Limited Investor Attention and Stock Market Misreactions to Accounting Information

David Hirshleifer*, Sonya S. Lim**, and Siew Hong Teoh*

We provide a model in which a single psychological constraint, limited investor attention, explains both under- and over-reaction to different earnings components. Investor neglect of information in current-period earnings about future earnings induces post-earnings announcement drift and the profit anomaly. Neglect of earnings components causes accruals and cash flows to predict abnormal returns. We derive new untested empirical implications relating the strength of the drift, accruals, cash flow, and profit anomalies to the forecasting power of current earnings-related information for future earnings, the degree of investor attention to different types of information, and the volatilities of and correlation between accruals and cash flows. We also show that owing to costs of attention, in equilibrium some investors may decide not to attend to the implications of earnings or its components.


**Department of Finance, DePaul University, 1 E. Jackson Blvd, Chicago, IL 60604, slim1@depaul.edu

Formerly entitled “Limited Investor Attention and Earnings-Related Under- and Over-reactions.” This paper was presented at the American Accounting Association Annual Meetings in San Francisco. We thank Jennifer Altimuro (the conference discussant), Phil Davies, Danling Jiang, Kuan-Hui Lee, Rose Liao, Angie Low, Lin Sun, and Yinglei Zhang for helpful discussions and comments.
Limited Investor Attention
and
Stock Market Misreactions
to Accounting Information

Abstract

We provide a model in which a single psychological constraint, limited investor attention, explains both under- and over-reaction to different earnings components. Investor neglect of information in current-period earnings about future earnings induces post-earnings announcement drift and the profit anomaly. Neglect of earnings components causes accruals and cash flows to predict abnormal returns. We derive new untested empirical implications relating the strength of the drift, accruals, cash flow, and profit anomalies to the forecasting power of current earnings-related information for future earnings, the degree of investor attention to different types of information, and the volatilities of and correlation between accruals and cash flows. We also show that owing to costs of attention, in equilibrium some investors may decide not to attend to the implications of earnings or its components.
1 Introduction

Market reactions to earnings and earnings components present a striking puzzle. Stock prices on average under-react to earnings surprises (post-earnings announcement drift), but over-react to the operating accruals component of earnings.\(^1\) Earnings- and accruals-related patterns of return predictability are often referred to as ‘anomalies,’ ‘under-’ and ‘over-reactions,’ or as reflecting investor ‘optimism,’ ‘pessimism,’ or ‘naiveté’. Such labels offer little guidance as to the sources of these effects. Furthermore, a procedure of conjecturing a separate psychological bias for each misreaction pattern creates a problem of model overfitting; explanatory power is bought at the expense of predictive power.

We explore here the implications of a single well-established psychological constraint, limited investor attention, for the ability of earnings and earnings components to forecast abnormal returns. We argue that limited attention offers a parsimonious explanation for both under- and over-reactions to earnings and earnings components. The model of limited attention that we provide is consistent with post-earnings announcement drift, the accrual anomaly, the cash flow anomaly, and the profit anomaly. In addition to providing an integrated approach to earnings-related anomalies, our model offers new empirical implications about the determinants of the strength of mispricing effects. Specifically, we derive new untested implications for how return predictability is related to the volatilities, correlations, and earnings forecasting power of these components.

Attention requires effort, and since the amount of information available is vast, must be selective (see, e.g., Kahneman (1973)). There is evidence from the experimental laboratory that limited attention affects how both naive and sophisticated individual investors, as well as financial professionals, interpret accounting data (see the review of Libby, Bloomfield, and Nelson (2002)). Such evidence suggests that attentional constraints may be a source of investor misvaluation of accounting information.

Empirical research provides several indications that limited investor attention contributes to post-earnings announcement drift. Market reactions to earnings surprises are muted when the news is released to the media during low attention periods. After-hours

\(^1\)Evidence of post-earnings announcement drift is provided by Ball and Brown (1968) and Bernard and Thomas (1989); on the accrual and cash flow anomalies, see Sloan (1996) and Collins and Hribar (2000). Most empirical tests that control for possible risk factors have not supported a rational risk premium explanation for these return patterns (see the abovementioned studies, and Bernard, Thomas, and Wahlen (1997)).
earnings announcements are impounded into price gradually in the days after the disclosure (Francis, Pagach, and Stephan (1992)). The volume-reaction and two-day stock price reaction to news that is released to the media on Fridays and on days with many competing announcements are weaker and post-earnings announcement drift is stronger than when news is released on other days of the week and on days with few competing announcements (DellaVigna and Pollet (2009), Hirshleifer, Lim, and Teoh (2009)).

There are also indications that limited investor attention may play a role in the accrual anomaly. Institutional investors, as professionals, should be more attentive to earnings components than individuals. Consistent with this, the accrual anomaly is stronger among stocks with lower ownership by active institutional investors (Collins, Gong, and Hribar (2003)). Furthermore, managers seem to use their accounting discretion to exploit investors’ neglect of accruals information. Analysts, whether for agency or psychological reasons, tend to neglect accruals information in forming their forecasts (Teoh and Wong (2002)). Finally, the accrual anomaly is not present among the subset of firms that disclose the level of accruals at the date of the original earnings announcement (Chen, DeFond, and Park (2002)). This suggests that the accrual anomaly is caused by a subset of investors who attend to the earnings announcement but not to the later financial reporting of accruals.

Contributing to a focus on earnings by both individual investors and financial professionals, the business media discuss earnings much more than cash flow and accruals numbers. As related by the well-known stock analyst Abby Joseph Cohen, “Many participants in the investment business still rely on EPS [earnings per share], to the exclusion of important measures of firm performance, such as revenues and cash flow…” (Cohen (2005)). This emphasis is useful for investors with limited attention and processing power. If an investor must select between earnings and cash flow to focus upon, the best choice is the signal that is most informative about firm value. Empirically, stock returns are more strongly related to news about contemporaneous earnings than about cash flow (Dechow (1994)). Nevertheless, since cash flow is incrementally informative relative to earnings, the neglect of how earnings is divided between cash flow and accruals causes

\[2\]

Furthermore, firms seem to take limited investor attention into account in timing the release of earnings news. There is evidence that firms defer the release of bad earnings news to Fridays rather than other weekdays (Bagnoli, Clement, and Watts (2005), DellaVigna and Pollet (2009)), and to after-hours rather than trading hours (Patell and Wolfson (1982), Bagnoli, Clement, and Watts (2005)).

\[3\]

This is confirmed in an extensive literature on earnings management prior to new issues of equity (Teoh, Welch and Wong (1998b, 1998a)), in order to meet benchmarks (DeGeorge, Patel, and Zecchhauser (1999)), and for other purposes; and evidence that abnormal accruals (a proxy for the use of discretion by firms) are misvalued (Teoh, Welch and Wong (1998b, 1998a) and Xie (2001)).
systematic bias.

Consistent with some past literature on limited attention, in our model some investors condition only on subsets of publicly available information signals in valuing a stock. Risk averse investors who are fully attentive to the relevant information item are willing to bear only a limited amount of risk in order to exploit mispricing. In consequence, equilibrium stock prices reflect a weighted average of the beliefs of investors who attend to different signals, with weights that depend on the relative numbers in each investor group and their risk tolerances; see also the discussion in footnote 9. This approach, which is by now familiar, serves as the building block for our new findings about the effects of limited attention toward earnings and earnings components.

In the model, some investors attend to the implications of current-period earnings for future prospects, and a subset of these investors also attend separately to the accrual and cash flow components of earnings. (Further motivation of the assumption that some investors neglect earnings information is provided in Section 2 and Section 5.) Using the signals they attend to, investors form valuations of the firm. In equilibrium, prices underreact to earnings surprises because of the subset of investors who do not incorporate this information into their expectations of future earnings.

Investors who do attend to earnings but do not distinguish between earnings components misvalue firms with abnormal levels of accruals. Empirically, the level of accruals is a less favorable forecaster than cash flow of firm profitability (Sloan (1996)). Regardless of whether this difference in forecasting power is a consequence of earnings management or of the general nature of the accounting system, a rational investor should take this fact into account in valuing firms. In our model, an investor who attends to earnings but does not impound the information in earnings components into his valuation overvalues high-accruals firms and undervalues low-accruals firms. Since misvaluation is eventually corrected, this effect can, under appropriate conditions, cause high accruals to predict low subsequent abnormal returns, and high cash flows to predict high subsequent abnormal returns. Thus, the analysis reconciles underreaction to earnings with overreaction to accruals.

Using this framework, we provide untested empirical implications about the direction and strength of the forecasting power of earnings surprises, accruals, and cash flow for future returns. The intuition starts from the fact that, conditional upon high accruals, contemporaneous earnings also tend to be high, which causes investors to forecast higher future earnings. In this situation, an investor who attends to earnings but neglects accruals is overoptimistic because he forecasts high future earnings based upon high
current-period earnings. When cash flow is high, an individual who neglects this fact is overpessimistic (even if he also neglects earnings).

These effects imply the direction and strength of accruals or cash flows as return predictors depends on how favorably these variables predict future earnings relative to the forecasting power of current-period earnings as a predictor of future earnings; and on what information investors neglect. The predictive powers of these variables for future earnings are influenced by the relative variability of cash flows and accruals, the correlation of these earnings components, and the quality of accruals. In consequence, the more variable are accruals relative to cash flows, the stronger is the cash flow anomaly relative to the accrual anomaly; higher correlation between cash flows and accruals tends to weaken the accrual anomaly; and lower quality of accruals (a weaker incremental ability of accruals to forecast future earnings) tends to strengthen the accrual anomaly.

When some investors neglect earnings, and others attend to earnings but neglect accruals, stock prices underreact to earnings, and overreact to accruals relative to cash flow. If enough attention is paid to earnings, the relative overreaction to accruals out-weighs the general underreaction to earnings, so that overall price overreacts to accruals. Thus, in our setting the condition for the accrual anomaly to exist differs from that suggested by Sloan (1996), that accruals are a stronger forecaster than cash flows of future earnings. Furthermore, there tends to be stronger underreaction to cash flow than to earnings; the ratio of the cash flow effect on returns to the accruals effect on returns is greater in absolute value than the ratio of the variance of accruals to the variance of cash flow.

We also show that owing to investor neglect of the implications of earnings news for future earnings, the current level of earnings predicts stock returns even after controlling for the current earnings surprise. Thus, the model explains the profit anomaly (Balakrishnan, Bartov, and Faurel (2010), Chen, Novy-Marx, and Zhang (2010)). The demands of inattentive investors are based on prior beliefs about expected future earnings as reflected in indicators such as the previous year’s earnings or analyst forecasts of the current year’s earnings. These indicators are less accurate than realized current year earnings as predictors of future earnings. In consequence, mispricing is correlated with the current level of earnings.

