A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets

Harrison Hong and Jeremy C. Stein *

First draft: May 1997
This version: November 1998

Abstract: We model a market populated by two groups of boundedly rational agents: "newswatchers" and "momentum traders". Each newswatcher observes some private information, but fails to extract other newswatchers' information from prices. If information diffuses gradually across the population, prices underreact in the short run. The underreaction means that the momentum traders can profit by trend-chasing. However, if they can only implement simple (i.e., univariate) strategies, their attempts at arbitrage must inevitably lead to overreaction at long horizons. In addition to providing a unified account of under- and overreactions, the model generates several other distinctive implications.

* Stanford Business School and MIT Sloan School of Management and NBER. This research is supported by the National Science Foundation and the Finance Research Center at MIT. We are grateful to Denis Gromb, René Stulz, an anonymous referee, and seminar participants at MIT, Michigan, Wharton, Duke, UCLA, Berkeley, Stanford and Illinois for helpful comments and suggestions. Thanks also to Melissa Cunniffe for help in preparing the manuscript.

Over the last several years, a large volume of empirical work has documented a variety of ways in which asset returns can be predicted based on publicly available information. Although different studies have used a host of different predictive variables, many of the results can be thought of as belonging to one of two broad categories of phenomena. On the one hand, returns appear to exhibit continuation, or momentum, in the short to medium run. On the other hand, there is also a tendency toward reversals, or fundamental reversion, in the long run.[1]

It is becoming increasingly clear that traditional asset-pricing models--such as the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965), Ross's (1976) arbitrage pricing theory (APT) or Merton's (1973) intertemporal capital asset pricing model (ICAPM)--have a hard time explaining the growing set of stylized facts. In the context of these models, all of the predictable patterns in asset returns, at both short and long horizons, must ultimately be traced to differences in loadings on economically meaningful risk factors. And there is little affirmative evidence at this point to suggest that this can be done.

As an alternative to these traditional models, many are turning to "behavioral" theories, where "behavioral" can be broadly construed as involving some departure from the classical assumptions of strict rationality and unlimited computational capacity on the part of investors. But the difficulty with this approach is that there are a potentially huge number of such departures that one might entertain, so it is hard to know where to start.

In order to impose some discipline on the process, it is useful to articulate the criteria that a new theory should be expected to satisfy. There seems to be broad agreement that to be successful, any candidate theory should, at a minimum: 1) rest on assumptions about investor behavior that are either a priori plausible or consistent with casual observation; 2) explain the existing evidence in a parsimonious and unified way; and 3) make a number of further predictions which can be subject to "out-of sample" testing and which are ultimately validated. Fama (1998) puts particular emphasis on the latter two criteria: "Following the standard scientific rule, market efficiency can only be replaced by a better model...The alternative has a daunting task. It must specify what it is about investor psychology that causes simultaneous underreaction to some types of events and overreaction to others...And the alternative must present well-defined hypotheses, themselves potentially rejectable by empirical tests."

A couple of recent papers take up this challenge. Both Barberis, Shleifer and Vishny (BSV) (1998) and Daniel, Hirshleifer and Subrahmanyam (DHS) (1998) assume that prices are driven by a single representative agent, and then posit a small number of cognitive biases that this representative agent might have. They then investigate the extent to which these biases are sufficient to simultaneously deliver both short-horizon continuation and long-horizon reversals.[2]

In this paper, we pursue the same goal as BSV and DHS, that of building a unified behavioral model. However, we adopt a fundamentally different approach. Rather than trying to say much about the psychology of the representative agent, our emphasis is on the interaction between heterogeneous agents. To put it loosely, less of the action in our model comes from particular cognitive biases that we ascribe to individual traders, and more of it comes from the way these traders interact with one another.

More specifically, our model features two types of agents, whom we term "newswatchers" and "momentum traders". Neither type is fully rational in the usual sense. Rather, each is boundedly rational, with the bounded rationality being of a simple form: each type of agent is only able to "process" some subset of the available public information.[3] The newswatchers make forecasts based on signals that they privately observe about future fundamentals; their limitation is that they do not condition on current or past prices. Momentum traders, in contrast, do condition on past price changes. However, their limitation is that their forecasts must be "simple" (i.e., univariate) functions of the history of past prices.[4]

In addition to imposing these two constraints on the information processing abilities of our traders, we make one further assumption, which is more orthodox in nature: private information diffuses gradually across the newswatcher population. All our conclusions then flow from these three key assumptions. We begin by showing that when only newswatchers are active, prices adjust slowly to new information--there is underreaction but never overreaction. As is made clear later, this result follows naturally from combining gradual information diffusion with the assumption that newswatchers do not extract information from prices.

Next, we add the momentum traders. It is tempting to conjecture that because the momentum traders can condition on past prices, they arbitrage away any underreaction left behind by the newswatchers; with sufficient risk tolerance, one might expect that they would force the market to become approximately efficient. However, it turns out that this intuition is incomplete, if momentum traders are limited to simple strategies. For example, suppose that a momentum trader at time t must base his trade only on the price change over some prior interval, say from t-2 to t-1. We show that in this case, momentum traders' attempts to profit from the underreaction caused by newswatchers lead to a perverse outcome: the initial reaction of prices in the direction of fundamentals is indeed accelerated, but this comes at the expense of creating an eventual overreaction to any news. This is true even when momentum traders are risk neutral.

Again, the key to this result is the assumption that momentum traders use simple strategies --

i.e., do not condition on all public information. Continuing with the example, if a momentum trader's order at time t is restricted to being a function of just the price change from t-2 to t-1, it is clear that it must be an increasing function. <u>On average</u>, this simple trend-chasing strategy makes money. But if one could condition on more information, it would become apparent that the strategy does better in some circumstances than in others. In particular, the strategy earns the bulk of its profits early in the "momentum cycle"--by which we mean shortly after substantial news has arrived to the newswatchers--and loses money late in the cycle, by which time prices have already overshot long-run equilibrium values.

To see this point, suppose that there is a single dose of good news at time t and no change in fundamentals after that. The newswatchers cause prices to jump at time t, but not far enough, so that they are still below their long-run values. At time t+1 there is a round of momentum purchases, and those momentum buyers who get in at this time make money. But this round of momentum trading creates a further price increase, which sets off more momentum buying, and so on. Later momentum buyers--i.e., those buying at t+i for some i--lose money, because they get in at a price above the long-run equilibrium.

Thus a crucial insight is that "early" momentum buyers impose a negative externality on "late" momentum buyers.[5] Ideally, one uses a momentum strategy because a price increase signals that there is good news about fundamentals out there that is not yet fully incorporated into prices. But sometimes, a price increase is the result not of news but just of previous rounds of momentum trade. Because momentum traders cannot directly condition on whether or not news has recently arrived, they do not know whether they are early or late in the cycle. Hence they must live with this externality, and accept the fact that sometimes they buy when earlier rounds of momentum trading have pushed prices past long-run equilibrium values.

Although we make two distinct bounded-rationality assumptions, our model can be said to "unify" underreaction and overreaction in the following sense. We begin by modelling a tendency for one group of traders to underreact to private information. We then show that when a second group of traders tries to exploit this underreaction with a simple arbitrage strategy, they only partially eliminate it, and in so doing, create an excessive momentum in prices that inevitably culminates in overreaction. Thus, the very existence of underreaction sows the seeds for overreaction, by making it profitable for momentum traders to enter the market. Or said differently, the unity lies in the fact that our model gets both underreaction and overreaction out of just one primitive type of shock: gradually-diffusing news about fundamentals. There are no other exogenous shocks to investor sentiment and no liquidity-motivated trades.

In what follows, we develop a simple infinite-horizon model that captures these ideas. We begin in Section I by giving an overview of the empirical evidence that motivates our work. In Section II, we present and solve the basic model, and do a number of comparative statics experiments. Section III contains several extensions. In Section IV, we draw out the model's empirical implications. Section V discusses related work, and Section VI concludes.

I. <u>Evidence of Continuation and Reversals</u>

A <u>Continuation</u>

The continuation evidence can be decomposed along the following lines. First, returns tend to exhibit unconditional positive serial correlation at horizons on the order of three to twelve months. This is true both in cross-sections of individual stocks (Jegadeesh and Titman (1993)) and for a variety of broad asset classes (Cutler, Poterba and Summers (1991)).[6] One possible interpretation of this unconditional evidence--which fits with the spirit of the model below--is that information which is initially private is incorporated into prices only gradually.

Second, conditional on observable public events, stocks tend to experience post-event drift in the same direction as the initial event impact. The types of events that have been examined in detail and that fit this pattern include: earnings announcements (perhaps the most-studied type of event in this genre; see e.g. Bernard (1992) for an overview); stock issues and repurchases; dividend initiations and omissions; and analyst recommendations.[7] Recent work by Chan, Jegadeesh and Lakonishok (1996) shows that these two types of continuation are distinct: in a multiple regression, both past returns and public earnings surprises help to predict subsequent returns at horizons of six months and one year.

B Reversals and Fundamental Reversion

One of the first and most influential papers in the reversals category is DeBondt and Thaler (1985), who find that stock returns are negatively correlated at long horizons. Specifically, stocks that have had the lowest returns over any given five-year period tend to have high returns over the subsequent five years, and vice-versa.[8] A common interpretation of this result is that when there is a sustained streak of good news about an asset, its price overshoots its "fundamental value", and ultimately must experience a correction. More recent work in the same spirit has continued to focus on long-horizon predictability, but has used what are arguably more refined indicators of fundamental value, such as book-to-market, and cashflow-to-price ratios. (See, e.g., Fama and French (1992), Lakonishok, Shleifer and Vishny (1994).)[9]

C Is It Risk?

In principle, the patterns noted above could be consistent with traditional models, to the extent that they reflect variations in risk, either over time or across assets. Fama and French (1993, 1996) argue that many of the long-horizon results--such as return reversals, the book-to-market effect, and the cashflow-to-price effect--can be largely subsumed within a three-factor model that they interpret as a variant of the APT or ICAPM. However, this position has been controversial, since there is little affirmative evidence that the Fama-French factors correspond to economically meaningful risks. Indeed, several recent papers demonstrate that the contrarian strategies that exploit long-horizon overreaction are not significantly riskier than average.[10] There seems to be more of a consensus that the short-horizon underreaction evidence cannot be explained in terms of risk. Bernard and Thomas (1989) reject risk as an explanation for post-earnings-announcement drift. And Fama and French (1996) remark that the continuation results of Jegadeesh and Titman (1993) constitute the "main embarrassment" for their three-factor model.

II. The Model

A Price Formation With Newswatchers Only

As mentioned above, our model features two classes of traders, newswatchers and momentum traders. We begin by describing how the model works when only the newswatchers are present. At every time t, the newswatchers trade claims on a risky asset. This asset pays a single liquidating dividend at some later time T. The ultimate value of this liquidating dividend can be written as: $D_T = D_0 + \Sigma^T_{j=0} \epsilon_j$, where all the ϵ's are independently distributed, mean-zero normal random variables with variance σ^2. Throughout, we consider the limiting case where T goes to infinity. This simplifies matters by allowing us to focus on steady-state trading strategies--i.e., strategies that do not depend on how close we are to the terminal date.[11]

In order to capture the idea that information moves gradually across the newswatcher population, we divide this population into z equal-sized groups. We also assume that every dividend innovation ϵ_j can be decomposed into z independent sub-innovations, each with the same variance σ^2/z: $\epsilon_j = \epsilon^1_j + ... + \epsilon^z_j$. The timing of information release is then as follows. At time t, news about ϵ_{t+z-1} begins to spread. Specifically, at time t, newswatcher group 1 observes ϵ^1_{t+z-1}, group 2 observes ϵ^2_{t+z-1}, and so forth, through group z, which observes ϵ^z_{t+z-1}. Thus at time t, each sub-innovation of ϵ_{t+z-1} has been seen by a fraction 1/z of the total population.

Next, at time t+1, the groups "rotate", so that group 1 now observes ϵ^2_{t+z-1}, group 2 observes ϵ^3_{t+z-1}, and so forth, through group z, which now observes ϵ^1_{t+z-1}. Thus at time t+1 the information has spread further, and each sub-innovation of ϵ_{t+z-1} has been seen by a fraction 2/z of the total population. This rotation process continues until time t+z-1, at which point every one of the z groups has directly observed each of the sub-innovations that comprise ϵ_{t+z-1}. So ϵ_{t+z-1} has become totally public by time t+z-1. Although this formulation may seem unnecessarily awkward, the rotation feature is useful, because it implies that even as information moves slowly across the population, everybody is on average equally well-informed.[12] This symmetry makes it transparently simple to solve for prices, as is seen momentarily.

In this context, the parameter z can be thought of as a proxy for the (linear) rate of information flow--higher values of z imply slower information diffusion. Of course, the notion that information spreads slowly is more appropriate for some purposes than others. In particular, this construct is fine if our goal is to capture the sort of underreaction that shows up empirically as unconditional positive correlation in returns at short horizons. However, if we are also interested in capturing phenomena like post-earnings-announcement drift--where there is apparently underreaction even to data that is made available to everyone simultaneously--we need to embellish the model. We discuss this embellishment later on; for now it is easiest to think of the model as only speaking to the unconditional evidence on underreaction.

All the newswatchers have constant absolute risk aversion (CARA) utility with the same risk-aversion parameter, and all live until the terminal date T. The riskless interest rate is normalized to zero, and the supply of the asset is fixed at Q. So far, all these assumptions are completely orthodox. We now make two that are less conventional. First, at every time t, newswatchers formulate their asset demands based on the static-optimization notion that they buy and hold until the liquidating dividend at time T.[13] Second, and more critically, while newswatchers can condition on the information sets described above, they do not condition on current or past prices. In other words, our equilibrium concept is a Walrasian equilibrium with private valuations, as opposed to a fully revealing rational expectations equilibrium.

As suggested in the Introduction, these two unconventional assumptions can be motivated based on a simple form of bounded rationality. One can think of the newswatchers as having their hands full just figuring out the implications of the ϵ's for the terminal dividend D_T. This leaves them unable to also use current and past market prices to form more sophisticated forecasts of D_T (our second assumption); it also leaves them unable to make any forecasts of future price changes, and hence unable to implement dynamic strategies (our first assumption).

