ELSEVIER Journal of Financial Economics 49 (1998) 307–343

JOURNAL OF Financial ECONOMICS

A model of investor sentiment[1]

Nicholas Barberis[a], Andrei Shleifer[b,*], Robert Vishny[a]

[a] *Graduate School of Business, University of Chicago, Chicago, IL 60637, USA*
[b] *Harvard University, Cambridge, MA 02138, USA*

Received 29 January 1997; received in revised form 10 February 1998

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, or of how investors form beliefs, which 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.

JEL classification: G12; G14

Keywords: Investor sentiment; Underreaction; Overreaction

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 1–12 months, security prices underreact to news.[2]

* Corresponding author. Tel.: 617/495-5046; fax: 617/496-1708; e-mail: ashleifer@harvard.edu.

[1] We are grateful to the NSF for financial support, and to Oliver Blanchard, Alon Brav, John Campbell (a referee), John Cochrane, Edward Glaeser, J.B. Heaton, Danny Kahneman, David Laibson, Owen Lamont, Drazen Prelec, Jay Ritter (a referee), Ken Singleton, Dick Thaler, an anonymous referee, and the editor, Bill Schwert, for comments.

[2] Some of the papers in this area, discussed in more detail in Section 2, include Cutler et al. (1991), Bernard and Thomas (1989), Jegadeesh and Titman (1993), and Chan et al. (1997).

0304-405X/98/$19.00
PII S0304-405X(98)00027-0

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 news has power in predicting positive returns in the future. The overreaction evidence shows that over longer horizons of perhaps 3–5 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 news tend to become overpriced and have low average returns afterwards.[3] Put differently, securities with strings of good performance, however measured, receive extremely high valuations, and these valuations, on average, return to the mean.[4]

The 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. The most notable recent attempt to explain the evidence from the efficient markets viewpoint is Fama and French (1996). The authors believe that their three-factor model can account for the overreaction evidence, but not for the continuation of short-term returns (underreaction). This evidence also presents a challenge to behavioral finance theory because early models do not successfully explain the facts.[5] 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 statistical 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 Tversky and Kahneman (1974) on the important behavioral heuristic known as representativeness, or 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 a history of consistent

[3] Some of the papers in this area, discussed in more detail in Section 2, include Cutler et al. (1991), De Bondt and Thaler (1985), Chopra et al. (1992), Fama and French (1992), Lakonishok et al. (1994), and La Porta (1996).

[4] There is also some evidence of nonzero return autocorrelations at very short horizons such as a day (Lehmann, 1990). We do not believe that it is essential for a behavioral model to confront this evidence because it can be plausibly explained by market microstructure considerations such as the fluctuation of recorded prices between the bid and the ask.

[5] The model of De Long et al. (1990a) generates negative autocorrelation in returns, and that of De Long et al. (1990b) generates positive autocorrelation. Cutler et al. (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.

earnings growth, 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 the 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 the model why arbitrage fails to eliminate the mispricing. For the purposes of this paper, we rely on earlier work showing why deviations from efficient prices can persist (De Long et al., 1990a; Shleifer and Vishny, 1997). According to this work, an important reason why arbitrage is limited is that movements in investor sentiment are in part unpredictable, and therefore arbitrageurs betting against mispricing run the risk, at least in the short run, that investor sentiment becomes more extreme and prices move even further away from fundamental value. As a consequence of such 'noise trader risk,' arbitrage positions can lose money in the short run. When arbitrageurs are risk-averse, leveraged, or 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.[6]

While these earlier papers argue that mispricing can persist, they say little about the nature of the mispricing that might be observed. For that, we need a model of how people form expectations. The current paper provides one such model.

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

[6] The empirical implications of our model are derived from the assumptions about investor psychology or sentiment, rather than from those about the behavior of arbitrageurs. Other models in behavioral finance yield empirical implications that follow from limited arbitrage alone, without specific assumptions about the form of investor sentiment. For example, limited arbitrage in closed-end funds predicts average underpricing of such funds regardless of the exact form of investor sentiment that these funds are subject to (see De Long et al., 1990a; Lee et al., 1991).

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, and uses this information to update his beliefs about which state he is in. In his updating, the investor is Bayesian, although his model of the earnings 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 data.

Daniel et al. (1998) also construct a model of investor sentiment aimed at reconciling the empirical findings of overreaction and underreaction. They, too, use concepts from psychology to support their framework, although the underpinnings of their model are overconfidence and self-attribution, which are not the same as the psychological ideas we use. It is quite possible that both the phenomena that they describe, and those driving our model, play a role in generating the empirical evidence.

Section 2 of the paper summarizes the empirical findings that we try to explain. Section 3 discusses the psychological evidence that motivates our approach. Section 4 presents the model. Section 5 solves it and outlines its implications for the data. Section 6 concludes.

2. The evidence

In this section, we summarize the statistical evidence of underreaction and overreaction in security returns. We devote only minor attention to the behavior of aggregate stock and bond returns because these data generally do not provide enough information to reject the hypothesis of efficient markets. Most of the anomalous evidence that our model tries to explain comes from the cross-section of stock returns. Much of this evidence is from the United States, although some recent research has found similar patterns in other markets.

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:

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

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 discuss below, there is considerable evidence of underreaction to other types of news as well.

Empirical analysis of aggregate time series has produced some evidence of underreaction. Cutler et al. (1991) examine autocorrelations in excess 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 excess index returns over horizons of between one month and one year. For example, the average one-month autocorrelation in excess stock returns across the world is around 0.1 (and is also around 0.1 in the United States alone), and that in excess bond returns is around 0.2 (and around zero 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.

More convincing support 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 these studies 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 standardized 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 an earnings announcement. A general (and unsurprising) finding is that stocks with positive earnings surprises also earn relatively high returns in the period prior to the earnings announcement, as information about earnings is 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, over the 60 trading days *after* portfolio formation, stocks with the highest SUE earn a cumulative risk-adjusted return that is 4.2% 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 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, using a sample of 2626 firms, Bernard and Thomas (1990) find that these series exhibit an autocorrelation of about 0.34 at a lag of one quarter, 0.19 at two quarters, 0.06 at three quarters, and -0.24 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 conjectures that market participants do not recognize 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 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, which Bernard's and our analyses share, is that investors typically (but not always) believe that earnings are more stationary than they really are. As we show below, this idea has firm foundations in psychology.

Further evidence of underreaction comes from Jegadeesh and Titman (1993), who examine a cross-section of U.S. stock returns and find reliable evidence that over a 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.[7] More recent work by Rouwenhorst (1997) documents the presence of momentum in international equity markets. Chan et al. (1997) 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 six months and one year. That is, 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.

In addition to the evidence of stock price underreaction to earnings announcements and the related evidence of momentum in stock prices, there is also a body of closely related evidence on stock price drift following many other announcements and events. For example, Ikenberry et al. (1995) find that stock

[7] Early evidence on momentum is also contained in De Bondt and Thaler (1985).

prices rise on the announcement of share repurchases but then continue to drift in the same direction over the next few years. Michaely et al. (1995) find similar evidence of drift following dividend initiations and omissions, while Ikenberry et al. (1996) document such a drift following stock splits. Finally, Loughran and Ritter (1995) and Spiess and Affleck-Graves (1995) find evidence of a drift following seasoned equity offerings. Daniel et al. (1998) and Fama (1998) summarize a large number of event studies showing this type of underreaction to news events, which a theory of investor sentiment should presumably come to grips with.

2.2. Statistical evidence of overreaction

Analogous to the definition of underreaction at the start of the previous subsection, we now define overreaction as occurring when 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,

$$\begin{aligned} &\mathrm{E}(r_{t+1}|z_t = G, z_{t-1} = G, \ldots, z_{t-j} = G) \\ &\quad < \mathrm{E}(r_{t+1}|z_t = B, z_{t-1} = B, \ldots, z_{t-j} = B), \end{aligned}$$

where j is at least one and probably rather higher. The idea here is simply that after a series of announcements of good news, the investor becomes overly optimistic that future news announcements will also be good and hence overreacts, 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. Early papers include Fama and French (1988) and Poterba and Summers (1988); Cutler et al. (1991) examine some of this evidence for a variety of markets. The thrust of the evidence is that, over horizons of 3–5 years, there is a relatively slight negative autocorrelation in stock returns in many markets. Moreover, over similar horizons, some measures of stock valuation, such as the dividend yield, have predictive power for returns in a similar direction: a low dividend yield or high past return tend to predict a low subsequent return (Campbell and Shiller, 1988).