Finally, we consider a setting where individuals decide the level of attention by balancing the cost of attention against the expected benefits. We show that in equilibrium some investors may decide not to attend to the implications of public information about earnings or its components. This analysis explains why inattentive investors who trade
based upon their inattentive expectations (rather than simply remaining on the sidelines) can survive in the long run.

This paper is part of a recent theoretical literature on how constraints on information processing affect investor behavior. The approach followed here is similar in spirit to that of Hirshleifer and Teoh (2003), who study the effects on market prices of investors neglecting relevant accounting information or strategic aspects of the disclosure and reporting environment. A key difference here is that we examine the implications of limited attention for market misvaluation in relation to earnings surprises, the level of earnings, accruals, and cash flows. Other recent papers model the allocation of attentional resources (Gabaix and Laibson (2005), Peng (2005), and Hirshleifer, Lim, and Teoh (2008)), how limited learning capacity affects asset price comovement (Peng and Xiong (2006)) and the speed of price adjustment to fundamental shocks (Peng (2005) and Peng and Xiong (2006)), how delayed processing of new information affects the dynamics of asset price volatility (Peng and Xiong (2002)), how neglect of demographic information affects asset prices (DellaVigna and Pollet (2007)) and how informed parties make disclosure decisions when observers have limited attention (Hirshleifer, Lim, and Teoh (2008)).

2 The Economic Setting

We assume that there are three types of investors: those that ignore all current earnings-related information, those that attend only to a subset of that information, and those that attend to all publicly available information. All investors are ex ante identical, and differences in how much of the public information set they process depend on the cost they have incurred to obtain information. We assume that inattentive investors, apart from the specific signals that they ignore, update beliefs as rational Bayesians.

There are 2 dates. At date 1, cash flow $c_1$, accruals $a_1$, and earnings $e_1 = c_1 + a_1$ are realized. Investors update their prior beliefs based upon whatever public signals they observe. At date 2 terminal earnings, cash flows, and accruals, $e_2$, $c_2$, and $a_2$, are realized, where $e_2 = c_2 + a_2$, and the firm is liquidated. In general, accounting accruals must reverse out, so we assume that $a_2 = -a_1$.

Prices are set by trading in a securities market with no private information, so that a rational individual has nothing to learn from market price. An inattentive individual who is unaware of his signal neglect will also think he has nothing to learn from market price. We therefore assume that inattentive investors do not update their beliefs based
upon market price.\textsuperscript{4}

Nevertheless, if an investor understood his attentional limits, he could in principle adjust for them by deferring to the belief implicit in market price. Indeed, a discrepancy between an investor’s valuation and the market price could alert the investor to his information neglect. In general, however, the same constraints on processing power and memory that make it hard to attend to some public signal also make it hard to use price or other indicators to compensate optimally for the failure to attend to it.\textsuperscript{5} For example, an investor may just not think about the source of the discrepancy between the market price and his own valuation. So long as some fraction of inattentive investors have imperfect self-awareness, results similar to those derived here will obtain. Furthermore, even if individuals always attend to market price and are fully aware of their information neglect, similar results to those we derive here could be obtained so long as there is ‘noise’ in market price arising from liquidity trading.\textsuperscript{6}

In general, attending to more information is costly. Since investors have finite cognitive resources, attending to some information implies less time and resources for other activities. We assume that there are different investor groups indexed by $i$ who attend to different information sets. Fully attentive investors attend to all date 1 publicly available information; investors with limited attention attend to subsets of date 1 public information. The decision to incur a cost to attend to the information set precedes the trading decision. At this stage, we focus on the trade decision at date 1 taking the information set as given. Later in Section 4.2 we analyze the decision of how much information an individual attends to if there is a cost to attending to information.

\textsuperscript{4}Observing the ‘wrong’ price is an event which, as perceived by the investor, is not supposed to occur in equilibrium. In the Perfect Bayesian Equilibrium concept of game theory setting the individual’s posterior beliefs in such a situation equal to the prior belief can be consistent with equilibrium. Similar results would hold so long as some disagreement remains between the attentive and inattentive investors, i.e., inattentive investors do not always abandon their beliefs in favor of the information implicit in the market price.

\textsuperscript{5}Since in reality people face many relevant signals, they try to leverage their attention by focusing on more important information items. However, we cannot know perfectly which are more important before processing them, which makes it hard to determine how to optimally compensate for information neglect. Section 5 discusses evidence suggesting that individuals fail to compensate fully for the consequences of limited attention in making decisions.

\textsuperscript{6}In such a setting, an individual who attends to a given public signal in effect has a sort of ‘private’ information, so different individuals who attend to different public signals will trade and profit at the expense of liquidity traders, in the spirit of the models of Grossman and Stiglitz (1976) and Diamond and Verrecchia (1981). This approach is discussed more fully in Section 5.
We assume that investors have a mean-variance utility function

\[ E^i[C^i_2] - \left( \frac{A}{2} \right) \text{var}^i(C^i_2), \]  

where \( C^i_2 \) is terminal consumption, \( A \) is the coefficient of absolute risk aversion, and \( i \) superscripts denote the expectation or variance as formed by group \( i \).

We assume that the decision-making investors have an initial wealth endowment (i.e., claims to terminal consumption) of \( W \), and zero shares of the risky security. In addition, mechanistic noise traders bring to the market a stochastic per capita supply of the single risky security of \( x_0 \) (i.e., supply per member of the decision-making population), which is normally distributed with mean zero and variance \( V_x \), uncorrelated with the other exogenous random variables (earnings and its components).\(^7\) At date 1, the individual can buy or sell the security in exchange for ‘cash’ (claims to terminal consumption) at price \( S_1 \). The position in the security he attains is denoted \( x^i \). Let \( S_2 \) be the true value of the stock, which is conclusively revealed to all at date 2. Then letting \( k^i \) be the cost of attending to the information set used by attention group \( i \), the consumption of an individual in group \( i \) is

\[ C^i_2 = W - k^i + x^i(S_2 - S_1). \]  

(2)

Thus, an individual in attention group \( i \) solves

\[ \max_{x^i} x^i(E^i[S_2] - S_1) - \left( \frac{A}{2} \right) \text{var}^i(x^i S_2). \]  

(3)

As a preliminary building block, we verify in our setting a standard finding about stock prices as a weighted average of investor expectations of terminal cash flows as adjusted by a risk premium. Differentiating the objective with respect to \( x^i \), equating to zero and solving yields

\[ x^i = \frac{E^i[S_2] - S_1}{A \text{var}^i(S_2)}. \]  

(4)

In several behavioral models, there are both rational and imperfectly rational investors who are risk averse and who maximize their expected utilities with respect to their correct or incorrect beliefs. In this literature, prices in equilibrium reflect a weighted average of the valuations of the rational and irrational traders. Even if there

\(^7\)For most of the paper it is not necessary that \( x_0 \) be stochastic. However, in Section 4 ‘noise’ or ‘liquidity’ trading is needed for inattentive investors to earn positive expected profits from trading. As in Grossman and Stiglitz (1976), noise creates scope for other investors to earn expected profits from trading.
are no market imperfections, both groups influence prices significantly owing to the finite risk-bearing capacity of each group.

Letting $f^i$ denote the fraction of investors in attention group $i$, the security price is determined by the market clearing condition

$$\sum_i f^i x^i = x_0.$$  \hspace{1cm} (5)

Substituting for $x^i$ from (4), and solving for $S_1$ gives

$$S_1 = \frac{\sum_i \lambda^i E^i [S_2] - Ax_0}{\sum_i \lambda^i},$$  \hspace{1cm} (6)

where

$$\lambda^i \equiv \frac{f^i}{var^i(S_2)}.$$  \hspace{1cm} (7)

By normality, the $\lambda^i$’s are constants independent of the signal realizations used by investors to condition beliefs.

This confirms that in equilibrium prices are a weighted average of the beliefs about terminal cash flows of different investors adjusted by a risk premium ($Ax_0 / \sum_i \lambda^i$), with weight $\lambda^i / \sum_i \lambda^i$ on each information set. By (7), ceteris paribus $\lambda^i$ is increasing in $f^i$. Thus, the greater the likelihood of each investor being inattentive, the greater the weight that inattentive investors play in determining prices.

To focus on the effect of investor attention on investor mean beliefs, we follow the heuristic trading approach of Verrecchia (2001), in which the biases of traders are manifested in their expectations (first moments) but not second moments. So for tractability we model the expectations of attentive and inattentive investors (6) as depending on $i$, but the variances in (7) as independent of $i$, $var^i(S_2) = var(S_2)$.$^8$

So the equilibrium price simplifies to

$$S_1 = \sum_i f^i E^i [S_2] - A var(S_2)x_0.$$  \hspace{1cm} (8)

In this setting rational investors exploit a trading strategy that earns predictable abnormal returns relative to a fully rational asset pricing benchmark. Nevertheless, even though markets are perfect and there are no restrictions on either long positions or

$^8$It is not essential that the groups all assess variance as the unconditional variance; for example, all results would still hold if variance were assessed as the variance conditional upon all the available public signals.
short-selling, fully attentive investors do not completely arbitrage away the mispricing generated by inattentive investors. The reason they do not is because doing so is risky.

The intuition behind the traditional notion that rational investors dominate price is that rational investors trade to arbitrage away mispricing. However, if prices were set solely by the rational investors, imperfectly rational investors would perceive a profit opportunity to trade against what they regard as mispricing. If all investors are risk averse, the equilibrium outcome reflects a weighted average among these beliefs.\(^9\)

3 The Accrual Anomaly and Post-Earnings Announcement Drift

The next subsection describes the basic structure of cash flow, earnings, and accruals over time. Subsection 3.2 establishes that there is post-earnings announcement drift. Subsection 3.3 shows that the same setting also generates overreaction to the accruals component of earnings, and that cash flows positively predict subsequent abnormal returns. Subsection 3.4 derives new empirical predictions about the circumstances in which the accrual and cash flow anomalies will be stronger or weaker. Subsection 3.5 shows that limited attention also provides an explanation for the profit anomaly, the predictability of returns based upon the level of earnings.

3.1 The Relation of Earnings, Accruals and Cash Flow over Time

Suppose that at date 1 fraction \(f^u\) of investors ignore current-period earnings \(e_1\) and remain at their prior belief,\(^{10}\) that fraction \(f^e\) of investors attend to earnings \(e_1\) but not to accruals and cash flow separately, and that the remaining fraction \(f^a = 1 - f^u - f^e\) attend to both accruals and cash flow and therefore fully take into account the information in earnings, since

\[ e_1 = c_1 + a_1. \quad (9) \]

\(^9\)Weighted average forms are found in previous models; see, e.g., Kandel and Pearson (1995), who provides evidence supportive of prices reflecting average trader perceptions; and in the context of limited attention, Hirshleifer and Teoh (2003).