Given these assumptions, and the symmetry of our set-up, the conditional variance of fundamentals is the same for all newswatchers, and the price at time t is given by:

$$P_t = D_t + \{(z-1)\epsilon_{t+1} + (z-2)\epsilon_{t+2} + ... + \epsilon_{t+z-1}\}/z - \theta Q \qquad (1)$$

where θ is a function of newswatchers' risk aversion and the variance of the ϵ's. For simplicity, we normalize the risk aversion so that $\theta = 1$ hereafter. In words, equation (1) says that the new information works its way linearly into the price over z periods. This implies that there is positive serial correlation of returns over short horizons (of length less than z). Note also that prices never overshoot their long-run values, or equivalently, that there is never any negative serial correlation in returns at any horizon.

Even given the eminently plausible assumption that private information diffuses gradually across the population of newswatchers, the gradual-price-adjustment result in equation (1) hinges critically on the further assumption that newswatchers do not condition on prices. For if they did --and as long as Q is nonstochastic--the logic of Grossman (1976) would imply a fully revealing equilibrium, with a price P^*_t, following a random walk given by (for $\theta = 1$):[14]

$$P^*_t = D_{t+z-1} - Q \qquad (2)$$

We should therefore stress that we view the underreaction result embodied in equation (1) to be nothing more than a point of departure. As such, it raises an obvious next question: even if newswatchers are too busy processing fundamental data to incorporate prices into their forecasts, can't some other group of traders focus exclusively on price-based forecasting, and in so doing generate an outcome close to the rational expectations equilibrium of equation (2)? It is to this central question that we turn next, by adding the momentum traders into the mix.

B Adding Momentum Traders to the Model

Momentum traders also have CARA utility. Unlike the newswatchers, however, they have finite horizons. In particular, at every time t, a new generation of momentum traders enters the market. Every trader in this generation takes a position, and then holds this position for j periods --that is, until time t+j. For modelling purposes, we treat the momentum traders' horizon j as an exogenous parameter.

The momentum traders transact with the newswatchers by means of market orders. They submit quantity orders, not knowing the price at which these orders will be executed. The price is then determined by the competition among the newswatchers, who double as market-makers in this set-up. Thus in deciding the size of their orders, the momentum traders at time t must try to predict $(P_{t+j} - P_t)$. To do so, they make forecasts based on past price changes. We assume that these forecasts take an especially simple form: the only conditioning variable is the cumulative price change over the past k periods, i.e., $(P_{t-1} - P_{t-k-1})$.

As it turns out, the exact value of k is not that important, so in what follows we simplify things by setting k = 1, and using $(P_{t-1} - P_{t-2}) \equiv \Delta P_{t-1}$ as the time-t forecasting variable.[15] What is more significant is that we restrict the momentum traders to making univariate forecasts based on past price changes. If, in contrast, we allow them to make forecasts using n lags of price changes, giving different weights to each of the n lags, we suspect that for sufficiently large n, many of the results that we present below would go away. Again, the motivation is a crude notion of bounded rationality: momentum traders simply do not have the computational horsepower to run complicated multivariate regressions.

With k = 1, the order flow from generation-t momentum traders, F_t, is of the form:

$$F_t = A + \phi\Delta P_{t-1} \qquad (3)$$

where the constant A and the elasticity parameter ϕ have to be determined from optimization on the part of the momentum traders. This order flow must be absorbed by the newswatchers. We assume that the newswatchers treat the order flow as an uninformative supply shock. This is consistent with our prior assumption that the newswatchers do not condition on prices. Given that the order flow is a linear function of past price changes, if we allowed the newswatchers to extract information from it, we would be indirectly allowing them to learn from prices.

To streamline things, the order flow from the newswatchers is the <u>only</u> source of supply variation in the model. Given that there are j generations of momentum traders in the market at any point in time, the aggregate supply S_t absorbed by the newswatchers is given by:

$$S_t = Q - \Sigma^j_{i=1} F_{t+1-i} = Q - jA - \Sigma^j_{i=1} \phi\Delta P_{t-i} \qquad (4)$$

We continue to assume that, at any time t, the newswatchers act as if they buy and hold until the liquidating dividend at time T. This implies that prices are given exactly as in equation (1), except that the fixed supply Q is replaced by the variable S_t, yielding:

$$P_t = D_t + \{(z-1)\epsilon_{t+1} + (z-2)\epsilon_{t+2} + ... + \epsilon_{t+z-1}\}/z - Q + jA + \Sigma^j_{i=1} \phi\Delta P_{t-i} \qquad (5)$$

In most of the analysis, the constants Q and A play no role, so we disregard them when it is convenient to do so.

As noted previously, newswatchers' behavior is time-inconsistent. Although at time t they base their demands on the premise that they do not retrade, they violate this to the extent that they are active in later periods. We adopt this time-inconsistent shortcut because it dramatically simplifies the analysis. Otherwise, we face a complex dynamic programming problem, with newswatcher demands at time t depending not only on their forecasts of the liquidating dividend D_T, but also on their predictions for the entire future path of prices.

Two points can be offered in defense of this time-inconsistent simplification. First, it fits with the basic spirit of our approach, which is to have the newswatchers behave in a simple, boundedly rational fashion. Second, we have no reason to believe that it colors any of our important qualitative conclusions. Loosely speaking, we are closing down a "frontrunning" effect, whereby newswatchers buy more aggressively at time t in response to good news, since they know that the news will kick off a series of momentum trades and thereby drive prices up further over the next several periods.[16] Such frontrunning by newswatchers may speed the response of prices to information, thereby mitigating underreaction, but in our set-up it can never wholly eliminate either underreaction or overreaction.[17]

C The Nature of Equilibrium

With all of the assumptions in place, we are now ready to solve the model. The only task is to calculate the equilibrium value of ϕ. Disregarding constants, optimization on the part of the momentum traders implies:

$$\phi\Delta P_{t-1} = \gamma E_M(P_{t+j} - P_t)/\mathrm{var}_M(P_{t+j} - P_t) \tag{6}$$

where γ is the aggregate risk tolerance of the momentum traders, and E_M and var_M denote the mean and variance given their information, which is just ΔP_{t-1}. We can rewrite equation (6) as:

$$\phi = \gamma \mathrm{cov}(P_{t+j} - P_t, \Delta P_{t-1})/\{\mathrm{var}(\Delta P)\mathrm{var}_M(P_{t+j} - P_t)\} \tag{7}$$

The definition of equilibrium is a fixed point such that ϕ is given by equation (7), while at the same time price dynamics satisfy equation (5). We restrict ourselves to studying covariance-stationary equilibria. In the appendix, we prove that a necessary condition for a conjectured equilibrium process to be covariance stationary is that $|\phi| < 1$. Such an equilibrium may not exist for arbitrary parameter values, and we are also unable to generically rule out the possibility of multiple equilibria. However, we prove in the appendix that existence is guaranteed so long as the risk tolerance γ of the momentum traders is sufficiently small, since this in turn ensures that $|\phi|$ is sufficiently small. Moreover, detailed experimentation suggests that a unique covariance-stationary equilibrium does in fact exist for a large range of the parameter space.[18]

In general, it is difficult to solve the model in closed form, and we have to resort to a computational algorithm to find the fixed point. For an arbitrary set of parameter values, we always begin our numerical search for the fixed point at $j = 1$. Given this restriction, we can show that the condition $|\phi| < 1$ is both necessary and sufficient for covariance-stationarity. We also start with a small value of risk tolerance and an initial guess for ϕ of zero. The solutions in this region of the parameter space are well-behaved. Using these solutions, we then move to other regions of the parameter space. This procedure ensures that if there are multiple covariance-stationary equilibria, we would always pick the one with the smallest value of ϕ. We also have a number of sensible checks for when we have moved outside the covariance-stationary region of the parameter space. These are described in the appendix.

Even without doing any computations, we can make several observations about the nature of equilibrium. First, we have:

<u>Lemma 1</u>: In any covariance-stationary equilibrium, $\phi > 0$. That is, momentum traders must rationally behave as trend-chasers.

The lemma is proved in the appendix, but it is trivially easy to see why $\phi = 0$ cannot be an equilibrium. Suppose to the contrary it is. Then prices are given as in the all-newswatcher case in equation (1). And in this case, $\mathrm{cov}(P_{t+j} - P_t, \Delta P_{t-1}) > 0$, so that equation (7) tells us that $\phi > 0$, establishing a contradiction.

We are now in a position to make some qualitative statements about the dynamics of prices. First, let us consider the impulse response of prices to news shocks. The thought experiment here is as follows. At time t, there is a one-unit positive innovation ϵ_{t+z-1} that begins to diffuse among newswatchers. There are no further news shocks from that point on. What does the price path look like?

The answer can be seen by decomposing the price at any time into two components: that attributable to newswatchers, and that attributable to momentum traders. Newswatchers' aggregate estimate of D_T rises from time t to time t+z-1, by which time they have completely incorporated the news shock into their forecasts. Thus by time t+z-1, the price is just right in the absence of any order flow from momentum traders. But with $\phi > 0$, any positive news shock must generate an initially positive impulse to momentum-trader order flow. Moreover, the cumulative order flow must be increasing until at least time t+j, since none of the momentum trades stimulated by the shock begin to be unwound until t+j+1. This sort of reasoning leads to the following conclusions:

Proposition 1: In any covariance-stationary equilibrium, given a positive one-unit shock ϵ_{t+z-1} that first begins to diffuse among newswatchers at time t:

i) there is always overreaction, in the sense that the cumulative impulse response of prices peaks at a value that is strictly greater than one;

ii) if the momentum traders' horizon j satisfies $j \geq z-1$, the cumulative impulse response peaks at t+j, and then begins to decline, eventually converging to one;

iii) if $j < z-1$, the cumulative impulse response peaks no earlier than t+j, and eventually converges to one.

In addition to the impulse response function, it is also interesting to consider the autocovariances of prices at various horizons. We can develop some rough intuition about these autocovariances by considering the limiting case where the risk tolerance of the momentum traders γ goes to infinity. In this case, equation (7) implies that the equilibrium must have the property that $\text{cov}(P_{t+j} - P_t, \Delta P_{t-1}) = 0$. Expanding this expression, we can write:

$$\text{cov}(\Delta P_{t+1}, \Delta P_{t-1}) + \text{cov}(\Delta P_{t+2}, \Delta P_{t-1}) + \ldots\ldots + \text{cov}(\Delta P_{t+j}, \Delta P_{t-1}) = 0 \qquad (8)$$

Equation (8) allows us to state the following:

Proposition 2: In any covariance-stationary equilibrium, if price changes are positively correlated at short horizons (e.g., if cov(ΔP_{t+1}, ΔP_{t-1}) > 0), then with risk-neutral momentum traders they are negatively correlated at a horizon no longer than j+1--i.e., it must be that cov(ΔP_{t+i}, ΔP_{t-1}) < 0 for some $i \leq j$.

It is useful to explore the differences between Propositions 1 and 2 in some detail, since at first glance, it might appear that they are somewhat contradictory. On the one hand, Proposition 1 says that in response to good news, there is continued upward momentum in prices for at least j periods, and possibly more (if j < z-1). On the other hand, Proposition 2 suggests that price changes begin to be reversed within j+1 periods, and quite possibly sooner than that.

The two propositions can be reconciled by noting that the former is a conditional statement--i.e., it talks about the path of prices from time t onward, conditional on there having been a news shock at time t. Thus Proposition 1 implies that if a trader somehow knows for sure that there is a news shock at time t, he could make a strictly positive expected profit by buying at this time and holding until time t+j. One might term such a strategy "buying early in the momentum cycle"--i.e., buying immediately on the heels of news arrival. But of course, such a strategy is not available to the momentum traders in our model, since they cannot condition directly on the ϵ's.

In contrast, Proposition 2 is an unconditional statement about the autocovariance of prices. It flows from the requirement that if a trader buys at time t in response to an unconditional price increase at time t-1, and then holds until t+j, he makes zero profits on average. This zero-profit requirement in turn must hold when momentum traders are risk-neutral, because the unconditional strategy <u>is</u> available to them.

There is a simple reason why an unconditional strategy of buying following any price increase does not work as well as the conditional strategy of buying only following directly observed good news: not all price increases are news-driven. In particular, a trader who buys based on a price increase observed at time t runs the following risk. It may be "late" in the momentum cycle, in the sense that there has not been any good news for the last several periods. Say the last good news hit at t-i. If this is the case, the price increase at time t is just due to a late round of momentum buying. And those earlier momentum purchases kicked off by the news at t-i will begin to be unwound in the very near future (specifically, at t-i+j+1) causing the trader to experience losses well before the end of his trading horizon.

This discussion highlights the key spillover effect that drives our results. A momentum trader who is fortunate enough to buy shortly after the arrival of good news imposes a negative externality on those that follow him. He does so by creating a further price increase that the next generation partially misinterprets as more good news. This causes the next generation to buy, and so on. At some point, the buying has gone too far, and the price overshoots the level warranted by the original news. Given the inability of momentum traders to condition directly on the ϵ's, everybody in the chain is behaving as rationally as possible, but the externality creates an apparently irrational outcome in the market as a whole.

D <u>Winners and Losers</u>

A natural question is whether the bounded rationality of either the newswatchers or the momentum traders causes them to systematically lose money. In general, both groups can earn positive expected returns as long as the net supply Q of the asset is positive. Consider first the case where Q = 0. In this case, it can be shown that the momentum traders earn positive returns, as long as their risk tolerance is finite. Because with Q = 0, this is a zero-sum game, it must therefore be that the newswatchers lose money. The one exception is when momentum traders are risk-neutral, and both groups break even.[19]

When Q > 0, the game becomes positive-sum, as there is a return to risk-sharing that can be divided between the two groups. Thus even though the newswatchers may effectively lose some money on a trading basis to the momentum traders, this can be more than offset by their returns from risk-sharing, and they can make a net profit. Again, in the limit where the momentum traders become risk-neutral, both groups break even. The logic is similar to that with Q = 0, because risk-neutrality on the part of momentum traders dissipates all the risk-sharing profits, restoring the zero-sum nature of the game.