As before, the more convincing evidence comes from the cross-section of stock returns. In an early important paper, De Bondt and Thaler (1985) discover from looking at U.S. data dating back to 1933 that portfolios of stocks with extremely poor returns over the previous five years dramatically outperform portfolios of stocks with extremely high returns, even after making the standard risk adjustments. De Bondt and Thaler's findings are corroborated by later work (e.g., Chopra et al., 1992). In the case of earnings, Zarowin (1989) finds that firms that have had a sequence of bad earnings realizations subsequently outperform firms with a sequence of good earnings. This evidence suggests that stocks with a consistent record of good news, and hence extremely high past returns, are overvalued, and that an investor can therefore earn abnormal returns by betting against this overreaction to consistent patterns of news. Similarly, stocks with a consistent record 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 (De Bondt and Thaler, 1987; Fama and French, 1992), market value to cash flow (Lakonishok et al., 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. find spreads of 8–10% per 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 relative to what the subsequent returns actually deliver. Recent research extends the evidence on value stocks to other markets, including those in Europe, Japan, and emerging markets (Fama and French, 1998; Haugen and Baker, 1996).

The economic interpretation of this evidence has 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. In a direct attempt to distinguish risk and overreaction, La Porta (1996) sorts stocks on the basis of long-term growth rate forecasts made by professional analysts, and finds evidence that analysts are excessively bullish about the stocks they are most optimistic about and excessively bearish about the stocks they are most pessimistic about. In particular, stocks with the highest growth forecasts earn much lower future returns than stocks with the lowest 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 an efficient market, stocks with optimistic growth forecasts should not earn low returns.

Finally, La Porta et al. (1997) 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, presents rather reliable regularities. These regularities taken in their entirety are difficult to reconcile with the efficient markets hypothesis. More important for this paper, the two regularities challenge behavioral finance to provide a model of how investors form beliefs that can account for the empirical evidence.

3. Some psychological evidence

The model we present below is motivated by two important phenomena documented by psychologists: *conservatism* and the *representativeness heuristic*. In this subsection, we briefly describe this psychological evidence as well as a recent attempt to integrate it (Griffin and Tversky, 1992).

Several psychologists, including Edwards (1968), have identified a phenomenon known as conservatism. Conservatism states that individuals are slow to change their beliefs in the face of new evidence. Edwards benchmarks a subject's reaction to new evidence against that of an idealized rational Bayesian in experiments in which the true normative value of a piece of evidence is well defined. In his experiments, individuals update 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)

Conservatism is extremely suggestive of the underreaction evidence described above. Individuals subject to conservatism might disregard the full information content of an earnings (or some other public) announcement, perhaps because they believe that this number contains a large temporary component, and still cling at least partially to their prior estimates of earnings. As a consequence, they might adjust their valuation of shares only partially in response to the announcement. Edwards would describe such behavior in Bayesian terms as a failure to properly aggregate the information in the new earnings number with investors' own prior information to form a new posterior earnings estimate. In particular, individuals tend to underweight useful statistical evidence relative to the less useful evidence used to form their priors. Alternatively, they might be characterized as being overconfident about their prior information.

A second important phenomenon documented by psychologists is the representativeness heuristic (Tversky and Kahneman, 1974): "A person who follows this heuristic evaluates the probability of an uncertain event, or a sample, by the 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 tends 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 Tversky and Kahneman, is that people think they see patterns in truly random sequences. This aspect of the representativeness heuristic is suggestive of the overreaction evidence described above. When a company has a consistent history of earnings growth over several years, accompanied as it may be by salient and enthusiastic descriptions of its products and management, investors might conclude that the past history is representative of an underlying earnings growth potential. While a consistent pattern of high growth may be nothing more than a random draw for a few lucky firms, investors see 'order among chaos' and infer from the in-sample growth path that the firm belongs to a small and distinct population of firms whose earnings just keep growing. As a consequence, investors using the representativeness heuristic might disregard the reality that a history of high earnings growth is unlikely to repeat itself; they will overvalue the company, and be disappointed in the future when the forecasted earnings growth fails to materialize. This, of course, is what overreaction is all about.

In a recent study, Griffin and Tversky (1992) attempt to reconcile conservatism with representativeness. In their framework, people update their beliefs based on the 'strength' and the 'weight' of new evidence. Strength refers to such aspects of the evidence as salience and extremity, whereas weight refers to statistical informativeness, such as sample size.[8] According to Griffin and Tversky, in revising their forecasts, people focus *too much* on the strength of the evidence, and *too little* on its weight, relative to a rational Bayesian. In the Griffin–Tversky framework, conservatism like that documented by Edwards would occur in the face of evidence that has high weight but low strength: people are unimpressed by the low strength and react mildly to the evidence, even though its weight calls for a larger reaction. On the other hand, when the evidence has high strength but low weight, overreaction occurs in a manner consistent with representativeness. Indeed, representativeness can be thought of as excessive attention to the strength of particularly salient evidence, in spite of its relatively low weight.

[8] To illustrate these concepts, Griffin and Tversky use the example of a recommendation letter. The 'strength' of the letter refers to how positive and warm its content is; 'weight' on the other hand, measures the credibility and stature of the letter-writer.

In the context at hand, Griffin and Tversky's theory suggests that individuals might underweight the information contained in isolated quarterly earnings announcements, since a single earnings number seems like a weakly informative blip exhibiting no particular pattern or strength on its own. In doing so, they ignore the substantial *weight* that the latest earnings news has for forecasting the level of earnings, particularly when earnings are close to a random walk. At the same time, individuals might overweight consistent multiyear patterns of noticeably high or low earnings growth. Such data can be very salient, or have high strength, yet their weight in forecasting earnings growth rates can be quite low.

Unfortunately, the psychological evidence does not tell us quantitatively what kind of information is strong and salient (and hence is overreacted to) and what kind of information is low in weight (and hence is underreacted to). For example, it does not tell us how long a sequence of earnings increases is required for its strength to cause significant overpricing. Nor does the evidence tell us the magnitude of the reaction (relative to a true Bayesian) to information that has high strength and weight, or low strength and weight. For these reasons, it would be inappropriate for us to say that our model is derived from the psychological evidence, as opposed to just being motivated by it.

There are also some stock trading experiments that are consistent with the psychological evidence as well as with the model presented below. Andreassen and Kraus (1990) 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's 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's results whether subjects actually switch from bucking trends to chasing them, although their findings certainly suggest it.

De Bondt (1993) nicely complements Andreassen and Kraus's findings. Using a combination of classroom experiments and investor surveys, De Bondt 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 analyzes 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 forecasted change in price level is higher following a series of previous price increases than following price decreases, suggesting that investors indeed chase trends once they think they see them.

4. A model of investor sentiment

4.1. Informal description of the model

The model we present in this section attempts to capture the empirical evidence summarized in Section 2 using the ideas from psychology discussed in Section 3. We consider a model with a representative, risk-neutral investor with discount rate δ. We can think of this investor's beliefs as reflecting the 'consensus', even if different investors have different beliefs. There is only one security, which pays out 100% of its earnings as dividends; in this context, the equilibrium price of the security is equal to the net present value of future earnings, as forecasted by the representative investor. In contrast to models with heterogeneous agents, there is no information in prices over and above the information already contained in earnings.

Given the assumptions of risk-neutrality and a constant discount rate, returns are 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, the investor must be 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 autocorrelated (Bernard and Thomas, 1990). 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. This relative misperception is the key to underreaction.

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, so that 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, shocks are more likely to be followed by another shock of the same sign.

The idea that the investor believes that the world is governed by one of the two incorrect models is a crude way of capturing the psychological phenomena of the previous section. Model 1 generates effects identical to those predicted by conservatism. An investor using Model 1 to forecast earnings reacts too little to an individual earnings announcement, as would an investor exhibiting conservatism. From the perspective of Griffin and Tversky (1992), there is insufficient reaction to individual earnings announcements because they are low in strength. In fact, these announcements have extremely high weight when earnings follow a random walk, but investors are insensitive to this aspect of the evidence.