\(^{10}\)It is not essential for the main conclusions of this paper that these investors completely ignore earnings. They could attend to the direct effect but ignore the positive implications of higher immediate earnings for future earnings. Taking this into account would slightly modify the algebra but would not substantively affect the models qualitative predictions.
Then by equation (8), the date 1 stock price is

\[ S_1 = f^u E[S_2] + f^e E[S_2|e_1] + f^a E[S_2|a_1, c_1] - \text{Var}(S_2)x_0. \] (10)

For tractability, we assume multivariate normality of the stochastic variables. As a result, date 2 earnings can be expressed as a linear function of the date 1 variables \( a_1 \) and \( c_1 \) and a noise term \( \delta \) as

\[ e_2 = \beta_0 + \beta_1 c_1 + \beta_2 a_1 + \delta, \] (11)

where \( \text{cov}(\delta, c_1) = \text{cov}(\delta, a_1) = 0 \). Consistent with past empirical evidence (Decow (1994)), we assume that cash flow is a more positive predictor of future earnings than accruals, \( \beta_1 > \beta_2 > 0 \).\(^{11}\) Since earnings \( e_1 = a_1 + c_1 \), equation (11) directly implies that in a regression of future earnings \( e_2 \) on current period earnings and accruals, there is a negative coefficient on \( a_1 \). Similarly, in a regression of future earnings on current period earnings and cash flow, there is a positive coefficient on cash flow.

Letting bars denote unconditional expectations, by (11) we have

\[ \bar{e}_2 = \beta_0 + \beta_1 \bar{c}_1 + \beta_2 \bar{a}_1, \] (12)

so the deviation of earnings from its unconditional mean is

\[ e_2 - \bar{e}_2 = \beta_1(c_1 - \bar{c}_1) + \beta_2(a_1 - \bar{a}_1) + \delta. \] (13)

The terminal realized value of the stock is the sum of the cash flows at the two periods. Under clean surplus accounting, this is also equal to the sum of the earnings, so

\[ S_2 = c_1 + c_2 = e_1 + e_2 \]

\[ E[S_2] = \bar{e}_1 + \bar{e}_2 = \bar{e}_1 + \bar{e}_2. \] (14)

High earnings at date 1 is linearly associated with high earnings at date 2. The strength of this relation is given by the regression coefficient, which by (9) and (11) is

\[ \beta_{e_2,e_1} = \frac{\text{cov}(\beta_1 c_1 + \beta_2 a_1 + \delta, c_1 + a_1)}{V_{e_1}} \]

\[ = \left( \frac{V_c + C}{V_c + V_a + 2C} \right) \beta_1 + \left( \frac{V_a + C}{V_c + V_a + 2C} \right) \beta_2, \] (15)

\(^{11}\)An extra unit of either cash flow or accruals at date 1 increases date 1 earnings by one dollar. Thus, precisely because \( \beta_1 > \beta_2 > 0 \), expected firm value \( E[S_2|a_1, c_1] = E[e_1 + e_2|a_1, c_1] \) increases more when cash flow increases by one dollar than when accruals does.
where \( C \equiv \text{cov}(a_1, c_1) \), and where \( V \) denotes the variance of a variable, with the abbreviations \( V_a \) for \( V_{a_1} \) and \( V_c \) for \( V_{c_1} \).

Thus, \( \beta_{e_2e_1} \) is a weighted average of \( \beta_1 \) and \( \beta_2 \). Since \( \beta_1 > \beta_2 \), it follows that \( \beta_1 > \beta_{e_2e_1} \) > \( \beta_2 \) so long as both weights are positive. Both weights are positive under the assumption that each of higher accruals and higher cash flow is associated with higher contemporaneous earnings.

\[
\begin{align*}
\text{cov}(e_1, c_1) &= V_c + C > 0 \\
\text{cov}(e_1, a_1) &= V_a + C > 0.
\end{align*}
\]

This is plausible; accruals and cash flows are components of earnings, and empirical studies have documented that earnings is positively correlated with accruals and cash flows (e.g., Dechow and Ge (2006), Pincus, Rajgopal, and Venkatachalam (2007)).

We also assume that accruals are positively related to the sum of date 1 and date 2 earnings,

\[
\text{cov}(e_1 + e_2, a_1) = (1 + \beta_2)V_a + (1 + \beta_1)C > 0.
\]

This assumption is intuitively obvious; it requires that accounting adjustments not be completely meaningless, so that a higher accrual is on average associated with higher long-run total firm value \( e_1 + e_2 \). For example, as discussed above, earnings and accruals are positively contemporaneously correlated, so a sufficient condition for this to hold is that accruals be positively correlated with future earnings, \( \text{cov}(e_2, a_1) > 0 \). Although Sloan (1996) finds that the accrual component of earnings is less persistent than the cash flow component, his estimates also indicate that high accruals are associated with higher future earnings.\(^\text{12}\) Furthermore, condition (17) is also essentially equivalent (under efficient markets) to the finding of Dechow (1994) that stock returns covary more strongly with innovations in earnings than cash flows.\(^\text{13}\)

Under these conditions, (15) indicates that high current-period cash flow is a more favorable forecaster of future earnings than is high current-period earnings, which in

\(^\text{12}\)In a regression of one year-ahead earnings on accruals and cash flows, Sloan (1996) and Dechow and Ge (2006) reports that the coefficient on both variables are significantly positive.

\(^\text{13}\)Since \( e_1 + e_2 \) is total firm value, the covariance in (17) is \( \text{cov}(e_1 + e_2, e_1 - c_1) = \text{cov}(e_1 + e_2, e_1) - \text{cov}(e_1 + e_2, c_1) \). By the law of iterated expectations, this last expression is equal to \( \text{cov}(E[e_1 + e_2], e_1) - \text{cov}(E[e_1 + e_2], c_1) \), where the inner expectation is conditional on all publicly available date 1 information. In an efficient market the inner expectation is just the stock price, so if \( \text{cov}(S_1, e_1) - \text{cov}(S_1, c_1) > 0 \) (which is basically the Dechow finding, since at date 0 any prior level of earnings, cash flows, or stock price are constant) then (17) holds. In our model markets are not efficient, but as long as some investors are attentive to earnings components, the general point still holds that, \( \text{eteris paribus} \), a tendency of high accruals to covary positively with total firm value will tend to increase the covariance of stock returns with earnings as compared to cash flow.
turn is a more favorable forecaster than high current-period accruals. It follows that the relation between date 2 and date 1 earnings is stronger the more variable are cash flows relative to accruals.

### 3.2 Post-Earnings Announcement Drift

We next examine the relation between the date 1 earnings surprise and expected future abnormal stock returns. We define the earnings surprise as \( e_1 - \bar{e}_1 \), where \( \bar{e}_1 \) is the prior expectation of date 1 earnings. Since \( \bar{e}_1 \) is a constant, conditioning on the earnings surprise is equivalent to conditioning on the date 1 earnings \( e_1 \). We begin by calculating, conditional on earnings \( e_1 \), the expected future value of the stock \( E[S_2|e_1] \).

For tractability, we examine price changes rather than percentage returns, as is standard in much of the literature on information in securities markets (see, e.g., Verrecchia (2001) and Peng (2005)). Using standard properties of conditional expected values with multivariate normal distributions,

\[
E[S_2|e_1] = e_1 + E[e_2|e_1] \\
= e_1 + \bar{e}_2 + (e_1 - \bar{e}_1)\beta_{e_2e_1} = E[S_2] + (1 + \beta_{e_2e_1})(e_1 - \bar{e}_1). \tag{18}
\]

To calculate the conditional expected return given \( e_1 \), we also need \( E[S_1|e_1] \). Drift in our model means that the expected change in the stock price conditional on the earnings surprise increases with \( e_1 \), with zero effect when realized earnings is equal to its prior expectation. We calculate the expected post-event return as

\[
E[S_2 - S_1|e_1] = E[S_2|e_1] - E[S_1|e_1]. \tag{19}
\]

An iterated conditional expectation where one conditioning information set is coarser than the other is equivalent to a single expectation conditional on the coarser information set (e.g., Ash (1972), Theorem 6.5.10). Since \( e_1 = a_1 + c_1 \), the information provided by \( e_1 \) is coarser than that provided by \((a_1, c_1)\), so

\[
E[E[S_2|a_1, c_1]|e_1] = E[S_2|e_1]. \tag{20}
\]

\( E[S_2|e_1] \) is given in equation (18). It follows from (10) and (18) that the expected return after earnings announcement is

\[
E[S_2 - S_1|e_1] = E[S_2|e_1] - \{f^u E[S_2] + (f^e + f^a)E[S_2|e_1]\} \\
= f^u (E[S_2|e_1] - E[S_2]) \\
= f^u (1 + \beta_{e_2e_1})(e_1 - \bar{e}_1). \tag{21}
\]
We can define misvaluation in the current stock price as the difference between \( S_1 \) and \( S_1^* \), the price that would be set if all investors were attentive \((f^a = 1)\). From (10) and (20),
\[
E[S_1^*|e_1] = E[E[S_2|a_1, e_1] - Avar(S_2)x_0|e_1] = E[S_2|e_1].
\]
Therefore, since
\[
E[S_1^* - S_1|e_1] = E[S_2 - S_1|e_1]
\]
there is over (under)-valuation if and only if the expected price change is negative (positive).

From (10), the average immediate price reaction to the date 1 earnings announcement conditional upon an earnings realization \( e_1 \) is
\[
E[S_1|e_1] - E[S_1] = (1 - f^u)(E[S_2|e_1] - E[S_2]) = (1 - f^u)(1 + \beta_{e_2e_1})(e_1 - \bar{e}_1). 
\]
(22)

This proves:

**Proposition 1**

1. If fraction \( f^u > 0 \) of investors neglect the information in current-period earnings as well as cash flow and accruals, then more positive earnings surprises are associated with greater undervaluation of the firm, and more negative surprises with greater overvaluation; and there is post-earnings announcement drift.

2. The average immediate price reaction to the earnings surprise, the extent to which the firm is undervalued (or less overvalued), and average post-event abnormal returns increase with the earnings surprise \( e_1 - \bar{e}_1 \).

3. The strength of the post-earnings announcement drift anomaly, and the sensitivity of misvaluation and the average immediate price reaction to the earnings surprise, increase with the persistence of earnings, \( \beta_{e_2e_1} \).

4. The higher the fraction of investors who neglect earnings information, \( f^u \), the weaker is the average immediate reaction to a given earnings surprise, the stronger is the relation of misvaluation to the earnings surprise, and the stronger is post-earnings announcement drift.