E Numerical Comparative Statics

In order to develop a better feeling for the properties of the model, we perform a variety of numerical comparative statics exercises.[20] For each set of parameter values, we calculate the following five numbers: i) the equilibrium value of ϕ; ii) the unconditional standard deviation of monthly returns ΔP; iii) the standard deviation of the pricing error relative to a rational expectations benchmark, $(P_t - P^*_t)$; iv) the cumulative impulse response of prices to a one-unit ϵ shock; and v) the autocorrelations of returns. The detailed calculations are shown in the appendix; here we use plots of the impulse responses to convey the broad intuition.

We begin in Figure 1 by investigating the effects of changing the momentum traders' horizon j. We hold the information-diffusion parameter z fixed at 12 months, and set the standard deviation of the fundamental ϵ shocks equal to 0.5 per month. Finally, we assume that the aggregate risk tolerance of the momentum traders, γ, equals 1/3.[21] We then experiment with values of j ranging from six to 18 months. As a baseline, focus first on the case where j = 12 months. Consistent with Proposition 1, the impulse response function peaks 12 months after an ϵ shock, reaching a value of 1.342. In other words, at the peak, prices overshoot the change in long-run fundamentals by 34.2 percent. After the peak, prices eventually converge back to 1.00, although not in a monotonic fashion--rather, there are a series of damped oscillations as the momentum-trading effects gradually wring themselves out.

Now ask what happens as j is varied. As can be seen from the figure, the effects on the impulse response function are non-monotonic. For example, with j = 6, the impulse response peaks at 1.265, while with j = 18, the peak reaches 1.252, neither as high as in the case where j = 12. This non-monotonicity arises because of two competing effects. On the one hand, an increase in j means that there are more generations of momentum traders active in the market at any one time; hence their cumulative effect should be stronger, all else equal. On the other hand, the momentum traders rationally recognize the dangers of having a longer horizon--there is a greater risk that they get caught trading late in the momentum cycle. As a result, they trade less aggressively, so that ϕ is decreasing in j.

A more clear-cut result appears to emerge when we consider the effect of j on the time pattern of autocorrelations. As suggested by Figure 1--and confirmed by the calculations in the appendix--the smaller is j, the faster the autocorrelations begin to turn negative. For example, with j = 6, the first negative autocorrelation occurs at a lag of 6 months, while with j = 18, the first negative autocorrelation occurs at a lag of 12 months. Thus the intuition from Proposition 2 seems to carry over to the case of non-zero risk aversion.

In Figure 2, we examine the effect of changing momentum traders' risk tolerance. (This experiment can equivalently be thought of as varying the relative proportions of momentum traders and newswatchers.) We set j = z = 12 months, and allow γ to vary. As risk tolerance increases, momentum traders respond more aggressively to past price changes--i.e., ϕ increases. This causes the impulse response function to reach higher peak values. Also, the unconditional volatility of monthly returns rises monotonically.[22] It turns out, however, that the effect of risk tolerance on the pricing error (P_t - P^*_t) is U-shaped: the pricing error first falls, and then rises, as risk tolerance is increased. On the one hand, more momentum trading accelerates the reaction of prices to information, which reduces underreaction and thereby decreases pricing errors. On the other hand, more momentum trading also exacerbates overreaction, which increases pricing errors. Evidently, the two effects interact so as to produce a non-monotonic pattern.[23]

Finally, in Figure 3, we allow the information-diffusion parameter z to vary. Increasing z has a monotonic effect on the intensity ϕ of momentum trade: the slower the newswatchers are to figure things out, the greater are the profit opportunities for momentum traders. In the range of the parameter space where $j \geq z-1$, the induced increase in ϕ in turn has a monotonic effect on the peak impulse response--more aggressive momentum trade leads to more pronounced overshooting, and correspondingly, to negative autocorrelations that are generally larger in absolute value during the reversal phase.[24]

III. Extensions of the Basic Model: More Rational Arbitrage

We now consider a few extensions of the basic model. The overall spirit here is to ask: what happens as we allow for progressively more rational behavior by arbitrageurs?

A Contrarian Strategies

A.1 Contrarians and Momentum Traders Are Two Separate Groups

We have emphasized repeatedly that our results are attributable to the assumption that momentum traders make "simple" forecasts--i.e., they can only run univariate regressions. But even if one accepts this restriction at face value, it begs the following question: why do all traders have to use the same single forecasting variable? Why not allow for some heterogeneity in trading styles, with different groups focusing on different predictive variables?

Given the existence of the newswatchers and the underreaction that they create, it is certainly natural to begin an examination of simple arbitrage strategies with the sort of momentum-trading style that we have considered thus far. However, once it is understood that the momentum traders must--if they are the only arbitrageurs active in the market--ultimately cause prices to overreact, we then ought to think about the effects of second-round "contrarian" strategies that might be designed to exploit this overreaction.

To incorporate such contrarian strategies into our model, we assume that there is a total risk tolerance of γ available to engage in arbitrage activity. We also continue to assume that all arbitrageurs have horizons of j periods. But there are now two arbitrage styles. A fraction w of the arbitrageurs are momentum traders, who use ΔP_{t-1} to forecast $(P_{t+j} - P_t)$. The remaining (1-w) are contrarians, who use ΔP_{t-1-c} to forecast $(P_{t+j} - P_t)$. If we choose the lag length c properly, the contrarians will in equilibrium put <u>negative</u> weight on ΔP_{t-1-c} in making these forecasts.

Suppose provisionally that one takes the fraction w as fixed. Then the equilibrium is a natural generalization of that seen above. In particular, prices will be given by:

$$P_t = D_t + \{(z-1)\epsilon_{t+1} + (z-2)\epsilon_{t+2} + ... + \epsilon_{t+z-1}\}/z + \Sigma^{j}_{i=1} (\phi^M \Delta P_{t-i} + \phi^C \Delta P_{t-c-i}) \qquad (9)$$

where ϕ^M and ϕ^C now denote the trading elasticities of the momentum traders and the contrarians respectively. These elasticities in turn satisfy:

$$\phi^M = w\gamma cov(P_{t+j} - P_t, \Delta P_{t-1})/\{var(\Delta P)var_M(P_{t+j} - P_t)\} \qquad (10)$$

$$\phi^C = (1-w)\gamma cov(P_{t+j} - P_t, \Delta P_{t-1-c})/\{var(\Delta P)var_C(P_{t+j} - P_t)\} \qquad (11)$$

Equilibrium now involves a two-dimensional fixed point in (ϕ^M, ϕ^C) such that prices are given by equation (9), while at the same time equations (10) and (11) are satisfied. Although this is a more complicated problem than before, it is still straightforward to solve numerically. Of course, this is no longer the end of the story, since we still need to endogenize w. This can be done by imposing an indifference condition: in an interior solution where $0 < w < 1$, the utilities of the momentum traders and contrarians must be equalized, so nobody wants to switch styles. It turns out that the equal-utility condition can be simply rewritten in terms of either conditional variances or covariances of prices. (See the appendix for a proof.) This gives us:

Proposition 3: In an interior solution with $0 < w < 1$, it must be that:

i) $\mathrm{var}(P_{t+j} - P_t \mid \Delta P_{t-1}) = \mathrm{var}(P_{t+j} - P_t \mid \Delta P_{t-1-c})$; or equivalently

ii) $|\mathrm{cov}((P_{t+j} - P_t), \Delta P_{t-1})| = |\mathrm{cov}((P_{t+j} - P_t), \Delta P_{t-1-c})|$; or equivalently

iii) $\mathrm{cov}(\Delta P_{t+1}, \Delta P_{t-1}) + \mathrm{cov}(\Delta P_{t+2}, \Delta P_{t-1}) +.....+\mathrm{cov}(\Delta P_{t+j}, \Delta P_{t-1}) =$

$-\mathrm{cov}(\Delta P_{t+1}, \Delta P_{t-1-c}) -\mathrm{cov}(\Delta P_{t+2}, \Delta P_{t-1-c}) -.....-\mathrm{cov}(\Delta P_{t+j}, \Delta P_{t-1-c})$.

The essence of the proposition is that in order for contrarians to be active in equilibrium (i.e., to have $w < 1$) there must be as much profit opportunity in the contrarian strategy as in the momentum strategy. Loosely speaking, this amounts to saying that the negative autocorrelations in the reversal phase must cumulatively be as large in absolute magnitude as the positive autocorrelations in the initial underreaction phase. Thus adding the option of a contrarian strategy to the model cannot overturn the basic result that if there is underreaction in the short run, there must eventually be overreaction at some later point.

As it turns out, for a large range of parameter values, we can make a much stronger statement: the contrarian strategy is not used at all, for any choice of c. Rather, we get a corner solution of $w = 1$, in which all arbitrageurs endogenously choose to use a momentum strategy. This is in fact the outcome for every set of parameters that appears in Figures 1-3. Thus our previous numerical solutions are wholly unaffected by adding contrarians to the mix.

In order to get contrarian strategies to be adopted in equilibrium, we have to crank up the aggregate risk tolerance γ to a very high value. This does two things: first, it drives down the expected profits to the momentum strategy; and second, it causes the degree of overreaction to increase. Both of these effects raise the relative appeal of being a contrarian to the point that some arbitrageurs eventually switch over from the momentum strategy. Figure 4 illustrates. The figure considers a situation where z = 3, j = 1, where the contrarians trade at a lag that is c = 2 periods greater than the momentum traders, and where the risk tolerance takes on the value 1/0.3.

Given these parameter values, w=0.786. That is, 78.6 percent of traders opt to play momentum strategies and the remaining 21.4 percent become contrarians. The contrarians appear to have a modest stabilizing impact--the impulse response function peaks at 1.197 when there are only momentum traders, and this figure declines somewhat, to 1.146, when we allow for contrarian strategies. Nevertheless, price dynamics are still remarkably similar to what we have seen throughout. This underscores our key point: Across a wide range of parameter values, allowing for contrarian strategies need not alter the important qualitative features of our model.

A.2. <u>Arbitrageurs Can Run Bivariate Regressions</u>

To further relax our assumptions in the direction of rationality, we now ask what happens if every arbitrageur becomes incrementally smarter, and can condition on not one, but two lags of past prices. Said differently, instead of forcing each arbitrageur to choose whether to play a momentum strategy (and condition on ΔP_{t-1}) or a contrarian strategy (and condition on ΔP_{t-1-c}), we now allow them all to play an optimal blended strategy.

The results of this experiment are also illustrated in Figure 4. Relative to the previous case of segregated momentum and contrarian trading, allowing for bivariate-regression-running arbitrageurs is more stabilizing. For example, keeping all other parameters the same as before, the impulse response function now reaches a peak of only 1.125, as compared to the value of 1.146 with segregated momentum and contrarian trading. Nevertheless, its qualitative shape continues to remain similar. Thus while increasing the computational power of arbitrageurs obviously attenuates the results, it does not appear that we are in a knife-edge situation where everything hangs on them being able to run only univariate regressions.

B Fully Rational Arbitrage

Finally, it is natural to ask whether our basic results are robust to the introduction of a class of fully rational arbitrageurs. To address this question, we extend the baseline model of Section II as follows. In addition to the newswatchers and the momentum traders, we add a third group of traders, whom we label the "smart money". To give these smart-money traders the best shot at making the market efficient, we consider an extreme case where they can observe and rationally condition on everything in the model that is observed by any other trader. Thus, at time t, the smart-money traders observe all the fundamental information that is available to any of the newswatchers--i.e., they see ϵ_{t+z-1} and all preceding news innovations in their entirety. They also can use the complete past history of prices in their forecasts. Like everyone else, the smart money have CARA utility. Finally, each generation has a one-period horizon.

Unlike in the cases with contrarian trading considered in Sections III.A.1 and III.A.2 above, it is very difficult to solve explicitly for the equilibrium with the smart-money traders, either analytically or via numerical methods. This is because in the context of our infinite-horizon model, the optimal forecasts of the smart money are a function of an unbounded set of variables, as they condition on the entire past history of prices. (They really have to be very smart in this model to implement fully rational behavior.) Nevertheless, as proven in the Appendix, we are able to make the following strong general statements about the properties of equilibrium:

Proposition 4: Assume that the risk tolerance of the smart-money traders is finite. In any covariance-stationary equilibrium, given a one-unit shock ϵ_{t+z-1} that begins to diffuse at time t:

i) there is always underreaction, in the sense that prices rise by less than one at time t;

ii) there is active momentum trading;

iii) there is always overreaction, in the sense that the cumulative impulse response of prices peaks at a value that is strictly greater than one.

If the risk tolerance of the smart money traders is infinite, prices follow a random walk, and there is no momentum trading: $\phi = 0$.

The proposition formalizes the intuitive point--common to many models in this genre--that risk-averse fully rational arbitrageurs attenuate, but do not eliminate, the effects induced by other less-than-rational traders. In our particular setting, all the key qualitative results about the dynamics of prices continue to apply.

IV. Empirical Implications

We will not belabor the fact that our model delivers the right first-order predictions for asset returns: positive correlations at short horizons, and negative correlations at longer horizons. After all, it is designed to do just that. More interesting are the auxiliary implications, which should allow it to be tested against other candidate theories of underreaction and overreaction.

A In What Stocks Do Momentum Strategies Work Best?

In our model, short-term return continuation is a consequence of the gradual diffusion of private information, combined with the failure of newswatchers to extract this information from prices. This gradual-information-diffusion story is logically distinct from the mechanism in other models, such as BSV, who emphasize a conservatism bias (Edwards (1968)) with respect to public information. Moreover, it has testable cross-sectional implications. If momentum in stock returns does indeed come from gradual information flow, then momentum strategies of the sort proposed by Jegadeesh and Titman (1993) should be most profitable among those stocks for which information moves most slowly across the investing public.