In contrast, the investor who believes in Model 2 behaves as if he is subject to the representativeness heuristic. After a string of positive or negative earnings changes, the investor uses Model 2 to forecast future earnings, extrapolating past performance too far into the future. This captures the way that representativeness might lead investors to associate past earnings growth too strongly with future earnings growth. In the language of Griffin and Tversky, investors overreact to the information in a string of positive or negative earnings changes since it is of high strength; they ignore the fact that it has low weight when earnings simply follow a random walk.

The investor also believes that there is an underlying regime-switching process that 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 in which 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 become apparent, 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 believes 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 puts more weight on Model 1, the mean-reverting regime.

Our model differs from more typical models of learning. In our framework, the investor never changes the model he is using to forecast earnings, but rather 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. This is the only sense in which he is learning from the data.[9]

We now provide some preliminary intuition for how investor behavior of the kind described above, coupled with the true random walk process for earnings, can generate the empirical phenomena discussed in Section 2. In particular, we show how our framework can lead to both underreaction to earnings announcements and long-run overreaction.

In our model, a natural way of capturing overreaction 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 earnings. Indeed, after our investor sees a series of positive earnings shocks, he puts a high probability on the event that Model 2 is generating 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 the shock to earnings next period to be positive again. 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 is large and negative as the investor is taken by surprise by the negative announcement.[10] The average realized return after a string of positive shocks is therefore 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. Underreaction obtains in our model as long as the investor places more weight on Model 1 than on Model 2, on average. Consider the realized return following a positive earnings shock. Since, by assumption, the investor on average believes Model 1, he on average believes that this positive earnings shock will be partly reversed in the next period. In reality, however, a positive shock is as likely to be followed by a positive as by a negative shock. If the shock is negative, the realized return is not large, since this is the earnings

[9] From a mathematical perspective, the investor would eventually learn the true random walk model for earnings if it were included in the support of his prior; from the viewpoint of psychology, though, there is much evidence that people learn *slowly* and find it difficult to shake off pervasive biases such as conservatism and representativeness.

[10] A referee has pointed out to us that this is exactly the empirical finding of Dreman and Berry (1995). They find that glamour stocks earn small positive event returns on positive earnings surprises and large negative event returns on negative earnings surprises. The converse holds for value stocks.

realization that was expected by the investor. If the shock is positive, the realized return is large and positive, since this shock is unexpected. Similarly, the average realized return following a negative earnings shock is negative, and hence the difference in the average realized returns is indeed positive, consistent with the evidence of post-earnings announcement drift and short-term momentum.

The empirical studies discussed in Section 2 indicate that underreaction may be a broader phenomenon than simply the delayed reaction to earnings documented by Bernard and Thomas (1989). Although our model is formulated in terms of earnings news, delayed reaction to announcements about dividends and share repurchases can be understood just as easily in our framework. In the same way that the investor displays conservatism when adjusting his beliefs in the face of a new earnings announcement, so he may also underweight the information in the announcement of a dividend cut or a share repurchase.

The mechanism for expectation formation that we propose here is related to that used by Barsky and De Long (1993) in an attempt to explain Shiller's (1981) finding of excess volatility in 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 closely resembles 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 also rises, leading him to forecast higher dividends in the future as well. Analogously, in our model, a series of positive shocks to earnings leads 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.

4.2. A formal model

We now present a mathematical model of the investor behavior described above, and in Section 5, we check that the intuition can be formalized. 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 $-y$. Assume that all earnings are paid out as dividends. 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	$y_{t+1} = y$	$y_{t+1} = -y$
$y_t = y$	π_L	$1 - \pi_L$
$y_t = -y$	$1 - \pi_L$	π_L

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

The key is that π_L is small and π_H is large. We shall think of π_L as falling between zero and 0.5, with π_H falling between 0.5 and one. 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_L and π_H; he is also sure that he is right about the underlying process controlling the switching from one regime to another, or equivalently from Models 1 to 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

	$s_{t+1} = 1$	$s_{t+1} = 2$
$s_t = 1$	$1 - \lambda_1$	λ_1
$s_t = 2$	λ_2	$1 - \lambda_2$

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. In particular, we assume that $\lambda_1 + \lambda_2 < 1$. We also think of λ_1 as being smaller than λ_2. Since the unconditional probability of being in state 1 is $\lambda_2/(\lambda_1 + \lambda_2)$, this implies that the investor thinks of Model 1 as being more likely than Model 2, on average. Our results do not depend, however, on λ_1 being smaller than λ_2. The effects that we 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. He observes earnings each period and 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 q_t, the probability that y_t was generated by Model 1, using the new data to update his estimate from the previous period, q_{t-1}. Formally, $q_t = \Pr(s_t = 1 | y_t, y_{t-1}, q_{t-1})$. We suppose that the updating follows Bayes Rule, so that

$$q_{t+1} = \frac{((1-\lambda_1)q_t + \lambda_2(1-q_t))Pr(y_{t+1}|s_{t+1}=1, y_t)}{((1-\lambda_1)q_t + \lambda_2(1-q_t))Pr(y_{t+1}|s_{t+1}=1, y_t) + (\lambda_1 q_t + (1-\lambda_2)(1-q_t))Pr(y_{t+1}|s_{t+1}=2, y_t)}.$$

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 q_{t+1} from q_t using

$$q_{t+1} = \frac{((1-\lambda_1)q_t + \lambda_2(1-q_t))\pi_{\mathrm{L}}}{((1-\lambda_1)q_t + \lambda_2(1-q_t))\pi_{\mathrm{L}} + (\lambda_1 q_t + (1-\lambda_2)(1-q_t))\pi_{\mathrm{H}}},$$

and we show in the Appendix that in this case, $q_{t+1} < q_t$. 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,

$$q_{t+1} = \frac{((1-\lambda_1)q_t + \lambda_2(1-q_t))(1-\pi_{\mathrm{L}})}{((1-\lambda_1)q_t + \lambda_2(1-q_t))(1-\pi_{\mathrm{L}}) + (\lambda_1 q_t + (1-\lambda_2)(1-q_t))(1-\pi_{\mathrm{H}})},$$

and in this case, $q_{t+1} > q_t$ and the weight on Model 1 increases.

To aid intuition about how the model works, we present a simple example shown in Table 1. 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., q_0, is 0.5. For a randomly generated earnings stream over the next 20 periods, the table below presents the investor's belief q_t that the time t shock to earnings is generated by Model 1. The particular parameter values chosen here are $\pi_{\mathrm{L}} = \frac{1}{3} < \frac{3}{4} = \pi_{\mathrm{H}}$, and $\lambda_1 = 0.1 < 0.3 = \lambda_2$. Note again that the earnings stream is generated using the true process for earnings, a random walk.

In periods 0–4, 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 q_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, q_t falls through period 14 to a low of 0.36.

Table 1

t	y_t	q_t	t	y_t	q_t
0	y	0.50			
1	$-y$	0.80	11	y	0.74
2	y	0.90	12	y	0.56
3	$-y$	0.93	13	y	0.44
4	y	0.94	14	y	0.36
5	y	0.74	15	$-y$	0.74
6	$-y$	0.89	16	y	0.89
7	$-y$	0.69	17	y	0.69
8	y	0.87	18	$-y$	0.87
9	$-y$	0.92	19	y	0.92
10	y	0.94	20	y	0.72

One feature that is evident in the above example is that q_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 $t-1$.