As a special case, suppose that earnings follows a random walk, and consider a previous date 0 at which time the firm earns \( e_0 = \bar{e}_1 \). Then the surprise is the change in earnings \( e_1 - e_0 \). So the above predictions can be applied to one-year changes in earnings.

A recent body of research provides evidence which is generally supportive of the persistence prediction of Part 3. Livnat (2003) provides evidence that higher earnings persistence is associated with stronger drift, consistent with Part 3. Furthermore, since
revenues have greater persistence than expenses, drift should be especially strong if
the earnings surprise comes largely from a surprise in revenues (Jegadeesh and Livnat
(2006)). Jegadeesh and Livnat (2006) and Gu, Jain, and Ramnath (2005) find that
analyst forecasts also do not fully reflect differences in persistence between revenues
and expenses. Chen (2009) finds, consistent with the prediction in Part 3, that drift
is strongest among high persistence firms, but reports a reverse drift relation for firms
with very low persistence, a phenomenon our model does not explain.

Part 4 implies that the immediate price reaction is weaker and drift is stronger
when a greater fraction $f^a$ of investors are inattentive to earnings. Possible proxies for
inattention to earnings include the occurrence of earnings announcements on Fridays
(DellaVigna and Pollet (2009)), high number of earnings announcements competing for
attention (Hirshleifer, Lim, and Teoh (2009)), and announcement after trading hours
(Bagnoli, Clement, and Watts (2005)). For these proxies, Part 4 is consistent with the
evidence discussed in the introduction that there is weaker average immediate reaction to
earnings news and stronger drift when attention is weaker. For given size of the earnings
surprise, greater neglect of earnings in our model causes a weaker immediate reaction
to earnings news, and a greater long term reaction (drift). In addition, attention to
earnings is likely to increase with investor sophistication. If institutional holdings are
viewed as a proxy for investor sophistication, then this implication is consistent with the
evidence that firms whose shares are held heavily by individuals instead of institutions
have stronger drift (Bartov, Krinsky, and Radhakrishnan (2000)).

The empirical implications of Proposition 1 regarding under- and over-valuation can
be tested using contemporaneous misvaluation measures. For example, the residual
income model of Ohlson (1995) provides a contemporaneous measure of fundamental
firm value. As a result, the ratio of market price to the residual income model valua-
tion provides a measure of misvaluation (as applied, for example, by Lee, Myers, and
Swaminathan (1999) and Dong, Hirshleifer, Richardson, and Teoh (2006)).

3.3 The Accruals and Cash Flow Anomalies

We now examine whether accruals or cash flows are forecasters of future returns.\footnote{Since earnings is the sum of cash flows and accruals, a bivariate regression of future returns on earnings and on accruals is essentially equivalent to a bivariate regression of future returns on earnings and on cash flows. However, the univariate regressions of returns on accruals or on cash flows are not equivalent.}

Specifically, we examine whether, conditional on the level of accruals (or cash flows), the
stock is mispriced, thereby implying subsequent abnormal returns. To do so, we calculate the expectation of the date 2 value of the stock conditional on accruals, \( E[S_2|a_1] \), and of the date 1 stock price conditional on accruals, \( E[S_1|a_1] \). The difference between these two values is the expected price change conditional on the level of accruals.

It follows from a property of conditional expectations (e.g., Ash (1972)) that

\[
E[E[S_2|a_1, c_1]|a_1] = E[S_2|a_1],
\]

as the information in \( a_1 \) is coarser than that in \((a_1, c_1)\). So the expected price change conditional on accruals \( a_1 \) is

\[
E[S_2 - S_1|a_1] = E[S_2|a_1] - (f^u E[S_2] + f^e E[S_2|e_1]|a_1] + f^a E[S_2|a_1]) = f^u (E[S_2|a_1] - E[S_2]) + f^e (E[S_2|a_1] - E[E[S_2|e_1]|a_1]).
\] (24)

From (13) and (14),

\[
S_2 = E[S_2] + (1 + \beta_1)(c_1 - \bar{c}_1) + (1 + \beta_2)(a_1 - \bar{a}_1) + \delta.
\] (25)

\[
E[S_2|a_1] = E[S_2] + (1 + \beta_1) \left( \frac{C}{V_a} \right) (a_1 - \bar{a}_1) + (1 + \beta_2)(a_1 - \bar{a}_1)
\]

\[
= E[S_2] + (a_1 - \bar{a}_1) \left[ (1 + \beta_1) \left( \frac{C}{V_a} \right) + (1 + \beta_2) \right].
\] (26)

By (18) and multivariate normality,

\[
E[E[S_2|e_1]|a_1] = E[S_2] + (a_1 - \bar{a}_1)(1 + \beta_{e_2 e_1}) \left( 1 + \frac{C}{V_a} \right).
\] (27)

It follows that

\[
E[S_2 - S_1|a_1] = (a_1 - \bar{a}_1) \left\{ f^u \left[ (1 + \beta_1) \left( \frac{C}{V_a} \right) + (1 + \beta_2) \right] + f^e \left[ (\beta_1 - \beta_{e_2 e_1}) \left( \frac{C}{V_a} \right) + (\beta_2 - \beta_{e_2 e_1}) \right] \right\}
\]

\[
= (a_1 - \bar{a}_1) \left\{ f^u \left[ (1 + \beta_1) \left( \frac{C}{V_a} \right) + (1 + \beta_2) \right] + f^e (\beta_1 - \beta_2) \frac{C^2 - V_a V_c}{V_a(V_a + V_c + 2C)} \right\}. (28)
\]

The second equality follows from equation (15).

On the right hand side of the last equality, the first term within the brackets reflects the misvaluations of the fraction \( f^u \) of investors who ignore earnings entirely. This leads to direct underreaction to the higher date 1 earnings associated with higher date 1 accruals \( a_1 \), and underreaction to the implications of a higher accrual for future earnings.
Since these investors ignore the implications of higher accruals and resulting earnings, they underreact. So if \( f^e = 0 \), the future return is increasing with the deviation of accruals from their mean.

The other term reflects the fraction \( f^e \) of investors who attend to earnings but not to how earnings is divided between accruals and cash flows. The expected error made by these investors given current-period accruals is the difference between the correct forecast given current-period accruals, and the forecast of future earnings based on the effect of current-period accruals on current-period earnings. If \( |\text{corr}(a_1, c_1)| \neq 1 \), the \( f^e \) term in (28) is negative since \( \beta_1 > \beta_2 \) and \( C^2 - V_a V_c = V_a V_c (\text{corr}^2(a_1, c_1) - 1) < 0 \).

The model is symmetric with respect to \( c_1 \) and \( a_1 \), except for the assumption that \( \beta_1 > \beta_2 \) in (11). Thus, replacing \( a_1 \) with \( c_1 \), \( c_1 \) with \( a_1 \), \( \beta_2 \) with \( \beta_1 \), and \( \beta_1 \) with \( \beta_2 \) in (28), we obtain the expected return conditional on the date 1 cash flow,

\[
E[S_2 - S_1 | c_1] = (c_1 - \bar{c}_1) \left\{ f^n \left[ (1 + \beta_1) + (1 + \beta_2) \left( \frac{C}{V_c} \right) \right] + f^e (\beta_1 - \beta_2) \frac{-C^2 + V_a V_c}{V_c (V_a + V_c + 2C)} \right\}
\]

(29)

The \( f^e \) term in (29) contributes to underreaction to cash flow since \( \beta_1 > \beta_2 \) and \( -C^2 + V_a V_c \geq 0 \). Since the \( f^n \) term does so as well (since, by (16), \( C + V_c > 0 \), and \( \beta_1 > \beta_2 \)), overall there is underreaction to cash flow.

The above analysis is summarized in the following proposition.

**Proposition 2** Suppose that at date 1 fraction \( f^n \) of investors do not attend to any of earnings \( c_1 \), cash flow \( c_1 \) and accruals \( a_1 \); that fraction \( f^e \) attend to earnings but not to its decomposition between cash flow and accruals; and that the remaining fraction \( f^a = 1 - f^n - f^e \) attend to all publicly available information. Then:

1. A firm with date 1 cash flows that are above (below) their unconditional mean is undervalued (overvalued), and subsequently on average earns positive (negative) abnormal returns.

2. If \( f^n \) is sufficiently small relative to \( f^e \), a firm with date 1 accruals that are above (below) their unconditional mean is overvalued (undervalued) and subsequently on average earns negative (positive) abnormal returns.

Part 1 indicates that under our assumption that cash flow is a more favorable incremental predictor of future earnings than is accruals (\( \beta_1 > \beta_2 \) in equation (11)), there is under-reaction to cash flow. As discussed after footnote 11, this assumption is equivalent to the assumption that after controlling for current-period earnings, accruals are a negative incremental predictor of future earnings.
Part 2 indicates that if enough attention is paid to earnings, there is overreaction to the accruals component of earnings. One case in which the condition for Part 2 applies is $f_u = 0$ (all investors attend to earnings). The ability of accruals to forecast returns in this case reflects the fact, as discussed at the end of Subsection 3.1, that $\beta_1 > \beta_{c_e 1} > \beta_2$. In other words, high current-period cash flow is a more favorable forecaster of future earnings than is high current-period earnings, which is a more favorable forecaster than high current-period accruals. Those investors who focus on earnings without attention to its components do not take into account that for a given level of earnings, the true expectation of future earnings is higher when cash flow is high than when it is low. In consequence, such investors undervalue firms with high cash flow and overvalue firms with low cash flow. Such misvaluation is subsequently corrected, causing abnormal returns.

Similarly, those investors who focus on earnings without attending to its components do not take into account that for a given level of earnings, the true expectation of future earnings is higher when accruals are low than when they are high. In consequence, such investors overvalue firms with high accruals and undervalue firms with low accruals, leading to subsequent abnormal returns.

When $f_u > 0$, so that some investors neglect earnings, there is general underreaction to earnings. Since cash flow is a more favorable predictor of earnings than accruals, $f_u > 0$ further implies underreaction to cash flow relative to accruals. This reinforces the underreaction to cash flows, which provides the intuition for the cash flow anomaly asserted in Part 1. Even though investors overreact to accruals relative to cash flow, if $f_u$ is sufficiently large relative to $f^e$, there is underreaction even to the accruals component of earnings. This explains why the Part 2 conclusion that high accruals are associated with low returns requires that the fraction of investors who attend only to earnings (and not its components) be sufficiently large relative to the fraction of investors who neglect earnings entirely.