In research conducted subsequent to the development of the model here, we attempt in Hong, Lim and Stein (1998) to test this hypothesis. To do so, we consider two different proxies for the rate of information diffusion. The first is firm size. It seems plausible that information about small firms gets out more slowly; this would happen if, e.g., investors face fixed costs of information acquisition, and choose to devote more effort to learning about those stocks in which they can take large positions. Of course, one must be careful in drawing inferences because size may also capture a variety of other factors, such as cross-stock differences in arbitrage costs.[25] In light of this concern, we use as a second--and hopefully purer--proxy for information flow a stock's residual analyst coverage, after controlling for size.[26]

With respect to size, we find that, once one moves past the very smallest-capitalization stocks (where price discreteness and/or very thin market-making capacity are issues) the profitability of Jegadeesh-Titman-style six-month momentum strategies declines sharply with market cap. With respect to residual analyst coverage, not only are momentum strategies substantially more profitable at a horizon of six months in low-analyst-coverage stocks, they are also profitable for <u>longer</u>--there is pronounced positive correlation of returns for up to about two years in these stocks, as opposed to less than one year in high-coverage stocks. Size and residual coverage also interact in an interesting and economically plausible fashion: the marginal impact of analyst coverage is most pronounced in smaller stocks, which have fewer analysts to begin with. While it may be possible to come up with alternative interpretations, all these pieces of evidence would seem to be strongly consistent with our emphasis in this paper on gradual information flow as the root cause of underreaction.

B Linking Momentum to Overreaction in the Cross Section

There is also a second, more subtle cross-sectional implication of our model pertaining to the rate of information flow. In Figure 3 we saw that not only does slower information diffusion lead to higher short-run return correlations, but by making stocks more attractive to momentum traders, it also (for a wide range of parameter values) leads to more pronounced overshooting and stronger reversals in the longer run. In other words, the same stocks that we find in Hong, Lim and Stein (1998) to be most "momentum-prone"--small stocks with relatively few analysts--should also be most "reversal-prone".

Although this prediction has not to our knowledge been subject to detailed investigation, it is broadly consistent with recent work which finds that much of the long-horizon predictability that has been documented in the stock market is attributable to smaller-cap companies.[27] As noted above, there is the caveat that size may be proxying for a number of other factors, so as in Hong, Lim and Stein, it would be desirable to create a sharper test, perhaps using analyst coverage or some other non-size measure of momentum-proneness.

C Differential Dynamics in Response to Public vs. Private News Shocks?

As we have stressed repeatedly, the most natural interpretation of the ϵ's in our model is that they represent information that is initially private, and that gradually diffuses across the population of investors. Thus our primary contribution is to show that the equilibrium impulse response to such private information must be hump-shaped, with underreaction in the short run giving way to eventual overreaction. But what about the impulse response to news that is simultaneously observed by all investors, such as earnings announcements?

It is easy to embellish our model so that it also generates short-run underreaction to public news. For example, one might argue that although the news announcement itself (e.g. "earnings are up by 10 percent") is public, it requires some other, private, information (e.g., knowledge of the stochastic process governing earnings) to convert this news into a judgement about value. If this is true, the market's response to public news involves the aggregation of private signals, and our previous underreaction results continue to apply.

On the one hand, this sort of patch adds an element of descriptive realism, given the large body of empirical evidence on post-event drift. But the more interesting and subtle question is this: If we augment the model so as to deliver short-run underreaction to public news, what does it have to say about whether there is overreaction in the longer run to this same news? Is the impulse response function hump-shaped as before, or do prices drift gradually to the correct level without going too far?

Unlike with private news, the answer is now less clear. This is because the inference problem for momentum traders is simplified. Recall from above that with private news, a momentum trader never knows whether he is buying early or late in the cycle--i.e. he cannot tell if a price increase is the result of recent news or of past rounds of momentum trade. But if momentum traders can condition on the fact that there was a public news announcement at some given date t, they can refine their strategies. In particular, they can make their strategies time-dependent, so that they only trend-chase aggressively in the periods right after public news, and lay low at other times. If they do this, there need be no overreaction to public news in equilibrium; rather, the impulse response function may be increasing everywhere.

Of course, it is conceivable that momentum traders are not so sophisticated, and continue to use strategies that do not depend on how recently public news was released. If so, the impulse response to public news is also hump-shaped. But the important point is that the logic of our model admits (even strongly suggests) the possibility that the response to public news looks different than that to private information. This is clearly a testable proposition.

D Trading Horizons and the Pattern of Return Autocorrelations

One novel feature of our model is that it explicitly links momentum traders' horizons to the time pattern of return autocorrelations. This link is loosely suggested by Proposition 2, and it emerges clearly in the comparative statics results of Figure 1: the longer the momentum traders' horizon j, the longer it takes for the autocorrelations to switch from positive to negative.

The first thing to note in this regard is that our model seems to get the average magnitudes about right. For example, Jegadeesh and Titman (1993) find that autocorrelations for stock portfolios are positive for roughly 12 months, and then turn systematically negative. According to our calculations (see the appendix), this is what one should expect if j is on the order of 12-18 months, which sounds like a plausible value for the horizon of a trading strategy.[28]

A second observation is that we can make cross-sectional predictions, to the extent that we can identify exogenous factors that influence the trading horizon j. One natural candidate for such a factor is trading costs. It seems plausible to conjecture that as trading costs increase, momentum traders choose to hold their positions for longer. If so, we would expect stocks with relatively high bid-ask spreads to have autocorrelations that stay positive for longer periods of time before turning negative. Or going across assets classes, we would expect the same thing for assets such as houses, and collectibles, where trading costs are no doubt significantly higher.[29] Some evidence on this latter point is provided by Cutler, Poterba and Summers (1991). They find that, in contrast to equities, the autocorrelations for house and farm prices are positive at lags of up to 3 years, and for collectibles, at lags of up to 2 years.

E Anecdotal Evidence on Professional Investment Strategies

In our model, momentum traders have two key characteristics: 1) aside from their inability to run multiple regressions, they are rational maximizers who make money on average; and 2) they impose a negative externality on others. The latter feature arises because someone entering the market at any time t does not know how heavily invested momentum traders are in the aggregate at this time, and hence cannot predict whether or not there will be large-scale unwinding of momentum positions in the near future.

Anecdotal evidence supports both of these premises. With regard to the near-rationality of momentum strategies, it should be noted that a number of large and presumably sophisticated money managers use what are commonly described as momentum approaches, that "emphasize accelerating sales, earnings, or even stock prices...and focus less on traditional valuation measures such as price-to-earnings ratios..."[30] This contrasts with the more pejorative view of positive-feedback trading that prevails in prior academic work such as Delong et al (1990).

With regard to the negative externalities, it seems that other professional investors do in fact worry a lot about the dangers of momentum traders unwinding their positions. The following quotes from money managers illustrate these concerns: "Before I look at a stock, I take a look at the (SEC) filings to see who the major shareholders are. If you see a large amount of momentum money in there, you have to accept that there's a high risk..."; "If you're in with managers who are very momentum oriented....you have to be aware that's a risk going in. They come barreling out of those stocks, and they're not patient about it."[31]

In addition to these two premises, anecdotal evidence is also consistent with one of our key predictions: That momentum traders are more active in small stocks, where analyst coverage is thinner and information diffuses more slowly. According to a leading pension fund consultant: "Most of the momentum players play in the small and mid-cap stocks." And a well-known momentum investor says that he typically focuses on small companies because "the market is inefficient for smaller companies."[32]

More broadly, the extended version of the model with contrarians fits with the observation that there are a variety of professional money-management "styles", each of which emphasizes a different subset of public information. Such heterogeneity cannot be understood in the context of the standard rational model, where there is only one "correct" style, that which processes <u>all</u> available information in an optimal fashion. But it is a natural feature of our bounded-rationality framework, which allows multiple styles to coexist and earn similar profits.

V. Comparison to Related Work

As noted in the Introduction, this paper shares the same goal as recent work by BSV (1997) and DHS (1997)--i.e., to construct a plausible model that delivers a unified account of asset-price continuations and reversals. However, the approach taken here is quite different. Both BSV and DHS use representative agent models, while our results are driven by the externalities that arise when heterogeneous traders interact with one another.[33] Consequently, many of the auxiliary empirical implications of our model are distinct.

First, it is impossible for a representative agent model to make predictions linking trading horizons to the temporal pattern of autocorrelations, as we do in Section IV.D. Second, neither the BSV nor the DHS model would seem to be able to easily generate our prediction that both continuations and reversals are more pronounced in stocks with thinner analyst coverage (Sections IV.A and IV.B). A further difference with BSV is that our model allows for a differential impulse response to public and private shocks (Section IV.C), while theirs only considers public news.

In its focus on the interaction of different types of traders--including those who behave in a trend-chasing fashion--this paper is closer to earlier models of positive-feedback trading by DeLong et al (1990) and Cutler, Poterba and Summers (1990). However, there are significant differences with this work as well. For example, in DeLong et al, the positive-feedback traders are extremely irrational, and get badly exploited by a group of rational frontrunners.[34] In our model, the momentum traders are very nearly rational, and actually manage to take advantage of the other group of traders, the newswatchers. This distinction is closely related to the fact that in DeLong et al, there is never any underreaction. There is short-run positive correlation of returns, but this reflects an initial overreaction, followed by <u>even more</u> overreaction.[35]

At a more general level, this paper revisits several themes that have been prominent in previous theoretical work. The notion that one group of optimizing traders might create a negative informational externality, and thereby destabilize prices even while they are making profits, also shows up in Stein (1987). Stretching a bit further, there is an interesting analogy here with the ideas of Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992) on informational cascades. In these models, agents move sequentially. In equilibrium, each rationally bases his decision on the actions of the agent before him, even though this inflicts a negative informational externality on those that follow. Very much the same thing could be said of the generations of momentum traders in this model.

VI. Conclusions

At the outset, we argued that any new "behavioral" theory of asset pricing should be judged according to three criteria: 1) It should rest on assumptions about investor behavior that are either a priori plausible or consistent with casual observation; 2) It should explain the existing evidence in a parsimonious and unified way; and 3) It should make a number of further predictions which can be subject to testing and which are ultimately validated.

How well have we done on these three scores? With respect to the first, we believe that our particular rendition of bounded rationality--as the ability to process a small subset of the available information in an unbiased way--is both plausible and intuitively appealing. Moreover, in our framework, this sort of bounded rationality implies a widespread reliance by arbitrageurs on simple momentum strategies. As we have discussed, this implication appears to be strongly consistent with what is observed in the real world.

In terms of the parsimony/unity criterion, it should be emphasized that everything in our model is driven by just one primitive type of shock: slowly-diffusing news about future fundamentals. There are no other exogenous sources of investor sentiment, and no liquidity disturbances. Our main conceptual contribution is to show that if there is ever any short-run underreaction to this kind of news on the part of one set of traders, then (given the simple nature of arbitrage strategies) there must eventually be overreaction in the longer run as well.

Finally, our model does deliver several testable auxiliary implications. Among the most noteworthy are the following: 1) Both short-run continuation and long-run reversals should be more pronounced in those (small, low-analyst-coverage) stocks where information diffuses more slowly; 2) There may be more long-run overreaction to information which is initially private than to public news announcements; and 3) There should be a relationship between momentum traders' horizons and the pattern of return autocorrelations. Evidence supportive of the first prediction is already emerging; we hope to explore some of the others in future work.

Appendix

A. Proofs

A.1 ARMA Representation of the Return Process

Let us begin by recalling Equation (5) from the text (suppressing constants):

$$P_t = D_t + \frac{(z-1)}{z}\epsilon_{t+1} + \ldots. + \frac{1}{z}\epsilon_{t+z-1} + \sum_{i=1}^{j} \phi \Delta P_{t-i} \qquad (A.1)$$

It follows that

$$\Delta P_t = \frac{\sum_{i=0}^{z-1} \epsilon_{t+i}}{z} + \phi \Delta P_{t-1} - \phi \Delta P_{t-(j+1)}. \qquad (A.2)$$

Assuming that ϕ satisfies proper conditions to be specified, ΔP_t is a covariance stationary process. Let

$$\alpha_k \equiv E[\Delta P_t \Delta P_{t-k}]$$

(i.e. the unconditional autocovariance lagged k periods). When k=0, we have the unconditional variance. The autocovariances of this process satisfy the following Yule-Walker equations.

$$\alpha_0 = E[\sum_{i=0}^{z-1} \frac{\epsilon_{t+i}}{z} \Delta P_t] + \phi \alpha_1 - \phi \alpha_{j+1}. \qquad (A.3)$$

And for k > 0, we have

$$\alpha_k = E[\sum_{i=0}^{z-1} \frac{\epsilon_{t+i}}{z} \Delta P_{t-k}] + \phi\alpha_{k-1} - \phi\alpha_{k-(j+1)}. \qquad (A.4)$$

It is not hard to verify that for k > z-1,

$$E[\sum_{i=1}^{z-1} \frac{\epsilon_{t+i}}{z} \Delta P_{t-k}] = 0. \qquad (A.5)$$

And for k≤ z-1, we have

$$E[\sum_{i=0}^{z-1} \frac{\epsilon_{t+i}}{z} \Delta P_{t-k}] = \frac{(z-k)\sigma^2}{z^2} + \phi E[\sum_{i=0}^{z-1} \frac{\epsilon_{t+i}}{z} \Delta P_{t-(k+1)}] - \phi E[\sum_{i=0}^{z-1} \frac{\epsilon_{t+i}}{z} \Delta P_{t-(k+j+1)}] \qquad (A.6)$$

where σ is the standard deviation of the ϵ's. Solving the Yule-Walker equations reduces to solving a system of j+2 linear equations. Next, the optimal strategies of the momentum traders are given by

$$\zeta_t^M = \frac{\gamma E[P_{t+j} - P_t | \Delta P_{t-1}]}{Var[P_{t+j} - P_t | \Delta P_{t-1}]}, \qquad (A.7)$$

where

$$P_{t+j} - P_t = \Delta P_{t+j} + ... + \Delta P_{t+1}.$$

In equilibrium,

$$\zeta_t^M = \phi \Delta P_{t-1}. \qquad (A.8)$$

Finally, it follows that

$$Cov(\Delta P_{t-1}, P_{t+j} - P_t) = \alpha_{j+1} + \ldots.. + \alpha_2,$$

and

$$Var(P_{t+j} - P_t) = j\alpha_0 + 2(j-1)\alpha_1 + \ldots. + 2(j-(j-1))\alpha_{j-1}.$$

Using these formulas, the problem is reduced to finding a fixed point in ϕ that satisfies the equilibrium condition (A.8). Given the equilibrium ϕ, we then need to verify that the resulting equilibrium ARMA process is in fact covariance stationary (since all of our formulas depend crucially on this assumption).