5. Model solution and empirical implications

5.1. Basic results

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

$$P_t = \mathrm{E}_t\left\{\frac{N_{t+1}}{1+\delta} + \frac{N_{t+2}}{(1+\delta)^2} + \cdots\right\}.$$

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, $\mathrm{E}_t(N_{t+j}) = N_t$, and price equals N_t/δ. In our model, price deviates from this correct 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* 4, *then prices satisfy*

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

where p_1 *and* p_2 *are constants that depend on* π_L, π_H, λ_1, and λ_2. *The full expressions for* p_1 and p_2 *are given in the Appendix.*[11]

The formula for P_t has a very simple interpretation. The first term, N_t/δ, is the price that would obtain if the investor used the true random walk process to forecast earnings. The second term, $y_t(p_1 - p_2 q_t)$, gives the deviation of price from this fundamental value. Later in this section we look at the range of values

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

of π_L, π_H, λ_1, and λ_2 that 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 forsake mathematical rigor in order to build intuition for those 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. Suppose the latest earnings shock y_t is a positive one. Underreaction means that, on average, the stock price does not react sufficiently to this shock, leaving the price below fundamental value. This means that, on average, $y(p_1 - p_2 q_t)$, the deviation from fundamental value, must be negative. If q_{avg} denotes an average value of q_t, this implies that we must have $p_1 < p_2 q_{\text{avg}}$. 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. Overreaction would require that price now be above fundamental value. Moreover, we know that after a series of shocks of the same sign, q_t is normally low, indicating a low weight on Model 1 and a high weight on Model 2. If we write q_{low} to represent a typical low value of q_t, overreaction then requires that $y(p_1 - p_2 q_{\text{low}})$ be positive, or that $p_1 > p_2 q_{\text{low}}$. This is the sense in which p_1 cannot be too small in relation to p_2. Putting the two conditions together, we obtain

$$p_2 q_{\text{low}} < p_1 < p_2 q_{\text{avg}}.$$

In Proposition 2, we provide sufficient conditions on p_1 and p_2 for prices to exhibit both underreaction and overreaction, 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, 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 number $J \geqslant 1$, such that for all $j \geqslant J$,

$$\mathrm{E}_t(P_{t+1} - P_t | y_t = y_{t-1} = \cdots = y_{t-j} = y) \\ - \mathrm{E}_t(P_{t+1} - P_t | y_t = y_{t-1} = \cdots = y_{t-j} = -y) < 0.$$

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

$$\mathrm{E}_t(P_{t+1} - P_t | y_t = +y) - \mathrm{E}_t(P_{t+1} - P_t | y_t = -y) > 0.$$

Proposition 2 below provides sufficient conditions on π_L, π_H, λ_1, and λ_2 for these two inequalities to hold.[12]

Proposition 2. *If the underlying parameters* π_L, π_H, λ_1, *and* λ_2 *satisfy*

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

$$p_2 \geqslant 0,$$

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_L, π_H, λ_1, *and* λ_2 (*the full expressions are given in the Appendix*).

We now examine the range of values of the fundamental parameters π_H, π_L, 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. Fig. 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 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 fixed, we want to know the range of values of π_L and π_H for which the conditions for underreaction and overreaction both hold. Given the way the model is set up, π_L and π_H are restricted to the ranges $0 < \pi_L < 0.5$ and $0.5 < \pi_H < 1$. We evaluate the conditions in Proposition 2 for pairs of (π_L, π_H) where π_L ranges from zero to 0.5 at intervals of 0.01 and π_H ranges from 0.5 to one, again at intervals of 0.01.

The graph at the top of Fig. 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 graph shows that the sufficient conditions do not hold if both π_L and π_H are near the high end of their feasible ranges, or if both π_L and π_H are near the low end of their ranges. The reason for this is the following. Suppose both π_L and π_H 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

[12] For 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 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 the expected change in earnings $\mathrm{E}_t(y_{t+1})$ always lies between $-y$ and $+y$, a positive earnings change is in fact a positive surprise. Therefore, the results are qualitatively the same in the two cases. In the simulations in Section 5.2, we calculate returns in the usual way, and condition on earnings surprises as well as raw earnings changes.

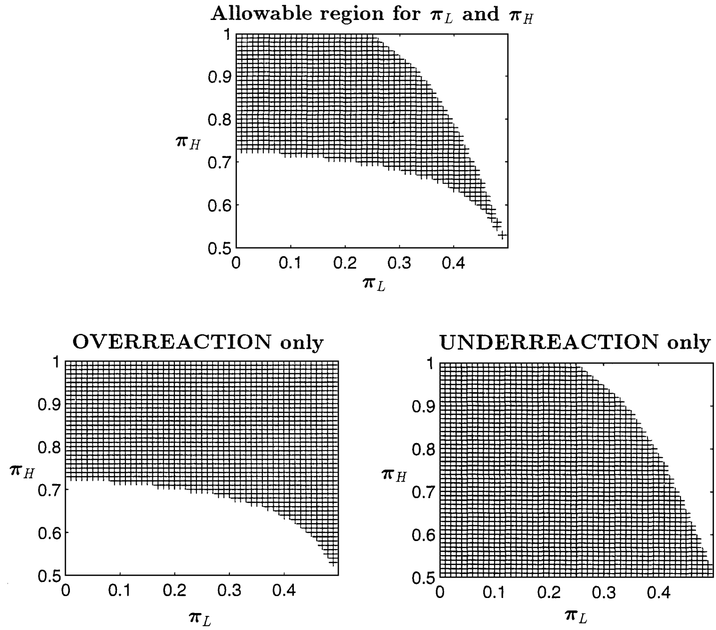

Fig. 1. Shaded area in graph at top marks the $[\pi_L, \pi_H]$ pairs which satisfy the sufficient conditions for both underreaction and overreaction, when $\lambda_1 = 0.1$ and $\lambda_2 = 0.3$. Graph at bottom left (right) shows the $[\pi_L, \pi_H]$ pairs that satisfy the condition for overreaction (underreaction) only. π_L (π_H) is the probability, in the mean-reverting (trending) regime, that next period's earnings shock will be of the same sign as last period's earnings shock. λ_1 and λ_2 govern the transition probabilities between regimes.

consequence of this is that overreaction certainly obtains, although underreaction might not. Following a positive shock, the investor on average expects another positive shock and since the true process is a random walk, returns are negative, on average. Hence the average return following a positive shock is lower than that following a negative shock, which is a characterization of overreaction rather than of underreaction.

On the other hand, if π_L and π_H 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 might not hold.

To confirm this intuition, we also show in Fig. 1 the ranges of (π_L, π_H) pairs for which only underreaction or only overreaction holds. The graph at bottom left shows the parameter values for which only overreaction obtains, while the graph to its right shows the values for which only underreaction holds. The intersection of the two regions is the original one shown in the graph at the top. These figures confirm the intuition that if π_L and π_H are on the high side, overreaction obtains, but underreaction might not.

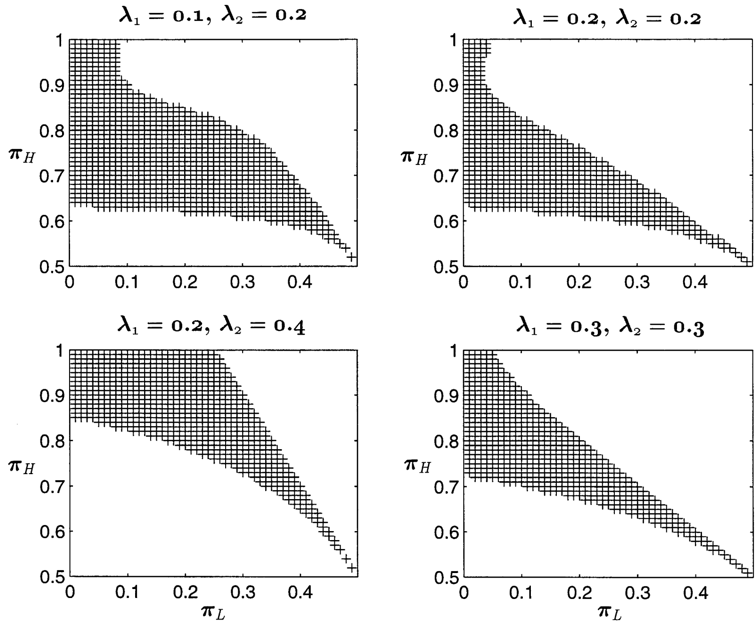

Fig. 2. Shaded area shows the $[\pi_L, \pi_H]$ pairs which satisfy the sufficient conditions for both underreaction and overreaction for a variety of different values of λ_1 and λ_2. π_L (π_H) is the probability, in the mean-reverting (trending) regime, that next period's earnings shock will be of the same sign as last period's earnings shock. λ_1 and λ_2 govern the transition probabilities between regimes.

Fig. 2 presents ranges of (π_L, π_H) pairs that generate both underreaction and overreaction for a number of other values of λ_1 and λ_2. In all cases, there are nontrivial ranges of (π_L, π_H) pairs for which the sufficient conditions hold.