Thus, the analysis is consistent with the accrual anomaly, but the condition under which it applies is different from that proposed by Sloan (1996), that accruals be a less favorable predictor than cash flows of future earnings ($\beta_1 > \beta_2$). Here $\beta_1 > \beta_2$ is a necessary but not sufficient condition for the accrual anomaly to apply. Thus, an empirical implication of Proposition 2 is that the accrual anomaly can vanish or even reverse even for sets of firms for which cash flow is a more positive predictor than accruals of future earnings, if the difference $\beta_1 - \beta_2$ is not too large and/or $f^e$ is not sufficiently large compared to $f_u$. A further empirical implication is that the positive
relation of cash flows to subsequent returns should be more robust across different firms, trading venues, and time periods than the negative relation of accruals to subsequent returns. Specifically, the accrual anomaly could reverse for firms that have greater neglect of earnings (higher \( f'^a \)). Consistent with this implication, Pincus, Rajgopal, and Venkatachalam (2007) find that the cash flow anomaly is reliably present in a much wider set of countries (12 out of 20) than the accrual anomaly (3, including the US).

Proposition 1 indicates that a proxy for \( f'^a \) that can be used for further empirical testing is the strength of the firm’s post-earnings-announcement drift. Alternatively, other proxies for the sensitivity of a firm’s misvaluation to its earnings surprises could be used based upon contemporaneous misvaluation proxies. Either type of proxy for \( f'^a \) can be used to test Part 2 of Proposition 2 as well.

### 3.4 The Relative Strength of Accruals- and Cash Flow-Based Return Predictability

Our analysis provides insight regarding the strength of the cash flow and accruals anomalies. Let \( b_a \) be the slope coefficient in the relation between accruals surprises \( a_1 - \bar{a}_1 \) and subsequent abnormal stock returns, and let \( b_c \) be the slope coefficient in the relation between cash flow surprises \( c_1 - \bar{c}_1 \) and subsequent abnormal stock returns. Then by equations (28) and (29),

\[
\begin{align*}
    b_a &= f'^a \left[ (1 + \beta_1) \left( \frac{C}{V_a} \right) + (1 + \beta_2) \right] + f^e (\beta_1 - \beta_2) \frac{C^2 - V_a V_c}{V_a (V_a + V_c + 2C)} \\
    b_c &= f'^a \left[ (1 + \beta_1) + (1 + \beta_2) \left( \frac{C}{V_c} \right) \right] + f^e (\beta_1 - \beta_2) \frac{-C^2 + V_a V_c}{V_c (V_a + V_c + 2C)}. \quad (30)
\end{align*}
\]

#### 3.4.1 Effects of Variations in Attention on the Accruals- and Cash Flow-Anomalies

Since the \( f'^a \) terms are positive in both cash flow and accruals equations in (30), while the \( f^e \) term is positive in the cash flow equation and negative in the accruals equation, we have:

**Proposition 3**

1. The strength of the relation between cash flow and subsequent abnormal returns, \( b_c \), is increasing in \( f'^a \) and \( f^e \).
2. If there is an accrual anomaly \( (b_a < 0) \), the strength of the relation between accruals and subsequent abnormal returns, \( |b_a| = -b_a \), is decreasing in \( f_u \) and increasing in \( f_e \).

3. A shift in investor probability mass from attending only to earnings to neglecting both earnings and its components (increasing \( f_u \) and decreasing \( f_e \) by the same amount) intensifies the cash flow anomaly.

Part 1 reflects the fact that neglect of earnings and neglect of earning components both reinforce the cash flow anomaly. A shift in investor probability mass from attending to both earnings and its components to either neglecting earnings components (increasing \( f_e \)) or to neglecting earnings as well as its components (increasing \( f_u \)) intensifies the cash flow anomaly. An untested empirical implication of Part 1 is that the cash flow anomaly will be stronger when attention declines. Proxies for attention or inattention from the literature such as trading volume, the number of distracting news events, and Friday announcements could be used to test this implication.

Intuitively, if the fraction of investors that attends to earnings but neglects accruals and cash flow, \( f_e \), increases (at the expense of fully attentive investors), then this neglect favors underreaction to cash flows and overreaction to accruals. Empirically, the complexity or length of the firm’s annual report is a proxy for the incremental cost of attending to earnings components (where the earnings level itself can be attended to without delving into these details). It is therefore a possible proxy for variation in \( f_e \), holding \( f_u \) constant. If the fraction of investors that neglects earnings, \( f_u \), increases, the underreaction to cash flows is reinforced, whereas the overreaction to accruals is weakened. As mentioned in the discussion following Proposition 2, the strength of post-earnings announcement drift can be used as an empirical proxy for \( f_u \) for a stock or set of stocks.

Another attention proxy is the share ownership of institutional versus individual investors. If institutional investors are attentive to accruals, cash flows, and earnings, then high institutional ownership and low individual ownership should be associated with low \( f_u \) (neglect of both earnings and accruals) and \( f_e \) (neglect of earnings). Low \( f_u \) and low \( f_e \) both weaken the cash flow anomaly. Furthermore, if institutional ownership is high enough to drive both \( f_u \) and \( f_e \) toward zero (so that few investors neglect the split of earnings between cash flow and accruals), the accrual anomaly should become arbitrarily weak. Empirically, Collins, Gong, and Hribar (2003) find that the accrual anomaly is stronger in stocks with lower ownership by active institutional investors.
With regard to Part 3, a shift in investor probability mass from attending only to earnings to neglecting earnings as well has two opposing effects on the cash flow anomaly, since both $f^e$ and $f^u$ terms contribute to underreaction to cash flow. However, decreasing $f^e$ and increasing $f^u$ by the same amount $\Delta$ increases $b_c$ by

$$\Delta \left( V_c + C \right) \left[ \beta_1 (V_c + C) + \beta_2 (V_a + C) + (V_a + V_c + 2C) \right] \frac{1}{V_c (V_a + V_c + 2C)} > 0.$$

When cash flow is high, for example, investors who attend only to earnings (fraction $f^e$) neglect the cash flow but draw a somewhat positive inference from the (on average) high earnings. However, investors who do not attend to earnings (fraction $f^u$) do not draw any inference from the cash flow or the earnings. Therefore, decreasing $f^e$ and increasing $f^u$ by the same amount strengthens the cash flow anomaly.

### 3.4.2 Effects of Earnings Quality on Accrual and Cash Flow Anomalies

It has often been alleged that earnings management reduces the quality of earnings, in the sense that earnings is a less accurate indicator of long-run firm performance. In our setting such a reduction in earnings quality associated with accruals would be reflected in a low $\beta_2$ in equation (11), so that current-period level of accruals becomes a much less favorable predictor of future earnings than current-period cash flow.

We now consider the implications of lower earnings quality in this sense for the strengths of the accrual and cash flow anomalies. Differentiating $b_a$ in (30) with respect to $\beta_2$ yields

$$\frac{\partial b_a}{\partial \beta_2} = f^u + f^e \left[ \frac{V_a V_c - C^2}{V_a (V_a + V_c + 2C)} \right] > 0 \quad (32)$$

where the final inequality holds because $V_a V_c \geq C^2$. This derivative indicates that if there is an accrual anomaly ($b_a < 0$), then when earnings quality increases (accruals becomes a more favorable incremental predictor of future earnings), the accrual anomaly becomes weaker (the negative incremental coefficient increases, moving it closer to zero). If there is no accrual anomaly or a reverse accruals effect ($b_a \geq 0$), then the effect becomes even more reversed.

Differentiating $b_c$ in (30) with respect to $\beta_2$ yields

$$\frac{\partial b_c}{\partial \beta_2} = f^u \left( \frac{C}{V_c} \right) + f^e \left[ \frac{C^2 - V_a V_c}{V_c (V_a + V_c + 2C)} \right] \quad (33)$$

The $f^u$ term is positive, but the $f^e$ term is not since $C^2 \leq V_a V_c$. This indicates that the cash flow anomaly could be either strengthened or weakened by an increase in earnings quality.
Proposition 4 Consider an increase in accruals quality that increases $\beta_2$, i.e., makes the incremental forecasting power of current-period accruals for future earnings closer to that of current-period cash flows. Then an increase in earnings quality:

1. Weakens the accrual anomaly; and

2. Can either weaken or strengthen the cash flow anomaly.

The $\beta_2$ parameter can be estimated by running the regression in (11). Therefore the prediction about the relation of accruals quality to the accrual anomaly is empirically testable.

Intuitively, in Part 1 the accrual anomaly is driven by the low quality of the accruals component of earnings, so if this component becomes a more favorable forecaster of earnings, a strong investor reaction to accruals becomes a less severe mistake. This is reinforced by a subtler effect. When earnings quality increases owing to an increase in $\beta_2$, the persistence of earnings, $\beta_{et}e_1$, increases (see equation (15)). Investor neglect of the implications of current-period earnings for future earnings (as reflected in $f^u$) therefore causes greater underreaction to earnings as the persistence of earnings increases (Proposition 1, Part 3). This greater general underreaction further weakens the overreaction to accruals.

In Part 2, the intuition for the effect of a change in accruals quality on the cash flow anomaly reflects the same two effects. On the one hand, higher accruals quality eases the misperceptions associated with the neglect of accruals and cash flow information. This is reflect in the negative $f^e$ term in (33), indicating a weakening of the cash flow anomaly.

On the other hand, more subtly, higher accruals quality makes high current period earnings a more favorable forecaster of future earnings. Therefore, the neglect of current period earnings, as reflected in the $f^u$ term in (33), tends to increase underreaction. This reinforces the underreaction to cash flows, strengthening the cash flow anomaly. Thus, the overall effect of improving accruals quality on the cash flow anomaly is ambiguous.

Testing of Proposition 4 requires controlling for other exogenous parameters such as the variances of accruals and cash flows. This is important, for example, in evaluating differences in earnings quality that derive from differing earnings management practices, since a firm that manages earnings heavily may have greater variability in accruals. We next examine the effects on anomalies of varying volatilities and correlations as well.
3.4.3 Effects of Accruals and Cash Flow Volatilities and Correlations on Accruals- and Cash Flow- Anomalies

The $f^e$ term, which reflects investor inattention to the division of earnings between cash flow and accruals, is the sole source of mispricing when $f^n = 0$. We define $b^0_a$ and $b^0_c$ as the values of the slope coefficients that come from neglect of earnings components, not from neglect of earnings itself (i.e., the $f^n$ term),

$$b^0_a = f^e (\beta_1 - \beta_2) \left[ \frac{C^2 - V_a V_c}{V_a (V_a + V_c + 2C)} \right]$$
$$b^0_c = f^e (\beta_1 - \beta_2) \left[ -\frac{C^2 + V_a V_c}{V_c (V_a + V_c + 2C)} \right].$$

(34)

It follows that the relative size of the slope coefficients in the accruals regression and in the cash flow regression is determined by the ratio of the variances of accruals and of cash flows.

$$\left| \frac{b^0_a}{b^0_c} \right| = -\frac{b^0_a}{b^0_c} = \frac{V_c}{V_a}.$$  (35)

For example, if $f^n = 0$ and $V_c = V_a$, then $b^0_a = -b^0_c$, so that the accruals and cash flow anomalies are equally strong.