A.2 Stationarity

We next provide a characterization for the covariance stationarity of a conjectured return process. This condition is just that the roots of

l i e

outsi

$$1 - \phi x + \phi x^{j+1} = 0 \qquad (A.9)$$

d e

the unit circle (see, e.g., Hamilton (1994)).

Lemma A.1 The return process specified in equation (A.2) is a covariance stationary process only if | ϕ | < 1.

Proof. Proof is by induction on j. For j=1, the return process follows an ARMA(2,z). So, the conditions for covariance stationarity are: -2 ϕ < 1; and -1 < ϕ < 1, (see, e.g. Hamilton (1994)). The stated result follows for j=1. Apply the inductive hypothesis and assume the result holds for j=k. From equation (A.9), it follows that the roots x of

$$1-\phi x+\phi x^{k+1}=0$$

must lie outside the unit circle (e.g. | x | > 1). It follows that

$$|1-\phi x|=|\phi||x|^{k+1}. \qquad (A.10)$$

Hence, as k increases, it follows that | ϕ | decreases for equation (A.10) to hold. The stated result follows for arbitrary j. QED

We use this result to characterize a number of properties of a conjectured covariance stationary equilibrium.

Proof of Lemma 1. We show that ϕ > 0 in a covariance stationary equilibrium by contradiction. Suppose it is not, so that $\phi \leq 0$. It is easy to verify from equation (A.4) and from Lemma A.1 that

$$\alpha_k \geq 0 \qquad \forall k \qquad \rightarrow \alpha_2+\alpha_3+....+\alpha_{j+1}>0$$

implying that ϕ > 0, leading to a contradiction. QED

A.3 Existence and Numerical Computation

An equilibrium ϕ satisfying the convariance stationary condition in Lemma A.1 does not exist for arbitrary parameter values. It is easy to verify however that a covariance stationary equilibrium does exist for sufficiently small γ.

Lemma A.2 For γ sufficiently small, there exists a covariance stationary equilibrium.

Proof. It is easy to show that for γ sufficiently small, we can apply Brouwer's fixed point theorem. QED

In general, the equilibrium needs to be solved numerically. For the case of j=1, we can always verify the resulting ϕ leads to covariance stationarity. For arbitrary j, we only have a necessary condition although the calculations for the autocovariances would likely explode for a ϕ which does not lead to a covariance stationary process. So, we always begin our calculations for j =1 and γ small and use the resulting solutions to bootstrap our way to other regions in the parameter space. The solutions are gotten easily. When we move outside the covariance stationary region of the parameter space, autocovariances take on non-sensible values such as negative values for the unconditional variance or autocovariances that do not satisfy the standard property that

$$|\alpha_k| < |\alpha_0|, \qquad k>0$$

in a covariance stationary equilibrium. In general, we have not had much problem finding fixed points for wide parameter regions around those exhibited in the text.

A.4 Proof of Proposition 3

The equilibrium condition to determine w is for the utilities from the two strategies to be equal. Given our assumptions on the preferences of the momentum and contrarian investors and the distributions of the ϵ's, it follows from Grossman and Stiglitz (1980) that this is equivalent to the conditional variance of the j-period returns being equal across the two strategies. Given that both momentum and contrarian investors have the same j-period horizon, it follows that this is equivalent to the conditional covariance of the j-period returns being equal across the two strategies. QED

A.5 Proof of Proposition 4

Suppose initially that there are only newswatchers and smart money investors (i.e. there are no momentum investors). Smart money investors have finite risk tolerance given by γ^S and maximize one-period returns. We conjecture the following equilibrium price function:

$$P_t = D_t + \frac{(z-1)}{z}\epsilon_{t+1} + ... + \frac{1}{z}\epsilon_{t+z-1} + \sum_{i=1}^{z-1} \beta_i \epsilon_{t+i}. \qquad (A.11)$$

Note that we are once again suppressing all calculations related to the constant. The holdings of the smart money investors are given by

$$\zeta_t^S = \frac{\gamma^S E[P_{t+1} - P_t | D_t, \epsilon_{t+1}, ..., \epsilon_{t+z-1}]}{Var[P_{t+1} - P_t | D_t, \epsilon_{t+1}, ..., \epsilon_{t+z-1}]}. \qquad (A.12)$$

At the conjectured equilibrium price given in equation (A.11), we have that

$$\zeta_t^S = \sum_{i=t+1}^{t+z-1} \beta_i \epsilon_i. \qquad (A.13)$$

Equation (A.13) then gives the following set of equations which determine the β's in equilibrium:

$$\beta_1 = \gamma^S \frac{\frac{1}{z} - \beta_1}{(\frac{1}{z} + \beta_{z-1})^2 \sigma^2} \qquad (A.14)$$

and

$$\beta_i = \gamma^S \frac{\frac{1}{z} + (\beta_{i-1} - \beta_i)}{(\frac{1}{z} + \beta_{z-1})^2 \sigma^2}, \qquad i = 2, \ldots, z-1. \qquad (A.15)$$

Using equations (A.14) and (A.15), it is not hard to show that in a covariance stationary equilibrium: (1) The returns still exhibit positive serial correlation for finite levels of smart money risk tolerance, $\gamma^S < \infty$, and (2) When smart money investors are risk neutral, prices follow a random walk.

Since returns are serially correlated when the risk tolerance of smart money is finite, $\phi = 0$ cannot be an equilibrium when we add momentum traders to the model. Since smart money investors have access to the entire history of past price changes, it follows from the logic of equation (A.11) that the conjectured price function with momentum traders is now

$$P_t = D_t + \frac{(z-1)}{z} \epsilon_{t+1} + \ldots. + \frac{1}{z} \epsilon_{t+z-1} + \sum_{i=1}^{z-1} \beta_i \epsilon_{t+i} + \sum_{i=1}^{\infty} \kappa_i \Delta P_{t-i} + \sum_{i=1}^{j} \phi \Delta P_{t-i}. \qquad (A.16)$$

Assuming that a covariance stationary equilibrium exists, the holding of the smart money is

$$\zeta_t^S = \frac{\gamma^S E[P_{t+1} - P_t | D_t, \epsilon_{t+1}, \ldots, \epsilon_{t+z-1}, \Delta P_{t-1}, \Delta P_{t-2}, \ldots, \Delta P_{-\infty}]}{Var[P_{t+1} - P_t | D_t, \epsilon_{t+1}, \ldots, \epsilon_{t+z-1}, \Delta P_{t-1}, \Delta P_{t-2}, \ldots, \Delta P_{-\infty}]}. \qquad (A.17)$$

while the holding of the momentum investors is given by equation (A.7). At the conjectured equilibrium price in equation (A.16), we have

$$\zeta_t^S = \sum_{i=t+1}^{t+z-1} \beta_i \epsilon_i + \sum_{i=1}^{\infty} \kappa_i \Delta P_{t-i} \qquad (A.18)$$

for the smart money and equation (A.8) for the momentum investors.

In general, the β's, κ's, and ϕ's in equation (A.16) have to be determined numerically as fixed points of equations (A.8) and (A.18) using the same methodology as described above in Section A.3. While solving for these parameters is computationally difficult, we can characterize certain behavior in a covariance stationary equilibrium. Given a positive one-unit shock that begins to diffuse among newswatchers at time t, the price underreacts at t for finite smart money risk tolerance, i.e.

$$\Delta P_t = \frac{1}{z} + \beta_{z-1} < 1.$$

The price eventually converges to one in a covariance stationary equilibrium. And in a covariance stationary equilibrium, the price must also overshoot one. To see this, suppose it does not. Then the serial correlation in returns would be positive at all horizons. Then this implies that momentum investors would have $\phi > 0$, which by our previous logic implies that there would be overreaction, thereby establishing a contradiction.

When the risk tolerance of smart money is infinite, it follows from the discussion above that without momentum traders prices follow a random walk. So, the expected return to momentum trading is zero. Hence, when the risk tolerance of smart money is infinite, prices following a random walk and no momentum trading is in fact a covariance-stationary equilibrium. QED

B. Numerical Comparative Statistics

[Insert Tables A1-A4 here]

Table A1
Comparative Statics With Respect To Momentum Traders' Horizon

Momentum traders' horizon j takes on values 3, 6, 9, 12, 15, and 18. ϕ is the intensity of momentum trade described in Equation (7). P_t is the stock price at time t. P_t^* is the rational expectations stock price in Equation (2). The other parameter values are set as follows: the information diffusion parameter z=12, the volatility of news shocks σ=0.5, and the risk tolerance γ=1/3.

Momentum Traders' Horizon		j=3	j=6	j=9	j=12	j=15	j=18
ϕ		0.5550	0.4455	0.3262	0.2605	0.2263	0.2015
Standard Deviation of ($P_t - P_{t-1}$)		0.2229	0.2322	0.2179	0.2028	0.1908	0.1833
Standard Deviation of ($P_t - P_t^*$) at $\phi=0$		0.9373	0.9373	0.9373	0.9373	0.9373	0.9373
Standard Deviation of ($P_t - P_t^*$)		0.8011	0.8438	0.9103	0.9365	0.9524	0.9604
Cumulative Impulse Response at Lag	0	0.0833	0.0833	0.0833	0.0833	0.0833	0.0833
	1	0.2129	0.2038	0.1939	0.1884	0.1855	0.1835
	2	0.3682	0.3408	0.3132	0.2991	0.2920	0.2870
	3	0.5377	0.4851	0.4355	0.4113	0.3994	0.3912
	4	0.6688	0.6328	0.5588	0.5238	0.5071	0.4955
	5	0.7530	0.7819	0.6823	0.6365	0.6147	0.5999
	6	0.7969	0.9317	0.8059	0.7492	0.7225	0.7042
	7	0.8105	1.0446	0.9296	0.8618	0.8302	0.8086
	8	0.8287	1.1246	1.0533	0.9745	0.9379	0.9129
	9	0.8753	1.1825	1.1769	1.0872	1.0456	1.0173
	10	0.9602	1.2273	1.2734	1.1999	1.1533	1.1217
	11	1.0830	1.2649	1.3522	1.3126	1.2610	1.2260
	12	1.1412	1.2151	1.3389	1.3420	1.2854	1.2471
	13	1.1475	1.1263	1.2947	1.3279	1.2909	1.2513
	14	1.1040	1.0364	1.2401	1.2969	1.2921	1.2522
	15	1.0116	0.9607	1.1820	1.2600	1.2924	1.2523
	16	0.9281	0.9012	1.1227	1.2211	1.2736	1.2524
	17	0.8782	0.8547	1.0630	1.1817	1.2462	1.2524
	18	0.8747	0.8173	1.0032	1.1420	1.2160	1.2524
	19	0.9240	0.8228	0.9433	1.1024	1.1848	1.2356
	20	0.9977	0.8648	0.8923	1.0627	1.1534	1.2120
	21	1.0663	0.9235	0.8500	1.0230	1.1219	1.1864
	22	1.1063	0.9834	0.8405	0.9833	1.0904	1.1603
	23	1.1012	1.0366	0.8518	0.9436	1.0589	1.1340
	24	1.0574	1.0810	0.8733	0.9038	1.0274	1.1076
	25	0.9951	1.1175	0.8993	0.8859	0.9959	1.0813
	26	0.9382	1.1313	0.9271	0.8848	0.9644	1.0550
	27	0.9096	1.1187	0.9557	0.8926	0.9329	1.0286
	28	0.9180	1.0870	0.9845	0.9043	0.9202	1.0023
	29	0.9572	1.0461	1.0134	0.9175	0.9161	0.9759
	30	1.0105	1.0042	1.0395	0.9312	0.9149	0.9496
	31	1.0560	0.9658	1.0618	0.9451	0.9146	0.9401
	32	1.0766	0.9324	1.0722	0.9590	0.9187	0.9373
	33	1.0663	0.9114	1.0719	0.9730	0.9259	0.9365
	34	1.0309	0.9076	1.0648	0.9870	0.9344	0.9364
	35	0.9861	0.9201	1.0540	1.0010	0.9433	0.9363
Return Autocorrelations at Lag	1	0.8630	0.9331	0.9458	0.9466	0.9465	0.9436
	2	0.5441	0.7888	0.8415	0.8504	0.8673	0.8634
	3	0.1706	0.6063	0.7048	0.7388	0.7813	0.7782
	4	(0.1203)	0.4039	0.5561	0.6220	0.6907	0.6920
	5	(0.2115)	0.1888	0.4031	0.5036	0.5870	0.6054
	6	(0.1035)	(0.0293)	0.2492	0.3846	0.4774	0.5182
	7	0.1284	(0.2138)	0.0963	0.2654	0.3658	0.4285
	8	0.3635	(0.3233)	(0.0522)	0.1464	0.2537	0.3270
	9	0.4735	(0.3637)	(0.1883)	0.0276	0.1415	0.2208
	10	0.4096	(0.3533)	(0.3027)	(0.0898)	0.0300	0.1137
	11	0.1910	(0.3051)	(0.3708)	(0.2027)	(0.0784)	0.0087
	12	(0.0958)	(0.2200)	(0.3955)	(0.2994)	(0.1741)	(0.0846)
	13	(0.3160)	(0.0850)	(0.3589)	(0.3385)	(0.2162)	(0.1215)
	14	(0.4026)	0.0574	(0.2985)	(0.3348)	(0.2452)	(0.1465)
	15	(0.3294)	0.1696	(0.2289)	(0.3088)	(0.2697)	(0.1690)
	16	(0.1297)	0.2376	(0.1560)	(0.2729)	(0.2873)	(0.1909)
	17	0.1034	0.2632	(0.0823)	(0.2332)	(0.2792)	(0.2125)
	18	0.2808	0.2532	(0.0098)	(0.1919)	(0.2595)	(0.2330)
	19	0.3387	0.2108	0.0582	(0.1502)	(0.2355)	(0.2485)
	20	0.2599	0.1318	0.1177	(0.1083)	(0.2096)	(0.2402)
	21	0.0869	0.0331	0.1594	(0.0663)	(0.1803)	(0.2224)
	22	(0.1076)	(0.0608)	0.1810	(0.0245)	(0.1488)	(0.2016)
	23	(0.2477)	(0.1329)	0.1762	0.0170	(0.1165)	(0.1801)
	24	(0.2817)	(0.1765)	0.1549	0.0573	(0.0838)	(0.1583)
	25	(0.2046)	(0.1914)	0.1252	0.0929	(0.0510)	(0.1363)

Table A2

Comparative Statics With Respect To Momentum Traders' Risk Tolerance

Momentum traders' risk tolerance γ takes on values 1/13, 1/11, 1/9, 1/7, 1/5 and 1/3. ϕ is the intensity of momentum trade described in Equation (7). P_t is the stock price at time t. P_t^* is the rational expectations stock price described in Equation (2). The other parameter values are set as follows: the information diffusion parameter z=12, the momentum traders' horizon j=12, and the volatility of news shocks σ=0.5.