5.2. Some simulation experiments

One way of evaluating our framework is to try to replicate the empirical findings of the papers discussed in Section 2 using artificial data sets of earnings and prices simulated from our model. First, we fix parameter values, setting the regime-switching parameters to $\lambda_1 = 0.1$ and $\lambda_2 = 0.3$. To guide our choice of π_L and π_H, we refer to Fig. 1. Setting $\pi_L = \frac{1}{3}$ and $\pi_H = \frac{3}{4}$ places us firmly in the region for which prices should exhibit both underreaction and overreaction.

Our aim is to simulate earnings, prices, and returns for a large number of firms over time. Accordingly, we choose an initial level of earnings N_1 and use the true random walk model to simulate 2000 independent earnings sequences, each one starting at N_1. Each sequence represents a different firm and contains six earnings realizations. We think of a period in our model as corresponding roughly to a year, so that our simulated data set covers six years. For the parameter values chosen, we can then apply the formula derived in Section 5.1 to calculate prices and returns.

One feature of the random walk model we use for earnings is that it imposes a constant volatility for the earnings shock y_t, rather than making this volatility proportional to the level of earnings N_t. While this makes our model tractable enough to calculate the price function in closed form, it also allows earnings, and hence prices, to turn negative. In our simulations, we choose the absolute value of the earnings change y to be small relative to the initial earnings level N_1 so as to avoid generating negative earnings. Since this choice has the effect of reducing the volatility of returns in our simulated samples, we pay more attention to the *sign* of the numbers we present than to their absolute magnitudes.

This aspect of our model also motivates us to set the sample length at a relatively short six years. For any given initial level of earnings, the longer the sample length, the greater is the chance of earnings turning negative in the sample. We therefore choose the shortest sample that still allows us to condition on earnings and price histories of the length typical in empirical analyses.

A natural starting point is to use the simulated data to calculate returns following particular realizations of earnings. For each n-year period in the sample, where n can range from one to four, we form two portfolios. One portfolio consists of all the firms with positive earnings changes in each of the n years, and the other of all the firms with negative earnings changes in each of the n years. We calculate the difference between the returns on these two portfolios in the year after formation. We repeat this procedure for all the n-year periods in the sample and calculate the time series mean of the difference in the two portfolio returns, which we call $r^n_+ - r^n_-$.

The calculation of $r^n_+ - r^n_-$ for the case of $n = 1$ essentially replicates the empirical analysis in studies such as that of Bernard and Thomas (1989). This quantity should therefore be positive, matching our definition of underreaction to news. Furthermore, to match our definition of overreaction, we need the average return in periods following a long series of consecutive positive earnings shocks to be *lower* than the average return following a similarly long series of negative shocks. Therefore, we hope to see $r^n_+ - r^n_-$ *decline* as n grows, or as we condition on a progressively longer string of earnings shocks of the same sign, indicating a transition from underreaction to overreaction. Table 2 below reports the results.

The results display the pattern we expect. The average return following a positive earnings shock is greater than the average return following a negative shock, consistent with underreaction. As the number of shocks of the same sign increases, the difference in average returns turns negative, consistent with overreaction.

Table 2

Earnings sort	
$r^1_+ - r^1_-$	0.0391
$r^2_+ - r^2_+$	0.0131
$r^3_+ - r^3_-$	– 0.0072
$r^4_+ - r^4_-$	– 0.0309

While the magnitudes of the numbers in the table are quite reasonable, their absolute values are smaller than those found in the empirical literature. This is a direct consequence of the low volatility of earnings changes that we impose to prevent earnings from turning negative in our simulations. Moreover, we report only point estimates and do not try to address the issue of statistical significance. Doing so would require more structure than we have imposed so far, such as assumptions about the cross-sectional covariance properties of earnings changes.

An alternative computation to the one reported in the table above would condition not on raw earnings but on the size of the surprise in the earnings announcement, measured relative to the investor's forecast. We have tried this calculation as well, and obtained very similar results.

Some of the studies discussed in Section 2, such as Jegadeesh and Titman (1993) and De Bondt and Thaler (1985), calculate returns conditional not on previous earnings realizations but on previous realizations of returns. We now attempt to replicate these studies.

For each n-year period in our simulated sample, where n again ranges from one to four, we group the 2000 firms into deciles based on their cumulative return over the n years, and compute the difference between the return of the best- and the worst-performing deciles for the year after portfolio formation. We repeat this for all the n-year periods in our sample, and compute the time series mean of the difference in the two portfolio returns, $r^n_{\mathrm{W}} - r^n_{\mathrm{L}}$.

We hope to find that $r^n_{\mathrm{W}} - r^n_{\mathrm{L}}$ decreases with n, with $r^1_{\mathrm{W}} - r^1_{\mathrm{L}}$ positive just as in Jegadeesh and Titman and $r^4_{\mathrm{W}} - r^4_{\mathrm{L}}$ negative as in De Bondt and Thaler. The results, shown in Table 3, are precisely these.

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 (E/P) ratios for the cross-section of returns. Each year, we group the 2000 stocks into deciles based on their E/P ratio and compute the difference between the return on the highest E/P decile and the return on the lowest E/P decile in the year after formation. We repeat this for each of the years in our sample and compute the time series mean of the difference in the two portfolio returns, which we call $r^{\mathrm{hi}}_{E/P} - r^{\mathrm{lo}}_{E/P}$. We find this statistic to be large and positive, matching

Table 3

Returns sort	
$r_W^1 - r_L^1$	0.0280
$r_W^2 - r_L^2$	0.0102
$r_W^3 - r_L^3$	– 0.0094
$r_W^4 - r_L^4$	– 0.0181

the empirical fact:

E/P sort	
$r_{E/P}^{\mathrm{hi}} - r_{E/P}^{\mathrm{lo}}$	0.0435

Note that this difference in average returns cannot be the result of a risk premium, since in our model the representative investor is assumed to be risk neutral.

5.3. *The event studies revisited*

We have already discussed the direct relationship between the concept of conservatism, the specification of regime 1 in our model, and the pervasive evidence of underreaction in event studies. We believe that regime 1 is consistent with the almost universal finding across different information events that stock prices tend to drift in the same direction as the event announcement return for a period of six months to five years, with the length of the time period dependent on the type of event.

An important question is whether our full model, and not just regime 1, is consistent with all of the event study evidence. Michaely et al. (1995) find that stock prices of dividend-cutting firms decline on the announcement of the cut but then continue falling for some time afterwards. This finding is consistent with our regime 1 in that it involves underreaction to the new and useful information contained in the cut. But we also know that dividend cuts generally occur after a string of bad earnings news. Hence, if a long string of bad earnings news pushes investors towards believing in regime 2, another piece of bad news such as a dividend cut would perhaps cause an overreaction rather than an underreaction in our model.[13]

While this certainly is one interpretation of our model, an alternative way of thinking about dividend announcements is consistent with both our model and

[13] Another study that presents a similar puzzle is by Ikenberry et al. (1996). They find that the positive price reaction to the announcement of a stock split is followed by a substantial drift in the same direction over the next few years. However, the split is also often preceded by a persistent run-up in the stock price, suggesting an overreaction that should ultimately be reversed.

the evidence. Specifically, our model only predicts an overreaction when the new information is part of a long string of similar numbers, such as earnings or sales figures. An isolated information event such as a dividend cut, an insider sale of stock, or a primary stock issue by the firm does not constitute part of the string, even though it could superficially be classified as good news or bad news like the earnings numbers that preceded it. Investors need not simply classify all information events, whatever their nature, as either good or bad news and then claim to see a trend on this basis. Instead, they may form forecasts of earnings or sales using the time series for those variables and extrapolate past trends too far into the future. Under this interpretation, our model is consistent with an overreaction to a long string of bad earnings news and the underweighting of informative bad news of a different type which arrives shortly afterwards.