So far in this discussion we have assumed that $f^n = 0$, which eliminates post-earnings announcement drift by Proposition 1. More generally, when $f^n > 0$ in (30), the slope coefficients $b_a$ and $b_c$ are larger than the corresponding $b^0_a$ and $b^0_c$. Since $b^0_c > 0$, it follows that if $f^n > 0$ and there is an accrual anomaly ($b_a < 0$),

$$-\frac{b_a}{b_c} = -\frac{f^n (1 + \beta_1) \left( \frac{C}{V_c} \right) + (1 + \beta_2)}{f^n (1 + \beta_1) + (1 + \beta_2) \left( \frac{C}{V_c} \right)} - \frac{b^0_a}{b^0_c} < -\frac{b^0_a}{b^0_c} = \frac{V_c}{V_a},$$

i.e., the relative strength of the cash flow effect is greater than what is implied by the variance ratio. For example, if cash flow and accruals are equally variable, then (35) indicates an equally strong cash flow and accrual effect. But if some investors ignore earnings, so that $f^n > 0$, by (36) the cash flow effect is stronger than the accruals effect. Thus, although the empirical literature has focused more on accruals than cash flow, our analysis predicts that the cash flow effect can be stronger.

This analysis is summarized in the following proposition.
Proposition 5

1. If investors always attend to earnings, so that \( f^u = 0 \), then the ratio of the absolute slopes in the univariate regressions of subsequent returns on accruals and on cash flows, \(-b_a^0/b_c^0\), is equal to the ratio of the variance of cash flow to the variance of accruals, \( V_c/V_a \).

2. If some investors neglect earnings, so that \( f^u > 0 \), and if there is an accrual anomaly (\( b_a < 0 \)), then the ratio of the absolute slopes in the univariate regressions of subsequent returns on accruals and on cash flows, \(-b_a/b_c\), is less than the ratio of the variance of cash flow to the variance of accruals, \( V_c/V_a \).

Intuitively, consider an investor who focuses on current-period earnings in forecasting future earnings. For simplicity, suppose that the covariance of current-period accruals and current-period cash flow is zero. High accruals increase current-period earnings, making the investor more optimistic about future earnings. The investor on average becomes too optimistic, because he forecasts high future cash flow based on the unconditional relation between current-period earnings and future earnings. Thus, the investor’s overoptimism will be greater to the extent that an extra unit of earnings is on average associated with higher future earnings more strongly than is an extra unit of accruals.

When accruals are highly variable relative to cash flow, most variation in earnings comes from variations in accruals. In this case, the forecasts of investors with limited attention who condition on earnings are almost equivalent to fully attentive forecasts based on accruals. High accruals bring about little over-optimism, so the slope \( b_a^0 \) approaches zero. In contrast, when the variance of accruals is low relative to the variance of cash flow, the forecasts of investors with limited attention who condition on a given deviation of earnings from mean earnings are similar to the forecasts that would be made based on a comparable deviation of cash flow from its mean. If in fact it is the level of accruals, not cash flow, that is high, then this misattribution brings about a great deal of overoptimism.

A higher value of \( b_a \) means that it is less negative, a weaker accruals effect. So Part 2 indicates that the cash flow effect can be stronger than the accruals effect even if the relative variability is equal.

One reason why a firm’s accruals might be highly variable is if it engages in extensive earnings management. If so, then accruals may be a less favorable predictor of future
earnings (lower $\beta_2$). It is therefore striking that the prediction of Proposition 5 is independent of the forecasting powers of cash flow or accruals, $\beta_1$ or $\beta_2$, for future earnings.\footnote{Intuitively, the prediction in (35) is about the \textit{relative} sizes of these anomalies. If $\beta_2 \ll \beta_1$, so that the neglect of accruals makes a big difference for value, the accrual anomaly will be stronger (as indicated by equation (32) in the derivation of Proposition 4). However, in such a case the cash flow anomaly also tends to be stronger (as indicated by the $f^e$ term in equation (33)).} The independence of the prediction of Proposition 5 from $\beta_1$ and $\beta_2$ simplifies empirical testing.

By Proposition 5, a higher variance of accruals relative to cash flow implies a weaker accruals effect. With regard to Part 1, if the accruals or cash flow effects are empirically evaluated using hedge profits based upon fractile sorts, the spread in the independent variables is affected by their variances. Under normality, among a set of ex ante identical firms, a higher variance of accruals implies a higher mean high-low fractile spread in accruals. Since the relative weakness in the slope is proportional to the relative variances, whereas the fractile spread increases only as the square root of the variances, there is partial cancellation. Overall, Proposition 5 indicates that for fractile spreads, among a set of ex ante identical firms that obtain different realizations of accruals and cash flows, the ratio of the hedge profits based upon accruals versus those based upon cash flow will be inversely proportional to the ratio of the standard deviations (instead of variances) of accruals and of cash flows.

We next evaluate the effects of variances on each of the slope coefficients separately.

**Proposition 6** If investors are fully attentive to earnings ($f^u = 0$), but some do not attend to accruals and cash flow separately ($f^e > 0$), then the slope coefficient $b^e_\theta$ is increasing, and the slope coefficient $b^c_\epsilon$ is decreasing, with $\rho$. In other words, as $\rho$ increases, both the accrual and cash flow anomalies become weaker.

The proof is in the Appendix. Empirically, the parameter $\rho$ can be estimated using accrual and cash flow data.

The intuition is that both accruals and cash flow effects derive from investors neglecting the differing implications for long-run earnings of current-period cash flow versus accruals. When accruals and cash flow are highly correlated with each other, each is highly correlated with earnings, so that ignoring the distinction between different earnings components is on average of little consequence.

If $f^u > 0$ so that some investors neglect earnings, not just its components, higher $\rho$
further weakens the accruals effect (increases the negative slope $b^0_a$):

$$b_a = f^u \left[(1 + \beta_1)\rho\sqrt{\frac{V_c}{V_a}} + (1 + \beta_2)\right] + b^0_a,$$

so

$$\frac{\partial b_a}{\partial \rho} = f^u(1 + \beta_1)\sqrt{\frac{V_c}{V_a}} + \frac{\partial b^0_a}{\partial \rho}.$$  (38)

Since the first term is positive, it follows that $\frac{\partial b_a}{\partial \rho} > 0$. The following proposition summarizes this finding:

**Proposition 7** The slope coefficient of future returns on accruals $b_a$ is increasing with the correlation $\rho$ between cash flows and accruals.

Intuitively, the first term reflects the fact that when the correlation between accruals and cash flow is higher, high accruals tend to be more positively (or less negatively) associated with high cash flow. Since the market underreacts to earnings ($f^u > 0$), there is underreaction to the higher accruals, which tends to oppose the pure accruals effect (make $b_a$ less negative).

The sign of $\frac{\partial b_c}{\partial \rho}$ is ambiguous. As shown earlier, the $f^e$ term of $b_c$, $b^e_c$, is decreasing with the correlation $\rho$ between accruals and cash flow. On the other hand, neglect of earnings ($f^u > 0$) intensifies the cash flow when $\rho$ is higher, as high cash flow is more positively associated with high accruals, both of which are positive predictors of date 2 earnings. Thus, the net effect of a higher $\rho$ on $b_c$ depends on parameter values.

### 3.5 The Profit Anomaly and Post-Earnings Announcement Drift

Balakrishnan, Bartov, and Faurel (2010) provide evidence that the level of (as distinguished from the surprise in) earnings predicts future stock returns, and that this effect is not subsumed by post-earnings announcement drift. There is also evidence that stocks with high earnings on average earn higher returns (Chen, Novy-Marx, and Zhang (2010)). In this subsection we show that investor inattention can explain the profit anomaly, the return predictability based on the level of earnings, and that, consistent with existing evidence, the effect is incremental to the post earnings announcement drift.

To illustrate the key intuition, we simplify the model by assuming that there are only two types of investors, those who attend to earnings alone and those who do not. In other words, we set the fraction who attend to earnings components, $f^u$, to zero. We focus here on the issue of return predictability based upon earnings or on earnings surprises,
not on earnings components. At some point prior to the date 1 earnings announcement, both groups of investors form their expectations $f_0$ of the upcoming earnings $e_1$ based upon available date 0 information. We can view $f_0$ as either an analyst forecast, or the previous level of earnings. After the date 1 earnings announcement, a fraction $\lambda$ of investors are inattentive and stick to their original beliefs, and a fraction $1 - \lambda$ of investors use earnings information to update their beliefs. Here $\lambda = f^u$ from before; we rename it here to avoid confusion with our terminology for the analyst forecast, $f_0$.

As before, we assume that the date 2 stock value is the sum of the date 1 earnings $e_1$ and the date 2 earnings $e_2$,

$$S_2 = e_1 + e_2,$$

where

$$e_2 = \bar{e}_2 + \beta_e (e_1 - \bar{e}_1) + \delta_e. \quad (40)$$

The date 1 earnings can be decomposed into the forecast $f_0$ and additional information $\varepsilon$,

$$e_1 = f_0 + \varepsilon, \quad (41)$$

where $f_0 \sim N(f_0, V_f)$ and $\text{cov}(f_0, \varepsilon) = 0$. We assume that forecasts are on average unbiased ($\bar{e}_1 = \bar{f}_0$).\footnote{In reality analyst forecasts are biased, but unbiasedness is a simple benchmark and is not crucial for our results.} Date 0 earnings forecasts of $e_1$ are made prior to the arrival of $e_1$, so we assume that they are less informative than $e_1$ itself about date 2 earnings. So comparing $\beta_f$ in the predictive regression

$$e_2 = \bar{e}_2 + \beta_f (f_0 - \bar{f}_0) + \delta_f, \quad (42)$$

with $\beta_e$ in equation (40), it follows that $\beta_e > \beta_f \geq 0$. Furthermore, the dominance of $e_1$ over $f_0$ as a predictor of future earnings implies that a bivariate regression of $e_2$ on $e_1$ and $f_0$ reduces to a univariate regression on $e_1$. So the expectations of attentive and inattentive investors are

$$E[S_2|e_1, f_0] = E[S_2|e_1] = E[S_2] + (1 + \beta_e)(e_1 - \bar{e}_1),$$

$$E[S_2|f_0] = E[S_2] + (1 + \beta_f)(f_0 - \bar{f}_0). \quad (43)$$

As before (see (10)), the date 1 stock price is a weighted average of the beliefs of attentive and inattentive investors. Recalling that $\lambda$ denotes the fraction of investors who do not attend to earnings, the date 1 stock price can be written as

$$S_1 = E[S_2] + (1 - \lambda)(1 + \beta_e)(e_1 - \bar{e}_1) + \lambda(1 + \beta_f)(f_0 - \bar{f}_0) - \text{Var}(S_2)x_0, \quad (44)$$
and the change in stock price is

\[ S_2 - S_1 = \lambda(1 + \beta_e)(e_1 - \bar{e}_1) - \lambda(1 + \beta_f)(f_0 - \bar{f}_0) + A\text{var}(S_2)x_0 + \delta e. \]  

