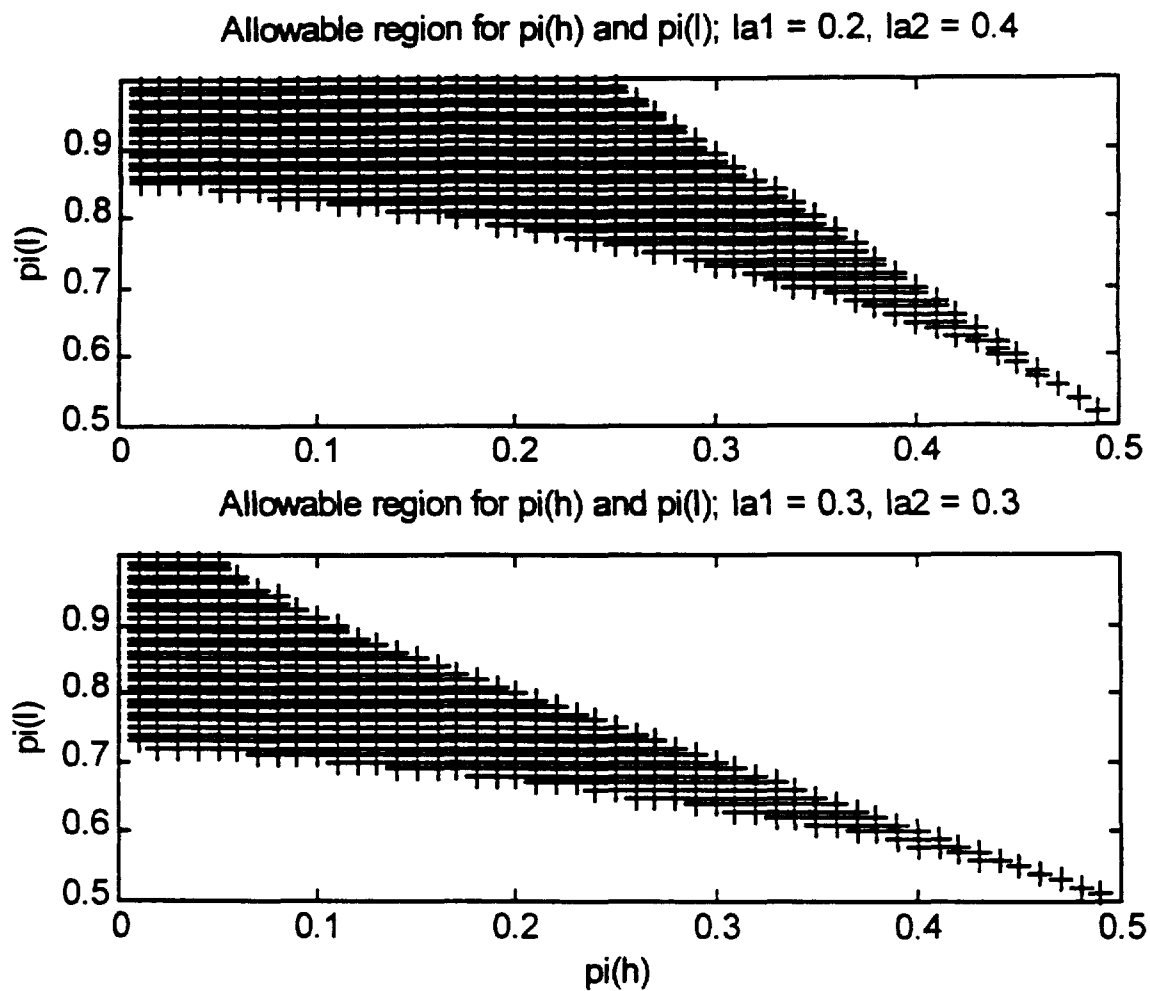


Figure 4: Shaded area shows the $(\pi_i(h), \pi_i(l))$ pairs which satisfy the sufficient conditions for both underreaction and overreaction