A related empirical finding is that even for extreme growth stocks that have had several consecutive years of positive earnings news, there is underreaction to quarterly earnings surprises. Our model cannot account for this evidence since it would predict overreaction in this case. To explain this evidence, our model needs to be extended. One possible way to extend the model is to allow investors to estimate the level and the growth rate of earnings separately. Indeed, in reality, investors might use annual earnings numbers over five to seven years to estimate the growth rate but higher frequency quarterly earnings announcements (perhaps combined with other information) to estimate earnings levels. Suppose, for example, that earnings have been growing rapidly over five years, so that an investor using the representativeness heuristic makes an overly optimistic forecast of the future growth rate. Suppose then that a very positive earnings number is announced. Holding the estimated long-run growth rate of earnings constant, investors might still underreact to the quarterly earnings announcement given the high weight this number has for predicting the *level* of earnings when earnings follow a random walk. That is, if such a model is contructed, it can predict underreaction to earnings news in glamour stocks. Such a model could therefore account for more of the available evidence than our simple model.

6. Conclusion

We have presented a parsimonious model of investor sentiment, or of how investors form expectations of future earnings. The model we propose is motivated by a variety of psychological evidence, and in particular by the idea of Griffin and Tversky (1992) that, in making forecasts, people pay too much attention to the strength of the evidence they are presented with and too little attention to its statistical weight. We have supposed that corporate announcements such as those of earnings represent information that is of low strength but significant statistical weight. This assumption has yielded the prediction that stock prices underreact to earnings announcements and similar events. We have further assumed that consistent patterns of news, such as series of good earnings announcements, represent information that is of high strength and low weight. This assumption has yielded a prediction that stock prices overreact to consistent patterns of good or bad news.

Our paper makes reasonable, and empirically supportable, assumptions about the strength and weight of different pieces of evidence and derives empirical implications from these assumptions. However, to push this research further, it is important to develop an priori way of classifying events by their strength and weight, and to make further predictions based on such a classification. The Griffin and Tversky theory predicts most importantly that, holding the weight of information constant, news with more strength would generate a bigger reaction from investors. If news can be classified on a priori grounds, this prediction is testable.

Specifically, the theory predicts that, holding the weight of information constant, one-time strong news events should generate an overreaction. We have not discussed any evidence bearing on this prediction in the paper. However, there does appear to be some evidence consistent with this prediction. For example, stock prices bounced back strongly in the few weeks after the crash of 1987. One interpretation of the crash is that investors overreacted to the news of panic selling by other investors even though there was little fundamental news about security values. Thus the crash was a high-strength, low-weight news event which, according to the theory, should have caused an overreaction. Stein (1989) relatedly finds that long-term option prices overreact to innovations in volatility, another potentially high-strength, low-weight event, since volatility tends to be highly mean-reverting. And Klibanoff et al. (1998) find that the price of a closed-end country fund reacts more strongly to news about its fundamentals when the country whose stocks the fund holds appears on the front page of the newspaper. That is, increasing the strength of the news, holding the weight constant, increases the price reaction. All these are bits of information consistent with the broader implications of the theory. A real test, however, must await a better and more objective way of estimating the strength of news announcements.

Appendix A.

Proposition 1. *If the investor believes that earnings are generated by the regime-switching model described in Section* 4, *then prices satisfy*

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

where p_1 and p_2 are given by the following expressions:

$$p_1 = \frac{1}{\delta}(\gamma_0'(1+\delta)[I(1+\delta)-Q]^{-1}Q\gamma_1),$$

$$p_2 = -\frac{1}{\delta}(\gamma_0'(1+\delta)[I(1+\delta)-Q]^{-1}Q\gamma_2),$$

where

$$\gamma_0' = (1,\ -1, 1,\ -1),$$

$$\gamma_1' = (0, 0, 1, 0),$$

$$\gamma_2' = (1, 0,\ -1, 0),$$

$$Q =$$

$$\begin{pmatrix} (1-\lambda_1)\pi_L & (1-\lambda_1)(1-\pi_L) & \lambda_2\pi_L & \lambda_2(1-\pi_L) \\ (1-\lambda_1)(1-\pi_L) & (1-\lambda_1)\pi_L & \lambda_2(1-\pi_L) & \lambda_2\pi_L \\ \lambda_1\pi_H & \lambda_1(1-\pi_H) & (1-\lambda_2)\pi_H & (1-\lambda_2)(1-\pi_H) \\ \lambda_1(1-\pi_H) & \lambda_1\pi_H & (1-\lambda_2)(1-\pi_H) & (1-\lambda_2)\pi_H \end{pmatrix}.$$

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:

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

Since

$$\mathrm{E}_t(N_{t+1}) = N_t + \mathrm{E}_t(y_{t+1}),$$

$$\mathrm{E}_t(N_{t+2}) = N_t + \mathrm{E}_t(y_{t+1}) + \mathrm{E}_t(y_{t+2}), \text{ and so on,}$$

we have

$$P_t = \frac{1}{\delta}\left\{N_t + \mathrm{E}_t(y_{t+1}) + \frac{\mathrm{E}_t(y_{t+2})}{1+\delta} + \frac{\mathrm{E}_t(y_{t+3})}{(1+\delta)^2} + \cdots\right\}.$$

So the key is to calculate $\mathrm{E}_t(y_{t+j})$. Define

$$q^{t+j} = (q_1^{t+j}, q_2^{t+j}, q_3^{t+j}, q_4^{t+j})',$$

where

$$q_1^{t+j} = \Pr(s_{t+j} = 1, y_{t+j} = y_t | \Phi_t),$$

$$q_2^{t+j} = \Pr(s_{t+j} = 1, y_{t+j} = -y_t | \Phi_t),$$

$$q_3^{t+j} = \Pr(s_{t+j} = 2, y_{t+j} = y_t | \Phi_t),$$

$$q_4^{t+j} = \Pr(s_{t+j} = 2, y_{t+j} = -y_t | \Phi_t),$$

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

Note that

$$\Pr(y_{t+j} = y_t | \Phi_t) = q_1^{t+j} + q_3^{t+j} = \bar{\gamma}' q^{t+j}$$

$$\bar{\gamma}' = (1, 0, 1, 0).$$

The key insight is that

$$q^{t+j} = Q q^{t+j-1},$$

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

$$Q' =$$

$$\begin{array}{ccccc}
 & (1) & (2) & (3) & (4) \\
(1) & (1-\lambda_1)\pi_{\mathrm{L}} & (1-\lambda_1)(1-\pi_{\mathrm{L}}) & \lambda_1\pi_{\mathrm{H}} & \lambda_1(1-\pi_{\mathrm{H}}) \\
(2) & (1-\lambda_1)(1-\pi_{\mathrm{L}}) & (1-\lambda_1)\pi_{\mathrm{L}} & \lambda_1(1-\pi_{\mathrm{H}}) & \lambda_1\pi_{\mathrm{H}} \\
(3) & \lambda_2\pi_{\mathrm{L}} & \lambda_2(1-\pi_{\mathrm{L}}) & (1-\lambda_2)\pi_{\mathrm{H}} & (1-\lambda_2)(1-\pi_{\mathrm{H}}) \\
(4) & \lambda_2(1-\pi_{\mathrm{L}}) & \lambda_2\pi_{\mathrm{L}} & (1-\lambda_2)(1-\pi_H) & (1-\lambda_2)\pi_{\mathrm{H}}
\end{array}$$

where, for example,

$$\Pr(s_{t+j} = 2, y_{t+j} = y_t | s_{t+j-1} = 1, y_{t+j-1} = y_t) = \lambda_1\pi_{\mathrm{H}}.$$

Therefore,

$$q^{t+j} = Q^j q^t = Q^j \begin{pmatrix} q_t \\ 0 \\ 1 - q_t \\ 0 \end{pmatrix}.$$

(Note the distinction between q_t and q^t). Hence,

$$\Pr(y_{t+j} = y_t | \Phi_t) = \bar{\gamma}' Q^j q^t$$

and

$$\mathrm{E}_t(y_{t+j}|\Phi_t) = y_t(\bar{\gamma}'Q^j q^t) + (-y_t)(\underline{\gamma}'Q^j q^t)$$

$$\underline{\gamma}' = (0, 1, 0, 1).$$

Substituting this into the original formula for price gives

$$p_1 = \frac{1}{\delta}(\gamma_0'(1+\delta)[I(1+\delta) - Q]^{-1}Q\gamma_1),$$

$$p_2 = -\frac{1}{\delta}(\gamma_0'(1+\delta)[I(1+\delta) - Q]^{-1}Q\gamma_2),$$

$$\gamma_0' = (1, -1, 1, -1),$$

$$\gamma_1' = (0, 0, 1, 0),$$

$$\gamma_2' = (1, 0, -1, 0).$$

Proposition 2. *Suppose the underlying parameters* π_{L}, π_{H}, λ_1, and λ_2 *satisfy*