Momentum Traders' Risk Tolerance		γ=1/13	γ=1/11	γ=1/9	γ=1/7	γ=1/5	γ=1/3
ϕ		0.1316	0.1453	0.1625	0.1848	0.2152	0.2605
Standard Deviation of (P_t - P_{t-1})		0.1662	0.1691	0.1731	0.1786	0.1872	0.2028
Standard Deviation of (P_t - P_t^*) at ϕ=0		0.9373	0.9373	0.9373	0.9373	0.9373	0.9373
Standard Deviation of (P_t - P_t^*)		0.9070	0.8998	0.8999	0.9021	0.9102	0.9365
Cumulative Impulse Response at Lag	0	0.0833	0.0833	0.0833	0.0833	0.0833	0.0833
	1	0.1776	0.1788	0.1802	0.1821	0.1846	0.1884
	2	0.2734	0.2760	0.2793	0.2836	0.2897	0.2991
	3	0.3693	0.3734	0.3787	0.3857	0.3957	0.4113
	4	0.4652	0.4709	0.4782	0.4879	0.5018	0.5238
	5	0.5612	0.5684	0.5777	0.5902	0.6080	0.6365
	6	0.6572	0.6659	0.6772	0.6924	0.7142	0.7492
	7	0.7531	0.7634	0.7767	0.7946	0.8203	0.8618
	8	0.8491	0.8609	0.8762	0.8968	0.9265	0.9745
	9	0.9450	0.9584	0.9757	0.9990	1.0327	1.0872
	10	1.0410	1.0559	1.0752	1.1013	1.1389	1.1999
	11	1.1369	1.1534	1.1747	1.2035	1.2451	1.3126
	12	1.1496	1.1676	1.1909	1.2224	1.2679	1.3420
	13	1.1403	1.1575	1.1800	1.2105	1.2549	1.3279
	14	1.1266	1.1422	1.1625	1.1900	1.2303	1.2969
	15	1.1122	1.1259	1.1435	1.1675	1.2024	1.2600
	16	1.0977	1.1093	1.1243	1.1444	1.1736	1.2211
	17	1.0832	1.0928	1.1050	1.1213	1.1446	1.1817
	18	1.0687	1.0762	1.0857	1.0981	1.1155	1.1420
	19	1.0541	1.0596	1.0664	1.0750	1.0864	1.1024
	20	1.0396	1.0430	1.0471	1.0518	1.0572	1.0627
	21	1.0251	1.0265	1.0278	1.0286	1.0281	1.0230
	22	1.0105	1.0099	1.0085	1.0055	0.9990	0.9833
	23	0.9960	0.9933	0.9892	0.9823	0.9699	0.9436
	24	0.9815	0.9767	0.9698	0.9591	0.9408	0.9038
	25	0.9779	0.9723	0.9641	0.9514	0.9296	0.8859
	26	0.9786	0.9731	0.9649	0.9521	0.9300	0.8848
	27	0.9805	0.9754	0.9679	0.9560	0.9354	0.8926
	28	0.9827	0.9781	0.9715	0.9609	0.9425	0.9043
	29	0.9849	0.9809	0.9752	0.9661	0.9503	0.9175
	30	0.9871	0.9838	0.9789	0.9713	0.9582	0.9312
	31	0.9893	0.9866	0.9827	0.9766	0.9662	0.9451
	32	0.9915	0.9894	0.9864	0.9818	0.9741	0.9590
	33	0.9937	0.9922	0.9901	0.9871	0.9821	0.9730
	34	0.9959	0.9950	0.9939	0.9923	0.9901	0.9870
	35	0.9981	0.9978	0.9976	0.9976	0.9981	1.0010
Return Autocorrelations at Lag	1	0.9337	0.9352	0.9370	0.9393	0.9423	0.9466
	2	0.8410	0.8419	0.8430	0.8445	0.8467	0.8504
	3	0.7437	0.7430	0.7422	0.7413	0.7401	0.7388
	4	0.6456	0.6432	0.6401	0.6360	0.6304	0.6220
	5	0.5474	0.5431	0.5376	0.5302	0.5198	0.5036
	6	0.4492	0.4430	0.4351	0.4244	0.4091	0.3846
	7	0.3510	0.3429	0.3326	0.3186	0.2983	0.2654
	8	0.2527	0.2429	0.2301	0.2127	0.1875	0.1464
	9	0.1545	0.1428	0.1276	0.1070	0.0769	0.0276
	10	0.0565	0.0430	0.0255	0.0017	(0.0330)	(0.0898)
	11	(0.0403)	(0.0554)	(0.0749)	(0.1013)	(0.1398)	(0.2027)
	12	(0.1281)	(0.1439)	(0.1644)	(0.1923)	(0.2329)	(0.2994)
	13	(0.1484)	(0.1662)	(0.1892)	(0.2203)	(0.2653)	(0.3385)
	14	(0.1424)	(0.1600)	(0.1830)	(0.2143)	(0.2599)	(0.3348)
	15	(0.1294)	(0.1456)	(0.1667)	(0.1956)	(0.2381)	(0.3088)
	16	(0.1149)	(0.1291)	(0.1477)	(0.1731)	(0.2105)	(0.2729)
	17	(0.1000)	(0.1122)	(0.1280)	(0.1495)	(0.1809)	(0.2332)
	18	(0.0852)	(0.0952)	(0.1082)	(0.1256)	(0.1508)	(0.1919)
	19	(0.0703)	(0.0782)	(0.0883)	(0.1016)	(0.1205)	(0.1502)
	20	(0.0554)	(0.0612)	(0.0684)	(0.0776)	(0.0901)	(0.1083)
	21	(0.0405)	(0.0442)	(0.0485)	(0.0537)	(0.0597)	(0.0663)
	22	(0.0257)	(0.0272)	(0.0286)	(0.0297)	(0.0294)	(0.0245)
	23	(0.0108)	(0.0102)	(0.0088)	(0.0058)	0.0008	0.0170
	24	0.0039	0.0066	0.0107	0.0177	0.0302	0.0573
	25	0.0174	0.0219	0.0285	0.0388	0.0566	0.0929

Table A3

Comparative Statics With Respect To Information Diffusion Parameter

Information diffusion parameter z takes on values 3, 6, 9, 12, 15, and 18. ϕ is the intensity of momentum trade described Equation (7). P_t is the stock price at time t. P_t^* is the rational expectations stock price. The other parameter values are set as follows: the momentum traders' horizon j=12, the volatility of news shocks σ=0.5, and the momentum traders' risk tolerance γ=1/3.

Information Diffusion Parameter		z=3	z=6	z=9	z=12	z=15	z=18
ϕ		0.0322	0.1293	0.2023	0.2605	0.3214	0.3785
Standard Deviation of ($P_t - P_{t-1}$)		0.2952	0.2317	0.2106	0.2028	0.1977	0.1907
Standard Deviation of ($P_t - P_t^*$) at ϕ=0		0.3727	0.6180	0.7935	0.9373	1.0620	1.1736
Standard Deviation of ($P_t - P_t^*$)		0.3744	0.6317	0.8087	0.9365	1.0331	1.0992
Cumulative Impulse Response at Lag	0	0.3333	0.1667	0.1111	0.0833	0.0667	0.0556
	1	0.6774	0.3549	0.2447	0.1884	0.1548	0.1321
	2	1.0218	0.5459	0.3828	0.2991	0.2497	0.2167
	3	1.0329	0.7373	0.5219	0.4113	0.3469	0.3042
	4	1.0333	0.9287	0.6611	0.5238	0.4448	0.3929
	5	1.0333	1.1201	0.8004	0.6365	0.5430	0.4821
	6	1.0333	1.1449	0.9397	0.7492	0.6412	0.5713
	7	1.0333	1.1481	1.0790	0.8618	0.7394	0.6607
	8	1.0333	1.1485	1.2183	0.9745	0.8376	0.7501
	9	1.0333	1.1486	1.2465	1.0872	0.9359	0.8395
	10	1.0333	1.1486	1.2522	1.1999	1.0341	0.9288
	11	1.0333	1.1486	1.2533	1.3126	1.1323	1.0182
	12	1.0333	1.1486	1.2536	1.3420	1.2306	1.1076
	13	1.0226	1.1270	1.2311	1.3279	1.3074	1.1760
	14	1.0111	1.0999	1.1996	1.2969	1.3704	1.2284
	15	0.9997	1.0717	1.1652	1.2600	1.3602	1.2718
	16	0.9989	1.0433	1.1302	1.2211	1.3256	1.3107
	17	0.9989	1.0148	1.0949	1.1817	1.2831	1.3474
	18	0.9989	0.9864	1.0596	1.1420	1.2378	1.3275
	19	0.9989	0.9795	1.0242	1.1024	1.1918	1.2862
	20	0.9989	0.9782	0.9889	1.0627	1.1454	1.2368
	21	0.9989	0.9780	0.9536	1.0230	1.0989	1.1842
	22	0.9989	0.9779	0.9407	0.9833	1.0524	1.1305
	23	0.9989	0.9779	0.9370	0.9436	1.0059	1.0763
	24	0.9989	0.9779	0.9360	0.9038	0.9594	1.0220
	25	0.9989	0.9779	0.9358	0.8859	0.9128	0.9676
	26	0.9992	0.9807	0.9402	0.8848	0.8732	0.9211
	27	0.9996	0.9846	0.9475	0.8926	0.8402	0.8837
	28	1.0000	0.9887	0.9560	0.9043	0.8329	0.8531
	29	1.0000	0.9929	0.9648	0.9175	0.8417	0.8268
	30	1.0000	0.9972	0.9737	0.9312	0.8581	0.8030
	31	1.0000	1.0014	0.9826	0.9451	0.8780	0.8015
	32	1.0000	1.0028	0.9916	0.9590	0.8992	0.8165
	33	1.0000	1.0032	1.0005	0.9730	0.9209	0.8409
	34	1.0000	1.0033	1.0095	0.9870	0.9428	0.8701
	35	1.0000	1.0033	1.0139	1.0010	0.9648	0.9014
Return Autocorrelations at Lag	1	0.6806	0.8661	0.9239	0.9466	0.9585	0.9666
	2	0.3410	0.6970	0.8155	0.8504	0.8794	0.9010
	3	0.0110	0.5233	0.6995	0.7388	0.7761	0.8178
	4	0.0004	0.3493	0.5787	0.6220	0.6640	0.7216
	5	0.0000	0.1775	0.4416	0.5036	0.5486	0.6116
	6	0.0000	0.0230	0.2981	0.3846	0.4321	0.4861
	7	0.0000	0.0000	0.1532	0.2654	0.3152	0.3547
	8	(0.0000)	(0.0230)	0.0112	0.1464	0.1983	0.2234
	9	(0.0000)	(0.0481)	(0.1148)	0.0276	0.0818	0.0979
	10	(0.0004)	(0.0739)	(0.1647)	(0.0898)	(0.0335)	(0.0202)
	11	(0.0110)	(0.0997)	(0.1983)	(0.2027)	(0.1455)	(0.1304)
	12	(0.0223)	(0.1249)	(0.2270)	(0.2994)	(0.2483)	(0.2311)
	13	(0.0330)	(0.1455)	(0.2482)	(0.3385)	(0.3352)	(0.3159)
	14	(0.0230)	(0.1309)	(0.2371)	(0.3348)	(0.3873)	(0.3692)
	15	(0.0117)	(0.1071)	(0.2130)	(0.3088)	(0.4071)	(0.3980)
	16	(0.0007)	(0.0815)	(0.1846)	(0.2729)	(0.3802)	(0.4097)
	17	(0.0000)	(0.0557)	(0.1544)	(0.2332)	(0.3356)	(0.4070)
	18	(0.0000)	(0.0302)	(0.1206)	(0.1919)	(0.2842)	(0.3855)
	19	(0.0000)	(0.0069)	(0.0847)	(0.1502)	(0.2302)	(0.3299)
	20	(0.0000)	(0.0009)	(0.0481)	(0.1083)	(0.1753)	(0.2591)
	21	0.0000	0.0029	(0.0120)	(0.0663)	(0.1201)	(0.1826)
	22	0.0000	0.0066	0.0208	(0.0245)	(0.0649)	(0.1062)
	23	0.0000	0.0104	0.0375	0.0170	(0.0101)	(0.0326)
	24	0.0004	0.0142	0.0477	0.0573	0.0435	0.0370
	25	0.0007	0.0180	0.0556	0.0929	0.0938	0.1015

Table A4

Comparative Static With Respect To Risk Tolerance With Contrarian Trading

The combined risk tolerance γ of momentum traders and contrarians takes on values 1/0.1, 1/0.3, and 1/0.5. w is the equilibrium fraction of momentum traders in the population. For each level of risk tolerance, the first column corresponds to the all-momentum equilibrium, (as in Section II) the second column to the equilibrium in which w is endogenously determined, (as in Section IIIA.1) and the third column to the equilibrium in which every trader can optimally condition on both the momentum and contrarian variables (as in Section IIIA.2). ϕ^M and ϕ^C are the intensity of momentum and contrarian trades respectively. The other parameter values are set as follows: the information diffusion parameter z=3, traders' horizon j=1, and the volatility of news shocks σ=1. The contrarians are assumed to trade based on returns from three periods ago: c=2.