(45)

In this setting, it is not hard to derive the profit anomaly, the finding that the level of earnings is a positive predictor of future returns. However, given the well-known finding of post earnings announcement drift, it is more interesting to derive that profit is a predictor of future returns even after controlling for the earnings surprise. We therefore run a ‘horse race’ between the profit anomaly and post earnings announcement drift by regressing the stock price change on both the level of date 1 earnings \( e_1 \), and the earnings surprise \( e_1 - f_0 \):

\[ S_2 - S_1 = b_0 + b_1 e_1 + b_2 (e_1 - f_0) + \epsilon. \]  

(46)

The coefficient on the level of earnings, \( b_1 \), captures the profit anomaly, and the coefficient on the earnings surprise, \( b_2 \), captures post earnings announcement drift. The regression coefficients \( b_1 \) and \( b_2 \) are

\[ b_1 = \lambda(\beta_e - \beta_f) \]

\[ b_2 = \lambda(1 + \beta_f). \]  

(47)

Since \( \beta_e > \beta_f \geq 0 \), \( b_1 > 0 \) and \( b_2 > 0 \). Therefore there are both a profit anomaly and post-earnings announcement drift, and neither subsumes the other.

The results imply that both anomalies stem from investor inattention to earnings information (\( \lambda \)). The effect of attention toward earnings news can be viewed as containing two parts. First is a shift in the mean belief about the current level of earnings \( e_1 \) by the amount of the surprise (\( \varepsilon = e_1 - f_0 \)). Second is greater certainty about the current level of earnings (\( e_1 \) for sure instead of \( f_0 \) on average), which causes the use of a different coefficient when investors forecast future earnings based on the information they have (use of \( \beta_e \) instead of \( \beta_f \)). When there is inattention, the mean shift causes a misperception (and subsequent return predictability) induced by the earnings surprise, i.e., post-earnings announcement drift. The second effect, a failure to update sufficiently strongly for any given level of date 1 expected earnings, induces mispricing in relation to the level of earnings.\(^{17}\) Therefore, the earnings level predicts returns (the profit anomaly) even after controlling for the earnings surprise.

\(^{17}\)Holding earnings surprise \( \varepsilon \) constant, a unit increase in earnings level \( e_1 = f_0 + \varepsilon \) is equivalent to a unit increase in \( f_0 \). This increase in earnings level increases the expected stock price by \( 1 + \beta_f \) for inattentive investors, which is smaller than the increase in the expectation of attentive investors, \( 1 + \beta_e \). So there is underreaction to variations in the earnings level above and beyond the underreaction to the earnings surprise.
A further untested empirical implication is that the profit anomaly is stronger when the current level of earnings is a much better predictor of future earnings than are prior earnings expectations such as analyst forecasts and prior earnings (higher $\beta_e - \beta_f$). The following proposition summarizes these results.

**Proposition 8**

1. If fraction $\lambda > 0$ of investors ignore the information in current-period earnings and hold onto their beliefs based on prior earnings expectations, both the level of earnings and earnings surprise (earnings minus prior earnings expectations) predict future stock returns.

2. There is a stronger profit anomaly when current earnings is a much better predictor of future earnings than prior earnings expectations.

4 Trading Profits and Equilibrium Inattention

It could be argued that inattentive investors should withdraw from the market instead of trading; and that if they do trade, they will on average lose money, and in the long run, tend to vanish. However, in standard models of securities markets, individuals do find it optimal to trade despite having informational disadvantages relative to some subsets of other traders. Furthermore, attending to more information is costly, so those who devote more cognitive resources to an information signal (such as a given firm’s earnings, accruals or cash flow) need not do better overall. Attending to a given signal allows individuals to value the firm more accurately, at the cost of withdrawing attentional resources from some other activity. For example, attention demands time, which has a monetary opportunity cost. Since there is both a benefit and a cost to attending to a given public signal, in a long-run equilibrium in which wealths of investors shift over time, it does not follow that fully attentive investors dominate.

In this section we show that if attention is costly, there can be an equilibrium in which inattentive investors find it optimal to trade; and in which inattentive and atten tive investors on average achieve the same expected utility, and therefore coexist. We first show that investors who do not attend to all publicly available information can nevertheless earn positive expected trading profits. Next, when individuals decide to attend by balancing the cost of attending against the expected utility benefits from making better-informed decisions, we show that there can exist an equilibrium where a positive fraction of investors decide not to attend to some public information. Since the
attentive and inattentive achieve equal ex ante welfare, there is no reason to expect the inattentive to vanish in the long run.

4.1 Expected Trading Profits

Investors who neglect the information contained in earnings or its components make inferior trades, which raises the question of whether such behavior can survive in equilibrium. We now show that inattentive investors can still earn positive expected profits owing to noise trading. We consider investors who make rational plans when they allocate attention ex ante, foreseeing the fact that if they choose not to attend to an information item then their subsequent beliefs will be biased. We first calculate the true expected trading profits to investors as a function of whether they attend to information signals.

Using the formula for the optimal security position given by (4), the expected profit of an investor who ignores earnings and its components in forming expectations is

$$E[x^u(S_2 - S_1)] = E \left[ \frac{E[S_2] - S_1}{\text{Avar}(S_2)} (S_2 - S_1) \right].$$

(48)

For an investor who attends to earnings information only, her expected profit is

$$E[x^e(S_2 - S_1)] = E \left[ \frac{E[S_2|e_1] - S_1}{\text{Avar}(S_2)} (S_2 - S_1) \right].$$

(49)

**Proposition 9** When there is a large amount of noise trading ($V_x$ sufficiently large), inattentive investors earn positive expected profits.

The proof is in the Appendix. We find that uninformed investors (those who ignore earnings) can earn positive expected profits owing to noise trading. Both noise trading and the presence of uninformed investors positively contribute to the expected trading profits of investors who attend to earnings but ignore its components.

4.2 Expected Utility and Characterization of Equilibrium

We now evaluate the expected utility of each type of investor. Given the mean-variance utility function in equation (1), the consumption (2), and the demand for the risky security in equation (4), the expected utility of investor group $i$ is as follows.

$$E[U^i(C_2)] = E \left[ (W - k^i) + \frac{(E^i[S_2] - S_i)^2}{2\text{Avar}(S_2)} \right].$$

(50)
Assume that the cost $k^i$ of attention is equal to zero for uninformed investors ($k^u = 0$) and that the cost of attending to earnings only is less than that of attending to cash flow and accruals separately ($k^e < k^a$).

From multivariate normality, $E^i[S_2] - S_1$ is normally distributed with mean zero since $E[S_1] = E[S_2] = E[E^i[S_2]]$. Thus $(E^i[S_2] - S_1)^2$ follows a chi-square distribution. Therefore, the expected utility of each group of investors is

\[
E[U^u(C_2)] = W + \frac{\text{var}(S_1)}{2A\text{var}(S_2)},
\]

\[
E[U^e(C_2)] = W - k^e + \frac{\text{var}(E[S_2|c_1] - S_1)}{2A\text{var}(S_2)},
\]

\[
E[U^a(C_2)] = W - k^a + \frac{\text{var}(E[S_2|c_1, a_1] - S_1)}{2A\text{var}(S_2)}.
\] (51)

In an equilibrium where the fraction of investors in each category is non-zero, the ability of investors to pay to enter higher attention categories will cause expected utility to be equated across all three categories:

\[
\frac{\text{var}(S_1)}{2A\text{var}(S_2)} = -k^e + \frac{\text{var}(E[S_2|c_1] - S_1)}{2A\text{var}(S_2)} = -k^a + \frac{\text{var}(E[S_2|c_1, a_1] - S_1)}{2A\text{var}(S_2)}. \] (52)

A numerical example of such an equilibrium is provided by the case $V_c = V_a = V_\delta = V_x = 1$, $A = 1$, $\text{cov}(a_1, c_1) = 0.3$, $\beta_1 = 0.5$, $\beta_2 = 0.3$, $f^u = 0.6$, $f^e = f^a = 0.2$, $k^e = 0.0834$, and $k^a = 0.0841$.

**Proposition 10** When attention is costly and is allocated endogenously, under some parameter values there are equilibrium fractions of investors who choose to neglect earnings and its components, and earn the same expected utility as fully attentive investors.

Intuitively, owing to the cost of attention as reflected in $k^i$, investors who allocate greater or lesser attention to earnings or its components can coexist in the long run since more attentive investors incur higher opportunity costs of attention. These opportunity costs can be viewed as the benefit from allocating cognitive resources to other decision domains, such as personal life or other investment categories. Previous research that studies the optimization decision for allocating attention (Hirshleifer, Lim, and Teoh (2008), Gabaix and Laibson (2005), Peng (2005), and Iliev and Welch (2010)) has identified the importance of such opportunity costs of allocating attention to any given domain. Proposition 10 shows that owing to the tradeoff between making well-informed decisions based upon earnings information versus in other domains (as reflected in the cost of attention function), a lack of attention to earnings can persist in the long run.
5 Do Investors Fully Compensate for Limited Attention?

A key assumption of our model is that individuals with limited attention trade based upon their beliefs. As a result, limited attention affects the equilibrium price. Casual observation suggests that investors often do make trades based on beliefs that do not fully reflect publicly available information. However, an individual who understands that he is neglecting a relevant public information item in forming his beliefs is free to retreat into autarky, thereby deferring completely to the market price. Such withdrawal into passivity could make sense if an individual believes that the market value is highly accurate. However, in reality, inattentive investors are not necessarily sophisticated enough to realize that they should be passive.

For both psychological and economic reasons, we do not expect that individuals with limited attention would or even should always withdraw into passivity. Evidence from both psychology and market behavior indicates that individuals often fail to adjust appropriately for the fact that they have left some information unprocessed.