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

$$p_2 \geqslant 0,$$

where

$$\underline{k} = \underline{q} + \tfrac{1}{2}\bar{\Delta}(\underline{q}),$$

$$\bar{k} = \underline{q}^e + \tfrac{1}{2}(c_1 + c_2 q_*),$$

$$c_1 = \frac{\bar{\Delta}(\underline{q})\bar{q} - \underline{\Delta}(\bar{q})\underline{q}}{\bar{q} - \underline{q}},$$

$$c_2 = \frac{\underline{\Delta}(\bar{q}) - \bar{\Delta}(\underline{q})}{\bar{q} - \underline{q}},$$

$$q_* = \begin{cases} \bar{q}^e & \text{if} \quad c_2 < 0, \\ \underline{q}^e & \text{if} \quad c_2 \geqslant 0, \end{cases}$$

where $\underline{q}^e$ *and* $\bar{q}^e$ *are bounds on the unconditional mean of the random variable* q_t. *Then the conditions for both underreaction and overreaction given in Section* 5.1. *are satisfied*. (Functions and variables not yet introduced 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 q_t, the probability assigned by the investor at time t to 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 $\bar{\Delta}(q_t)$ denote the increase in the probability assigned to being in regime 1, i.e.,

$$\bar{\Delta}(q) = q_{t+1} - q_t|_{y_{t+1}=-y_t,\ q_t=q}$$
$$= \frac{((1-\lambda_1)q+\lambda_2(1-q))(1-\pi_L)}{((1-\lambda_1)q+\lambda_2(1-q))(1-\pi_L)+((\lambda_1 q+(1-\lambda_2)(1-q))(1-\pi_H)} - q.$$

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

$$\underline{\Delta}(q) = q_t - q_{t+1}|_{y_{t+1}=y_t,\ q_t=q}$$
$$= q - \frac{((1-\lambda_1)q+\lambda_2(1-q))\pi_L}{((1-\lambda_1)q+\lambda_2(1-q))\pi_L+((\lambda_1 q+(1-\lambda_2)(1-q))\pi_H}.$$

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

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

Note that if $q_t \in [\underline{q}, \bar{q}]$, then $q_\tau \in [\underline{q}, \bar{q}]$ 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, q_t will fall every period, tending towards a limit of $\underline{q}$. From the updating formulas, this means that $\underline{q}$ satisfies

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

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

$$\bar{q} = \frac{((1-\lambda_1)\bar{q}+\lambda_2(1-\bar{q}))(1-\pi_L)}{((1-\lambda_1)\bar{q}+\lambda_2(1-\bar{q}))(1-\pi_L)+(\lambda_1 \bar{q}+(1-\lambda_2)(1-\bar{q}))(1-\pi_H)}.$$

In the case of the parameters used for the table in Section 4.2., $\underline{q} = 0.28$ and $\bar{q} = 0.95$.

There is no loss of generality in restricting the support of q_t to the interval $[\underline{q}, \bar{q}]$. Certainly, an investor can have prior beliefs that lie outside this interval, but with probability one, q_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,

$$\mathrm{E}_t(P_{t+1}-P_t|y_t=+y)-\mathrm{E}_t(P_{t+1}-P_t|y_t=-y)>0.$$

Overreaction means that the expected return following a series of positive shocks is smaller than the expected return following a series of negative shocks. In other words, there exists some number $J \geqslant 1$, such that for all $j \geqslant J$,

$$\begin{aligned}&\mathrm{E}_t(P_{t+1}-P_t|y_t=y_{t-1}=\cdots=y_{t-j}=y)\\&\quad-\mathrm{E}_t(P_{t+1}-P_t|y_t=y_{t-1}=\cdots=y_{t-j}=-y)<0.\end{aligned}$$

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

$$f(q)=\mathrm{E}_t(P_{t+1}-P_t|y_t=+y,q_t=q)-\mathrm{E}_t(P_{t+1}-P_t|y_t=-y,q_t=q).$$

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

$$\begin{aligned}P_{t+1}-P_t&=\frac{y_{t+1}}{\delta}+(y_{t+1}-y_t)(p_1-p_2q_t)-y_tp_2(q_{t+1}-q_t)\\&\quad-(y_{t+1}-y_t)p_2(q_{t+1}-q_t),\end{aligned}$$

we find

$$\begin{aligned}&\mathrm{E}_t(P_{t+1}-P_t|y_t=+y,q_t=q)=\frac{1}{2}\left(\frac{y}{\delta}+yp_2\underline{\Delta}(q)\right)\\&\quad+\frac{1}{2}\left(-\frac{y}{\delta}-2y(p_1-p_2q)-yp_2\bar{\Delta}(q)+2yp_2\bar{\Delta}(q)\right)\\&\quad=y(p_2q-p_1)+\tfrac{1}{2}yp_2(\bar{\Delta}(q)+\underline{\Delta}(q))\end{aligned}$$

Further, it is easily checked that

$$\mathrm{E}_t(P_{t+1}-P_t|y_t=+y,q_t=q)=-\mathrm{E}_t(P_{t+1}-P_t|y_t=-y,q_t=q)$$

and hence that

$$f(q)=2y(p_2q-p_1)+yp_2(\bar{\Delta}(q)+\underline{\Delta}(q)).$$

First, we show that a sufficient condition for overreaction is $f(\underline{q})<0$.

If this condition holds, it implies

$$\mathrm{E}_t(P_{t+1} - P_t | y_t = +y, q_t = \underline{q}) < \mathrm{E}_t(P_{t+1} - P_t | y_t = -y, q_t = \underline{q}).$$

Now as $j \to \infty$,

$$\mathrm{E}_t(P_{t+1} - P_t | y_t = y_{t-1} = \cdots = y_{t-j} = y) \to \mathrm{E}_t(P_{t+1} - P_t | y_t = +y, q_t = \underline{q})$$

and

$$\mathrm{E}_t(P_{t+1} - P_t | y_t = y_{t-1} = \cdots = y_{t-j} = -y)$$
$$\to \mathrm{E}_t(P_{t+1} - P_t | y_t = -y, q_t = \underline{q}).$$

Therefore, for $\forall j \geqslant J$ sufficiently large, it must be true that

$$\mathrm{E}_t(P_{t+1} - P_t | y_t = y_{t-1} = \cdots = y_{t-j} = y)$$
$$< \mathrm{E}_t(P_{t+1} - P_t | y_t = y_{t-1} = \cdots = y_{t-j} = -y),$$

which is nothing other than our original definition of overreaction.

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

$$2y(p_2\underline{q} - p_1) + yp_2(\bar{\Delta}(\underline{q}) + \underline{\Delta}(\underline{q})) < 0,$$

we obtain

$$p_1 > p_2\left(\underline{q} + \frac{\bar{\Delta}(\underline{q})}{2}\right) \tag{A.1}$$

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 $f(q)$ as

$$\mathrm{E}_q(f(q)) > 0,$$

where E_q denotes an expectation taken over the unconditional distribution of q. Rewriting this, we obtain:

$$2yp_2\mathrm{E}(q) - 2yp_1 + yp_2\mathrm{E}_q(\bar{\Delta}(q) + \underline{\Delta}(q)) > 0,$$

and hence,

$$p_1 < p_2\left(\mathrm{E}(q) + \frac{\mathrm{E}_q(\bar{\Delta}(q) + \underline{\Delta}(q))}{2}\right). \tag{A.2}$$

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. In the remainder of the proof, we construct a number $\bar{k}$ where

$$\bar{k} < \mathrm{E}(q) + \frac{\mathrm{E}_q(\bar{\Delta}(q) + \underline{\Delta}(q))}{2}.$$

This makes $p_1 < p_2\bar{k}$ a sufficient condition for Eq. (A.2). Of course, this assumes that $p_2 \geqslant 0$, and so we impose this as an additional constraint to be satisfied. In practice, we find that for the ranges of π_L, π_H, λ_1, and λ_2 allowed by the model, p_2 is always positive. However, we do not attempt a proof of this.