Risk Tolerance			γ=1/0.1			γ=1/0.3			γ=1/0.5	
Equilibrium (w)		1	0.6703		1	0.786		1	0.9079	
ϕ^M		0.4668	0.3078	0.2954	0.4167	0.3291	0.2904	0.3794	0.3466	0.282
ϕ^C		0.0000	(0.1514)	(0.1716)	0.0000	(0.0896)	(0.1498)	0.0000	(0.0351)	(0.1333)
Cumulative Impulse Response at Lag	0	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
	1	0.8223	0.7693	0.7651	0.8056	0.7764	0.7635	0.7931	0.7822	0.7607
	2	1.2282	1.1342	1.1276	1.1968	1.1458	1.1249	1.1744	1.1556	1.1205
	3	1.1895	1.0619	1.0499	1.1630	1.0917	1.055	1.1447	1.1177	1.057
	4	0.9819	0.9117	0.9029	0.9859	0.9425	0.9153	0.9887	0.9711	0.9251
	5	0.9031	0.8985	0.8944	0.9262	0.9178	0.9053	0.9408	0.9361	0.9148
	6	0.9632	1.0069	1.0108	0.9751	0.9967	1.0076	0.9818	0.9892	1.0056
	7	1.0281	1.0561	1.0596	1.0204	1.0393	1.0506	1.0156	1.0236	1.0432
	8	1.0303	1.0171	1.0159	1.0189	1.0162	1.014	1.0128	1.0131	1.012
	9	1.0010	0.9716	0.9671	0.9994	0.9853	0.974	0.9990	0.9945	0.9791
	10	0.9864	0.9785	0.9772	0.9919	0.9860	0.9819	0.9947	0.9923	0.9857
	11	0.9931	1.0080	1.0105	0.9969	1.0023	1.0078	0.9984	0.9996	1.006
	12	1.0032	1.0160	1.0182	1.0021	1.0081	1.0135	1.0014	1.0032	1.0101
	13	1.0047	1.0014	1.0005	1.0022	1.0019	1.0005	1.0011	1.0013	1.0003
	14	1.0007	0.9911	0.9891	1.0000	0.9965	0.9924	0.9999	0.9991	0.9945
	15	0.9981	0.9956	0.9953	0.9991	0.9977	0.9968	0.9995	0.9991	0.9978
	16	0.9988	1.0036	1.0049	0.9996	1.0010	1.0032	0.9999	1.0001	1.0022
	17	1.0003	1.0040	1.0048	1.0002	1.0016	1.0031	1.0001	1.0004	1.002
	18	1.0007	0.9994	0.9989	1.0002	1.0001	0.9993	1.0001	1.0001	0.9995
	19	1.0002	0.9974	0.9966	1.0000	0.9992	0.9979	1.0000	0.9999	0.9987
	20	0.9998	0.9993	0.9993	0.9999	0.9997	0.9996	1.0000	0.9999	0.9998
	21	0.9998	1.0013	1.0018	1.0000	1.0003	1.0011	1.0000	1.0000	1.0006
	22	1.0000	1.0009	1.0011	1.0000	1.0003	1.0006	1.0000	1.0001	1.0003
	23	1.0001	0.9996	0.9993	1.0000	1.0000	0.9996	1.0000	1.0000	0.9998
Return Autocorrelations at Lag	1	0.6433	0.6123	0.6053	0.6530	0.6346	0.6161	0.6595	0.6524	0.6245
	2	0.0268	0.0221	0.0140	0.0657	0.0600	0.0398	0.0937	0.0910	0.0623
	3	(0.2878)	(0.2404)	(0.2424)	(0.2447)	(0.2218)	(0.2248)	(0.2147)	(0.2068)	(0.2086)
	4	(0.1468)	(0.0221)	(0.0080)	(0.1294)	(0.0600)	(0.0194)	(0.1170)	(0.0910)	(0.0263)
	5	0.0658	0.1565	0.1707	0.0481	0.1047	0.1460	0.0371	0.0599	0.1264
	6	0.0992	0.0947	0.0968	0.0740	0.0794	0.0877	0.0584	0.0627	0.0792
	7	0.0156	(0.0521)	(0.0621)	0.0108	(0.0228)	(0.0477)	0.0081	(0.0031)	(0.0376)
	8	(0.0390)	(0.0722)	(0.0776)	(0.0263)	(0.0484)	(0.0641)	(0.0191)	(0.0281)	(0.0533)
	9	(0.0255)	0.0032	0.0081	(0.0155)	(0.0062)	0.0040	(0.0103)	(0.0088)	0.0019
	10	0.0063	0.0454	0.0526	0.0045	0.0231	0.0400	0.0033	0.0090	0.0311
	11	0.0149	0.0161	0.0158	0.0083	0.0119	0.0129	0.0052	0.0070	0.0103
	12	0.0040	(0.0204)	(0.0256)	0.0016	(0.0075)	(0.0181)	0.0007	(0.0014)	(0.0132)
	13	(0.0051)	(0.0176)	(0.0199)	(0.0028)	(0.0090)	(0.0144)	(0.0017)	(0.0035)	(0.0105)
	14	(0.0042)	0.0053	0.0080	(0.0018)	0.0005	0.0051	(0.0009)	(0.0007)	0.0035
	15	0.0004	0.0126	0.0153	0.0004	0.0049	0.0103	0.0003	0.0013	0.0071
	16	0.0022	0.0018	0.0012	0.0009	0.0016	0.0010	0.0005	0.0008	0.0007
	17	0.0008	(0.0068)	(0.0090)	0.0002	(0.0019)	(0.0056)	0.0001	(0.0003)	(0.0037)
	18	(0.0006)	(0.0038)	(0.0043)	(0.0003)	(0.0015)	(0.0027)	(0.0002)	(0.0004)	(0.0017)
	19	(0.0007)	0.0026	0.0038	(0.0002)	0.0004	0.0023	(0.0001)	0.0000	0.0014
	20	0.0000	0.0032	0.0041	0.0000	0.0010	0.0024	0.0000	0.0002	0.0015
	21	0.0003	(0.0003)	(0.0007)	0.0001	0.0001	(0.0004)	0.0000	0.0001	(0.0003)
	22	0.0002	(0.0020)	(0.0028)	0.0000	(0.0004)	(0.0016)	0.0000	0.0000	(0.0009)
	23	(0.0001)	(0.0007)	(0.0007)	0.0000	(0.0002)	(0.0004)	0.0000	(0.0001)	(0.0002)
	24	(0.0001)	0.0010	0.0015	0.0000	0.0001	0.0008	0.0000	0.0000	0.0004
	25	0.0000	0.0008	0.0010	0.0000	0.0002	0.0005	0.0000	0.0000	0.0003

References

Banerjee, Abhijit, 1992, A simple model of herd behavior, *Quarterly Journal of Economics* 107, 797-817.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, University of Chicago working paper, forthcoming in *Journal of Financial Economics.*

Bernard, Victor L., 1992. Stock price reactions to earnings announcements, in R. Thaler ed.: *Advances in Behavioral Finance* (Russell Sage Foundation, New York).

Bernard, Victor L., and J. Thomas, 1989, Post-earnings announcement drift: Delayed price response or risk premium? *Journal of Accounting Research,* 27, 1-48.

Bernard, Victor L., and J. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305-340.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1992, A theory of fads, fashions, custom and cultural change as informational cascades, *Journal of Political Economy* 100, 992-1026.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905-939.

Campbell, John Y., and Robert Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661-676.

Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *Journal of Finance* 51, 1681-1713.

Chopra, Navin, Josef Lakonishok, and Jay R. Ritter, 1992, Measuring abnormal performance: Do stocks overreact? *Journal of Financial Economics* 31, 235-268.

Cutler, David M., James M. Poterba, and Lawrence H. Summers, 1990, Speculative dynamics and the role of feedback traders, *American Economic Review Papers and Proceedings* 80, 63-68.

Cutler, David M., James M. Poterba, and Lawrence H. Summers, 1990, Speculative dynamics, *Review of Economic Studies* 58, 529-546.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 1998, A Theory of overconfidence, self-attribution, and security market under- and over-reactions, Northwestern University working paper, forthcoming in *Journal of Finance*.

Daniel, Kent D., and Sheridan Titman, 1997, Evidence on the characteristics of cross-sectional variation in stock returns, *Journal of Finance* 52, 1-33.

DeBondt, Werner F.M., and Richard H. Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793-808.

DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* 45, 379-395.

Edwards, W., 1968, Conservatism in human information processing, in B. Kleinmutz ed.: *Formal Representation of Human Judgment* (New York, John Wiley and Sons).

Fama, Eugene F., 1998, "Market efficiency, long-term returns, and behavioral finance, CRSP working paper #448, University of Chicago, forthcoming in *Journal of Financial Economics*.

Fama, Eugene F., and Kenneth R. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246-273.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene F., and Kenneth R. French, 1998, Value versus growth: The international evidence, *Journal of Finance* 53, 1975-1999.

Friedman, Milton, 1953. The case for flexible exchange rates. *Essays in Positive Economics* (University Of Chicago Press).

Grossman, Sanford J., 1976, On the efficiency of competitive stock markets where traders have diverse information, *Journal of Finance* 31, 573-85.

Grossman, Sanford J., and Joseph Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393-408.

Hamilton, James D., 1994. *Time Series Analysis* (Princeton University Press).

Hart, Oliver D., and David M. Kreps, 1986, Price destabilizing speculation, *Journal of Political Economy* 94, 927-52.

Haugen, Robert A. and Nardin L. Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401-439.

Hong, Harrison, and Jeremy C. Stein, 1997, A unified theory of underreaction, momentum trading and overreaction in asset markets, NBER working paper #6324.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 1998, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, NBER working paper #6553.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39, 181-208.

Ip, Greg, 1997, Red flag on wall street: Momentum investors, *Wall Street Journal* February 24.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jereski, Laura, and Deborah Lohse, 1996, Momentum-investing fever is sending some mutual funds to their sickbeds, *Wall Street Journal* July 26.

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1994, Contrarian investment, extrapolation and risk, *Journal of Finance* 49, 1541-1578.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23-51.

MacKinlay, A. Craig, 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3-28.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887.

Michaely, Roni, Richard H. Thaler, and Kent L. Womack, 1995, Price reactions to dividend initiations and omissions: Overreaction or drift? *Journal of Finance* 50, 573-608.

Poterba, James M., and Lawrence H. Summers, 1988, Mean reversion in stock returns: Evidence and implications, *Journal of Financial Economics* 22, 27-59.

Rehfeld, Barry, 1996, Mutual funds: What goes up, might go up more, *New York Times*, February 25.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341-360.

Rouwenhorst, K. Geert, 1998a, International momentum strategies, Yale University working paper, forthcoming in *Journal of Finance*.

Rouwenhorst, K. Geert, 1998b, Local return factors and turnover in emerging stock markets, Yale University working paper.

Scharfstein, David S., and Jeremy C. Stein, 1990, Herd behavior and investment, *American Economic Review* 80, 465-479.

Sharpe, William F., 1964, Capital Asset Prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Spiess, D. Katherine, and John Affleck-Graves, 1995, Underperformance in long-run stock returns following seasoned equity offerings, *Journal of Financial Economics* 38, 241-67.

Stein, Jeremy C., 1987, Informational externalities and welfare-reducing speculation, *Journal of Political Economy* 95, 1123-1145.

Womack, Kent L., 1996, Do brokerage analysts' recommendations have investment value? *Journal of Finance* 51, 137-168.

Footnotes

1. We discuss this empirical work in detail and provide references in Section I below.

2. We have more to say about these and other related papers in Section V below.

3. Although the model is simpler with just these two types of traders, the results are robust to the inclusion of a set of risk-averse, fully rational arbitrageurs, as shown in Section III.B.

4. The constraints that we put on traders' information-processing abilities are arguably not as well-motivated by the experimental psychology literature as the biases in BSV or DHS, and so may appear to be more ad hoc. However, they generate new and clear-cut asset-pricing predictions, some of which have already been supported in recent tests. See Section IV below.

5. As we discuss below, this "momentum externality" is reminiscent of the herding models of Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1992) and Scharfstein and Stein (1990).

6. Rouwenhorst (1998a,b) finds that Jegadeesh and Titman's (1993) U.S. results carry over to many other developed and emerging markets, though they are not statistically significant for every country individually. (See, e.g., Haugen and Baker (1996) on weak momentum in Japan.)

7. References include: Bernard and Thomas (1989, 1990) on earnings announcements; Loughran and Ritter (1995) and Spiess and Affleck-Graves (1995) on stock issues; Ikenberry, Lakonishok, and Vermaelen (1995) on repurchases; Michaely, Thaler, and Womack (1996) on dividend initiations and omissions; and Womack (1996) on analyst recommendations.

8. These results have been controversial, but seem to have stood up to scrutiny (Chopra, Lakonishok and Ritter (1992)). There are also direct analogs in the time series of aggregate market returns, although the statistical power is lower. See Fama and French (1988), Poterba and Summers (1988) and Cutler, Poterba and Summers (1991).

9. These results have also been found to be robust in international data. (Fama and French (1998), Rouwenhorst (1998b)) And again, there are analogous fundamental reversion patterns in the time-series literature on aggregate market predictability. (Campbell and Shiller (1988).)

10. See Lakonishok, Shleifer, and Vishny (1994), and MacKinlay (1995). Daniel and Titman (1997) directly dispute the idea that the book-to-market effect can be given a risk interpretation.

11. A somewhat more natural way to generate an infinite-horizon formulation might be to allow the asset to pay dividends every period. The only reason we push all the dividends out into the infinite future is for notational simplicity. In particular, when we consider the strategies of short-lived momentum traders below, our approach allows us to have these strategies depend only on momentum traders' forecasts of price changes, and we can ignore their forecasts of interim dividend payments.

12. Contrast this with a simpler setting where group 1 always sees all of ϵ_{t+z-1} first, then group 2 sees it second, etc. In this case, group 1 newswatchers are better-informed than their peers.

13. There is an element of time-inconsistency here, since in fact newswatchers may adjust their positions over time. Ignoring the dynamic nature of newswatcher strategies is more significant when we add momentum traders to the model, so we discuss this issue further in

Section II.B.

14. Strictly speaking, this result also requires that there be an initial "date 0" at which everybody is symmetrically informed.

15. In the NBER working paper version, we provide a detailed analysis of the comparative statics properties of the model with respect to k.

16. This sort of frontrunning effect is at the center of DeLong et al (1990).

17. See the NBER working paper version for a fuller treatment of this frontrunning issue.

18. Our experiments suggest that we only run into existence problems when <u>both</u> the risk tolerance γ and the information-diffusion parameter z <u>simultaneously</u> become very large--even an infinite value of γ poses no problem so long as z is not too big. The intuition will become clearer when we do the comparative statics, but loosely speaking, the problem is this: as z gets large, momentum trading becomes more profitable. Combined with high risk tolerance, this can make momentum traders behave so aggressively that our $|\phi|<1$ condition is violated.