The first step in bounding the expression $\mathrm{E}(q) + \frac{1}{2}\mathrm{E}_q(\bar{\Delta}(q) + \underline{\Delta}(q))$ is to bound $\mathrm{E}(q)$. To do this, note that

$$\begin{aligned}\mathrm{E}(q_t) &= \mathrm{E}(q_{t+1}) = \mathrm{E}_{q_t}(\mathrm{E}(q_{t+1}|q_t)) \\ &= \mathrm{E}_{q_t}(\tfrac{1}{2}(q_t + \bar{\Delta}(q_t)) + \tfrac{1}{2}(q_t - \underline{\Delta}(q_t))) \\ &= \mathrm{E}_q(g(q)).\end{aligned}$$

Consider the function $g(q)$ defined on $[\underline{q}, \bar{q}]$. 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 $g(q)$, namely $(\underline{q}, g(\underline{q}))$ and $(\bar{q}, g(\bar{q}))$. In other words, suppose that we bound $g(q)$ above by $\bar{g}(q) = a + bq$. The slope of this line is

$$b = \frac{g(\bar{q}) - g(\underline{q})}{\bar{q} - \underline{q}} = \frac{(\bar{q} - \underline{q}) - \frac{1}{2}(\underline{\Delta}(\bar{q}) + \bar{\Delta}(\underline{q}))}{\bar{q} - \underline{q}} < 1,$$

and a will be such that

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

Given that

$$\mathrm{E}_q(g(q) - q) = 0,$$

we must have

$$\mathrm{E}_q(\bar{g}(q) - q) \geqslant 0$$

or

$$\mathrm{E}(a + bq - q) \geqslant 0$$

$$\mathrm{E}(q) \leqslant \frac{a}{1 - b},$$

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

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

$$(\bar{\Delta} + \underline{\Delta})(q) > \left(\frac{q - \underline{q}}{\bar{q} - \underline{q}}\right)\underline{\Delta}(\bar{q}) + \left(\frac{\bar{q} - q}{\bar{q} - \underline{q}}\right)\bar{\Delta}(\underline{q}),$$
$$= c_1 + c_2 q$$

where

$$c_1 = \frac{\bar{\Delta}(\underline{q})\bar{q} - \underline{\Delta}(\bar{q})\underline{q}}{\bar{q} - \underline{q}},$$

$$c_2 = \frac{\underline{\Delta}(\bar{q}) - \bar{\Delta}(\underline{q})}{\bar{q} - \underline{q}}.$$

Therefore,

$$\mathrm{E}(q) + \tfrac{1}{2}\mathrm{E}(\bar{\Delta}(q) + \underline{\Delta}(q)) \geqslant \underline{q}^e + \tfrac{1}{2}\mathrm{E}(c_1 + c_2 q) \geqslant \underline{q}^e + \tfrac{1}{2}(c_1 + c_2 q_*),$$

where

$$q_* = \begin{cases} \bar{q}^e & \text{if} \quad c_2 < 0, \\ \underline{q}^e & \text{if} \quad c_2 \geqslant 0. \end{cases}$$

This completes the proof of the proposition.

References

Andreassen, P., Kraus, S., 1990. Judgmental extrapolation and the salience of change. Journal of Forecasting 9, 347–372.

Barsky, R., De Long, J.B., 1993. Why does the stock market fluctuate? Quarterly Journal of Economics 108, 291–311.

Bernard, V., 1992. Stock price reactions to earnings announcements In: Thaler, R. (Ed.), Advances in Behavioral Finance. Russell Sage Foundation, New York, pp. 303–340.

Bernard, V., Thomas, J., 1989. Post-earnings announcement drift: delayed price response or risk premium? Journal of Accounting Research, (Suppl.) 27, 1–36.

Bernard, V., Thomas, J., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics 13, 305–341.

Campbell, J.Y., Shiller, R., 1988. Stock prices, earnings, and expected dividends. Journal of Finance 43, 661–676.

Chan, L., Jegadeesh, N., Lakonishok, J., 1997. Momentum strategies. Journal of Finance 51, 1681–1713.

Chopra, N., Lakonishok, J., Ritter, J., 1992. Measuring abnormal performance: do stocks overreact? Journal of Financial Economics 31, 235–268.

Cutler, D., Poterba, J., Summers, L., 1991. Speculative dynamics. Review of Economic Studies 58, 529–546.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. A theory of overconfidence, self-attribution, and security market under- and over-reactions. Journal of Finance 53, in press.

De Bondt, W., 1993. Betting on trends: intuitive forecasts of financial risk and return. International Journal of Forecasting 9, 355–371.

De Bondt, W., Thaler, R., 1985. Does the stock market overreact? Journal of Finance 40, 793-808.

De Bondt, W., Thaler, R., 1987. Further evidence of investor overreaction and stock market seasonality. Journal of Finance 42, 557–581.

De Long, J.B., Shleifer, A., Summers, L., Waldmann, R., 1990a. Noise trader risk in financial markets. Journal of Political Economy 98, 703–738.

De Long, J.B., Shleifer, A., Summers, L., Waldmann, R., 1990b. Positive feedback investment strategies and destabilizing rational speculation. Journal of Finance 45, 375–395.

Dreman, D., Berry, M., 1995. Overreaction, underreaction, and the low-P/E effect. Financial Analysts Journal 51, 21–30.

Edwards, W., 1968. Conservatism in human information processing. In: Kleinmutz, B. (Ed.), Formal Representation of Human Judgment. John Wiley and Sons, New York, pp. 17–52.

Fama, E., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49, 283–306 (this issue).

Fama, E., French, K., 1988. Permanent and temporary components of stock prices. Journal of Political Economy 96, 246–273.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E., French, K., 1996. Multifactor explanations of asset pricing anomalies. Journal of Finance 51, 55–84.

Fama, E., French K., 1998. Value versus growth: the international evidence. Journal of Finance 53, in press.

Griffin, D., Tversky, A., 1992. The weighing of evidence and the determinants of confidence. Cognitive Psychology 24, 411–435.

Haugen, R., Baker, N., 1996. Commonality in the determinants of expected stock returns. Journal of Financial Economics 41, 401–439.

Ikenberry, D., Lakonishok, J., Vermaelen, T., 1995. Market underreaction to open market share repurchases. Journal of Financial Economics 39, 181–208.

Ikenberry, D., Rankine, G., Stice, E., 1996. What do stock splits really signal? Journal of Financial and Quantitative Analysis 31, 357–375.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.

Klibanoff, P., Lamont, O., Wizman, T., 1998. Investor reaction to salient news in closed-end country funds. Journal of Finance 53, 673–700.

Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–1578.

La Porta, R., 1996. Expectations and the cross-section of returns. Journal of Finance 51, 1715–1742.

La Porta, R., Lakonishok, J., Shleifer, A., Vishny, R., 1997. Good news for value stocks: further evidence on market efficiency. Journal of Finance 52, 859–874.

Lee, C., Shleifer, A., Thaler, R., 1991. Investor sentiment and the closed-end fund puzzle. Journal of Finance 46, 75–110.

Lehmann, B., 1990. Fads, martingales, and market efficiency. Quarterly Journal of Economics 105, 1–28.

Loughran, T., Ritter, J., 1995. The new issues puzzle. Journal of Finance 50, 23–51.

Michaely, R., Thaler, R., Womack, K., 1995. Price reactions to dividend initiations and omissions: overreaction or drift? Journal of Finance 50, 573–608.

Poterba, J., Summers, L., 1988. Mean-reversion in stock prices: evidence and implications. Journal of Financial Economics 22, 27–59.

Rouwenhorst, G., 1997. International momentum strategies. Journal of Finance 53, 267–284.

Samuelson, P., 1965. Proof that properly anticipated prices fluctuate randomly. Industrial Management Review 6, 41–49.

Shiller, R., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? American Economic Review 71, 421–436.

Shleifer, A., Vishny, R., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.

Spiess, K., Affleck-Graves, J., 1995. Underperformance in long-run stock returns following seasoned equity offerings. Journal of Financial Economics 38, 243–267.

Stein, J., 1989. Overreactions in the options market. Journal of Finance 44, 1011–1023.

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Science 185, 1124–1131.

Zarowin, P., 1989. Does the stock market overreact to corporate earnings information? Journal of Finance 44, 1385–1400.