19. This result is related to the fact that newswatchers have time-inconsistent strategies, so that in formulating their demands they ignore the fact that they will be transacting with momentum traders who will be trying to take advantage of them. Thus in some sense, the newswatchers are more irrational than the momentum traders in this model.

20. The appendix briefly discusses our computational methods.

21. Campbell, Grossman, and Wang (1993) suggest that this value of risk tolerance is about right for the market as a whole. Of course, for individual stocks, arbitrageurs may be more risk-tolerant, since they may not have to bear systematic risk. As we demonstrate below, our

results on overreaction tend to become more pronounced when risk tolerance is increased.

22. Although volatility rises with momentum trading, it is not necessarily (though it may be) "excessive" relative to a rational expectations benchmark. This is because we are starting from a point where there is underreaction, which leads to <u>lower</u> volatility than under a random walk.

23. The fact that momentum trading can increase both volatility and pricing errors serves as another counterexample to Friedman's (1953) famous claim that profitable speculation must stabilize prices. See also Hart and Kreps (1986), Stein (1987), and DeLong et al (1990).

24. When $j < z-1$, there is no longer a monotonic link between ϕ and the degree of overshooting. This is because the biggest momentum trades are already being unwound before newswatchers have fully incorporated a news shock into their forecasts.

25. Consequently, one might argue that virtually any behavioral model would be consistent with there being more predictability in small stocks.

26. Of course, analyst coverage is not an ideal proxy either, as it may be endogenously related to a number of other stock-specific factors besides size. So in various sensitivity tests, we also control for the correlation between analyst coverage and share turnover, industry factors, beta, and market-to-book.

27. Fama (1997) argues that this evidence is problematic for existing behavioral models, as they do not clearly predict that overreaction should be concentrated in smaller stocks.

28. As a benchmark, turnover on the NYSE has been in the range of 50-60 percent in recent years, implying an average holding period of 20-24 months. Of course, momentum traders

may have shorter horizons than the average investor.

29. Some care should be taken in testing this prediction, since assets with higher trading costs are likely to have more stale prices, which can induce spuriously positive autocorrelations in measured returns.

30. The quote is from Ip (1997). Among the large investors labeled momentum players are Nicholas-Applegate Capital Management, Pilgrim Baxter & Associates, Friess Associates, and Richard Driehaus, who was ranked first among 1,200 managers of all styles for the five years ended December 1995 by Performance Analytics, a pension advisory firm. See Rehfeld (1996).

31. See Ip (1997).

32. The consultant is Robert Moseson of Performance Analytics, quoted in Jereski and Lohse (1996). The momentum investor is Richard Driehaus, quoted in Rehfeld (1996).

33. BSV develop a regime-switching learning model, where investors wind up oscillating between two states: one where they think that earnings shocks are excessively transitory; and one where they think that earnings shocks are excessively persistent. DHS emphasize the idea that investors are likely to be overconfident in the precision of their private information, and that this overconfidence will vary over time as they learn about the accuracy of their past predictions.

34. In Cutler, Poterba, and Summers (1990), positive-feedback traders <u>can</u> make money, as there is background underreaction, like in our model. However, since the feedback behavior is assumed, rather than derived, their model does not yield many of the predictions discussed in Section IV.

35. Also, the model of DeLong et al (1990) does not really endogenously deliver reversals. Rather, prices are just forced back to fundamentals on a terminal date. In our model, the reversal phase is more endogenous, corresponding to the unwinding of momentum traders' positions. It also involves more complex dynamics, with the sort of damped oscillations seen in the figures.

Figure 1: Cumulative Impulse Response and Momentum Traders' Horizon

The momentum traders' horizon j takes on values of 6, 12, and 18. Base is the cumulative impulse response without momentum trac
other parameter values are set as follows: the information diffusion parameter z is 12, the volatility of news shocks is 0.5 and the risk
gamma is 1/3.

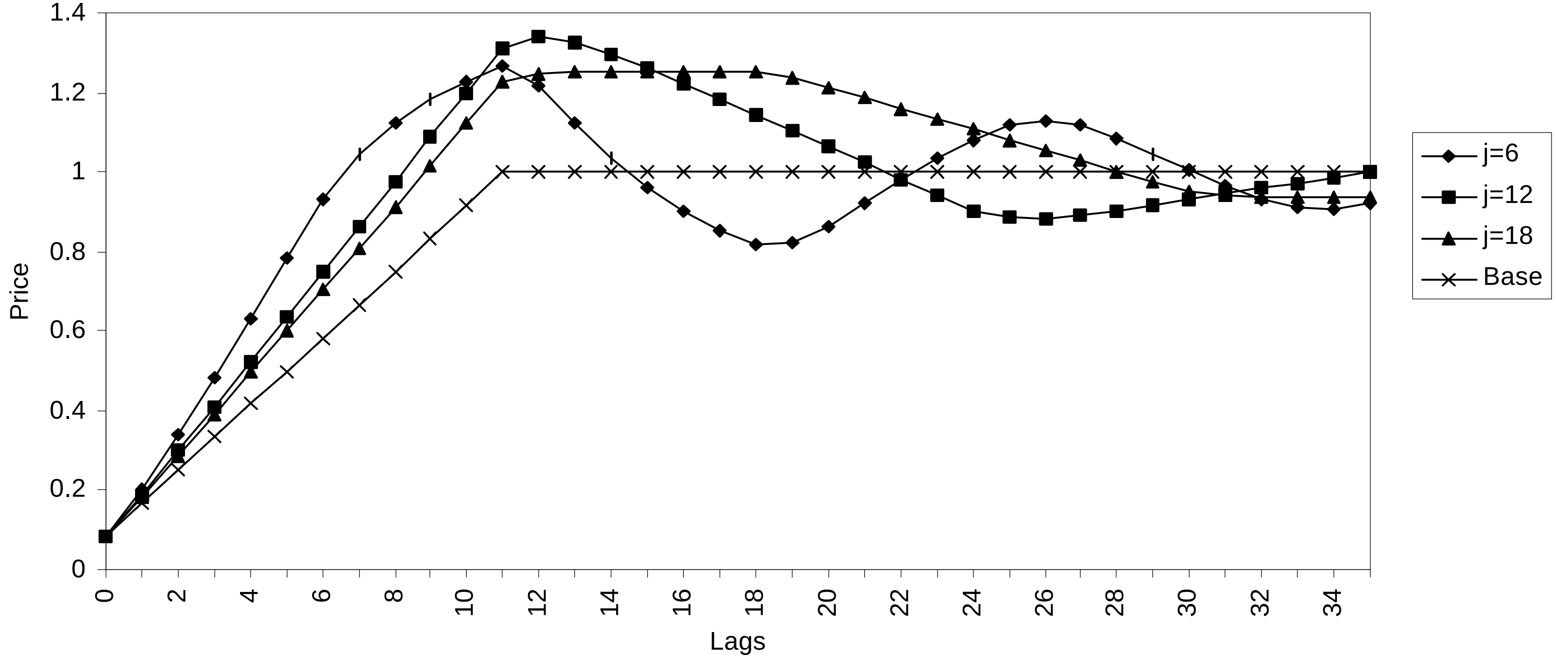

Figure 2: Cumulative Impulse Response and Momentum Traders' Risk Tolerance

The momentum traders' risk tolerance gamma takes on values of 1/11, 1/7 and 1/3. Base is the cumulative impulse response without momentum trading. The other parameter values are set as follows: the information diffusion parameter z is 12, the momentum traders' horizon j is 12 and the volatility of news shock is 0.5.

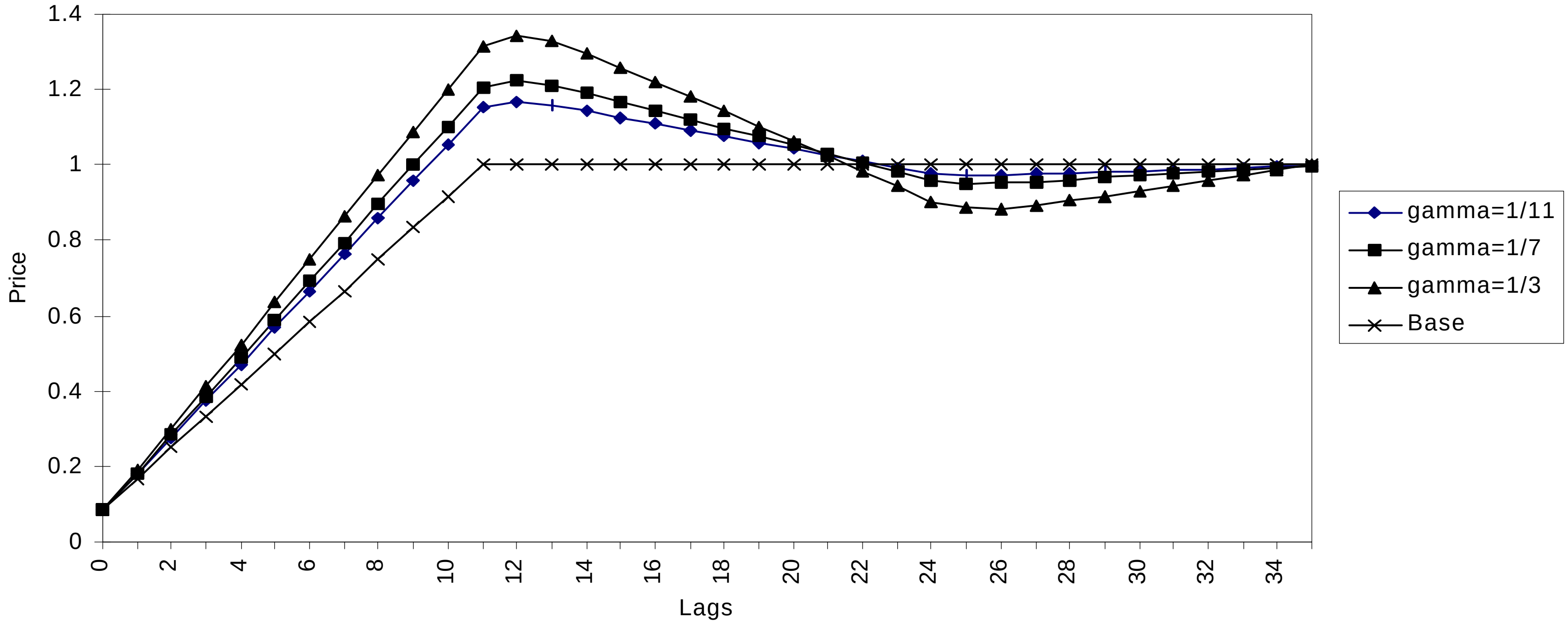

Figure 3: Cumulative Impulse Response and the Information Diffusion Parameter

The information diffusion parameter z takes on values of 3, 6, 9, and 12. The other parameter values are set as follows: momentum trade 12, the volatility of news shocks is 0.5, and momentum traders' risk tolerance gamma is 1/3.

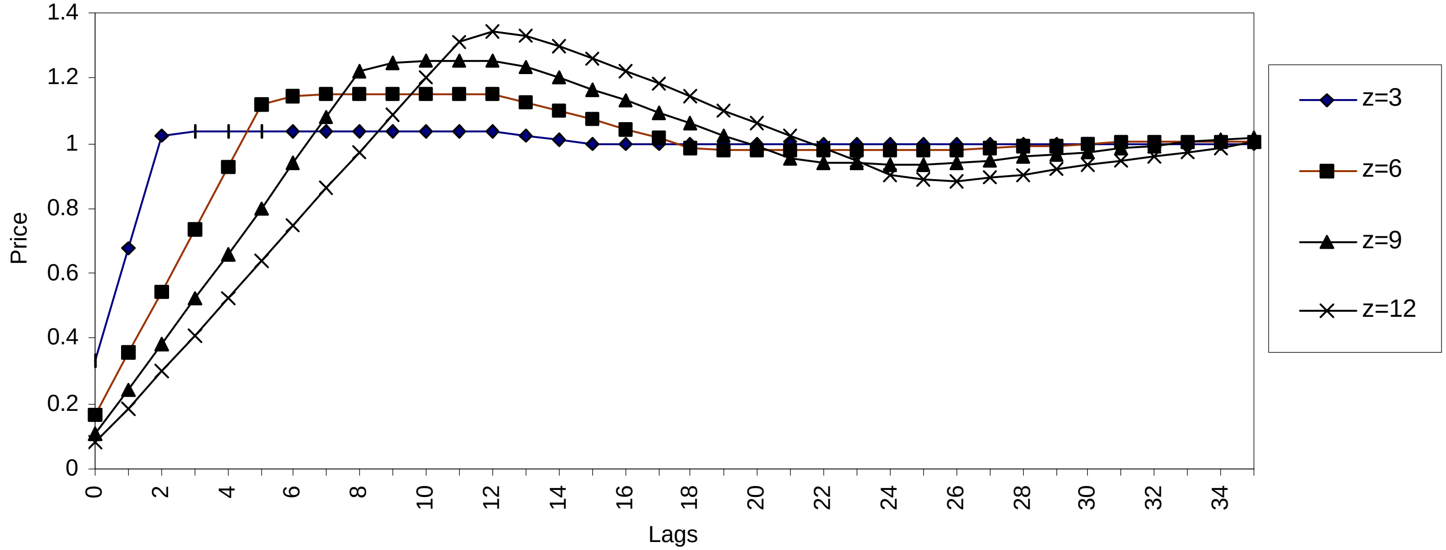

Figure 4: Cumulative Impulse Response and Contrarian Trading

Cumulative impulse responses for all-momentum trading equilibrium (w =1); the equilibrium in w hich traders endogenously choose w hether to follow either momentum or contrarian strategies (w =0.786); and the equilibrium in w hich traders can optimally condition on both momentum and contrarian variables ("Bo
The other parameter values are set as follow s: the information diffusion parameter z is 3, the momentum traders' horizon j is 1, the volatility of new s shock
and the risk tolerance gamma is 1/0.3. The contrarians are assumed to trade based on returns from three periods ago.