Such double neglect makes sense when processing power is limited. Attention and mental processing are needed to adjust for the fact that information items that are not being attended to in detail may affect the reliability of the judgments being formed. An extra cognitive step is required, after forming a valuation of the firm, to recognize that the market may be using useful information that was not a part of the information used by the individual. Indeed, the cognitive cost of calculating how to adjust for the fact that a relevant information item has been neglected can be as great or greater than the cognitive cost of simply processing the information item. So in many cases, ignoring the information item and failing to adjust for the fact that the item has been neglected go hand in hand.

Evidence that people both neglect signals and do not adjust for the fact that they are neglecting them is provided by studies that show that the form of presentation of information affects individuals judgments and decisions (see, e.g., Slovic (1972), Payne, Bettman, and Johnson (1993), and the review of Libby, Bloomfield, and Nelson (2002)). Experimental studies have found that different presentations of equivalent information about a firm affect the valuations and trades of investors and experienced financial analysts. In principle, if an investor understood that owing to limited attention certain

\[ \text{Presentation effects have been found in the context of recognition versus disclosure of pension liabilities (Harper, Mister, and Strawser (1987)), classification of the same hybrid financial instrument} \]
formats were hard to process, the investor could self-debias by, for example, mentally rearranging the format of presentation. However, such rearrangement itself requires mental processing.

Furthermore, experimental research has found that the presentation of one-sided arguments and evidence to subjects (call ‘jurors’) asked to judge a legal dispute were biased in favor of the side they heard (Brenner, Koehler, and Tversky (1996)). According to the authors, “The results indicate that people do not compensate sufficiently for missing information even when it is painfully obvious that the information available to them is incomplete.

Why don’t individuals fully adjust beliefs about their own precisions to take into account the set of cues they are neglecting? Because thinking about how the neglect of a cue creates bias or reduces precision requires an extra layer of cognitive processing. Just as cognitive resource constraints cause a cue to be neglected, they often also cause the individual to fail to take the extra step of assessing the consequences of that neglect. Even if on average the individual is correct, in those cases in which the cues he neglects are especially important, he will tend to overestimate the precision of his beliefs; and when the cues he neglects are minor, he will underestimate his precision. Such miscalibration encourages individuals who have neglected important cues to trade and influence price.

Our model shares with informal behavioral explanations for anomalies the assumption that investors who neglect public signals such as current-period earnings are important for price setting. For example, the seasonal-random-walk explanation for post-earnings-announcement drift of Bernard and Thomas (1989) assumes that investors who focus on year-ago earnings neglect the latest earnings surprise. There are a number of reasons why an individual who neglects important information such as the implications of current-period earnings would still trade actively. An investor who has a perspective about the firm’s long-term business strategy and competitive environment may trade based upon his valuation target without updating his forecasts of future earnings in response to every earnings report.19

19The tendency to hold a fixed valuation in mind may be reinforced by the practice of analysts of...
There is evidence that some individual investors are net providers of liquidity to the market, absorbing the demand by institutions for trading immediacy. If institutions (perhaps for agency reasons) sometimes make uninformed trades, a simple contrarian trading strategy will in general be profitable. Indeed, the evidence indicates that contrarian trading allows U.S. individual investors to earn positive excess returns in the month after their trades (Kaniel, Saar, and Titman (2008)).

A naive contrarian strategy of trading in opposition to market price movements without regard to their source will on average induce trading against earnings announcements. If institutions are willing to pay for trading immediacy, a naive contrarian strategy can be profitable on average. However, such a strategy will on average be unprofitable after an earnings announcement, because the price move is being triggered by actual news rather than by another investor’s demand for trading immediacy. Thus, limited attention on the part of investors traders who are following a generally-reasonable contrarian strategy can induce unprofitable contrarian trading in response to earnings announcements.

A further reason why an investor with limited attention would actively trade as a contrarian to earnings news is that analysts on average do not fully adjust their earnings forecasts in response to earnings announcements. Owing to limited attention, some investors may rely on analyst earnings forecasts without performing the additional cognitive processing needed to adjust for this state-contingent bias in analyst forecasts. The announcing ‘target’ prices below which the stock is recommended as a good buy. An investor who follows such a recommendation without adjusting the target to reflect further news about earnings fits the analysis. Plausible-sounding media sound bites can reinforce the tendency of investors with limited attention to trade as contrarians. There is much editorializing in the business media against ‘obsession’ with the short-term, and in favor of ‘buying on the dips.’ An investor can excuse a failure to update valuations frequently as ‘focusing on the long-term’ rather than ‘myopically’ focusing on quarterly earnings. A naive application of sound bites with high availability in the mass media is itself a possible consequence of limited attention.

For example, in the U.S. there is evidence that individuals as a group trade as contrarians to the previous week’s stock return (Kaniel, Saar, and Titman (2008)). A similar finding has been documented among Finnish and Korean individual traders (Grinblatt and Keloharju (2000), Choe, Kho, and Stulz (1999)).

For example, a limit order is one means by which an investor can implement a naive contrarian strategy. Since limit orders are triggered by moves in market price, placing such an order allows the investor to trade against market price moves, providing liquidity to the market, without continuously monitoring prices.

This phenomenon is well-documented (e.g., Abarbanell and Bernard (1992) and Zhang (2008)), and has theoretical motivation. Models of analyst reputation predict this phenomenon because an analyst who has a prior forecast tends to hold too strongly to that forecast in response to a public signal in order to maintain a reputation for having highly accurate private information. A general model of analyst neglect of relevant information is provided by Trueman (1994).
beliefs of such investors will therefore underreact to earnings surprises, just as analyst forecasts do.\textsuperscript{23} Consistent with this interpretation, there is evidence that firms that are followed by analysts whose forecasts are more responsive to earnings announcements have less post-earnings-announcement drift (Zhang (2008)).

Furthermore, the intuition from standard securities market models suggests that even an investor who is aware of his own limited attention and rationally adjusts for it should trade based upon his expectations rather than withdrawing into autarky. In Section 4, we show that investors who ignore earnings can profit owing to the presence of noise/liquidity traders. Standard models of information and securities markets such as Grossman and Stiglitz (1976) provide the insight that, owing to liquidity or noise trading, prices aggregate information imperfectly, and in equilibrium contain noise-induced mispricing. As a result, even an informationally disadvantaged individual should trade based upon his beliefs rather than deferring completely to market price.\textsuperscript{24} In a setting with limited attention, an isomorphic situation arises in which a set of investors has a kind of pseudo-private information. Attending to a public signal that some other investors neglect is analogous to observing a private signal. Neglecting a signal that other individuals attend to is akin to being uninformed about that signal. However, in standard models of information and securities markets, individuals with informational advantages and disadvantages all trade based upon their beliefs, and in so doing profit at the expense of liquidity traders.

Indeed, securities market models that are ostensibly about private information are often interpreted in ways that are more consistent with different individuals processing publicly available information differently. Applied discussions based on these models do not usually interpret private information as being limited to inside information (e.g., information obtained by rummaging through a firm's trash, eavesdropping on cell phone conversations, or persuading employees to leak information). Instead, investors become privately informed through some unspecified process—presumably involving analysis of publicly available signals such as financial statements and media reports. This broader interpretation of Grossman and Stiglitz (1976) implicitly involves trade based upon processing versus neglecting publicly available information.

\textsuperscript{23}A similar argument applies to investors who neglect accruals. Such investors may form beliefs based upon analysts' earnings forecasts, which underutilize the information contained in accruals (Teoh and Wong (2002)).

\textsuperscript{24}In forming his beliefs, an investor combines any private signals he may possess (none, for an uninformed trader), together with the information implicit in market price. The weight on his own signals and prior is positive because market price is noisy. The noise in market price derives both from liquidity shocks and, if there are multiple private signals, from confounding between signals.
In order to develop our implications in the simplest and most parsimonious possible way, we assume that investors with limited attention do not incorporate market price into their beliefs. In reality investors with limited attention probably do sometimes fail to process the information about future earnings implicit in market price. However, qualitatively identical results to those in our model could be derived in a setting in which investors with limited attention draw rational inferences from market price, and in which liquidity trading makes prices noisy. In such a setting liquidity trading causes investors rationally to place less weight upon market price in updating beliefs. Thus, such a setting would endogenize the conclusion that individuals with limited attention trade actively based upon their beliefs instead of deferring completely to market price. Such an analysis would have more algebraic details, but conceptually would be essentially equivalent to our model.

There is other evidence that limited attention affects capital markets; indeed, Daniel, Hirshleifer, and Teoh (2002) argue that limited attention may underlie a wide range of anomalous patterns in securities market trading and prices. Many short-horizon event studies confirm that stock markets react immediately to relevant news. Long-horizon event studies provide evidence suggesting that there is underreaction to various kinds of public news events (see, e.g., the review of Hirshleifer (2001)). However, there has been a great deal of debate as to the appropriate methodology for testing market efficiency using long-run abnormal returns. There is also evidence suggesting that investors’ and analysts’ assessments are influenced by the format and salience with which public signals are presented (see the discussion in Section 5, and Schrand and Walther (2000)).

Although this paper uses a pure limited attention approach, psychological studies also indicate that individuals tend to be overconfident about the precision of their beliefs. Overconfidence can reinforce the tendency for individuals who fail to process an information item to also fail to adjust for this neglect. An individual who overconfidently thinks that he has already taken into account the most important information would

25In an experimental setting, Gillette et al (1999) document investor misreactions to public information arrival. Perhaps the most striking indication of limited attention in public markets is that stock prices react to news that is already public information (Huberman and Regev (2001), and Ho and Michaely (1988)), and even to confusions in ticker symbols between stocks (Rashes (2001)). More broadly, Hong, Torous, and Vallanov (2007) report evidence that industry stock returns lead aggregate market returns, potentially consistent with gradual diffusion of information about fundamentals across markets. Hou and Moskowitz (2005) provide a measure of investor neglect of a stock, the lag in the relation between the return on the overall market and the stock’s return. They find that stocks with long delay (which can be viewed as low-attention stocks) have stronger post-earnings announcement drift.
underestimate the urgency of working hard to adjust for his neglect of other relevant signals.

It can be argued that arbitrageurs such as hedge funds will eliminate almost all of the mispricing created by limited attention. However, a literature in behavioral finance and accounting has argued that arbitrage is limited by risk-bearing capacity and market frictions; see, e.g., Shleifer and Vishny (1997), Hirshleifer (2001), and Lee (2001). We would argue that an important constraint on arbitrage is the aggregate attentional capacity in the market (see, e.g., Peng (2005)). Attention can be leveraged by such means as computers and financial intermediation, but even large institutions face tradeoffs in allocating their attentional resources.

6 Conclusion

This paper offers a model of stock market misreactions to earnings-related information based upon limited investor attention. It is motivated by evidence of stock return predictability based upon earnings surprises and the level of earnings, accruals, and cash flows, including the striking fact that there are both under-reactions and over-reactions to different kinds of earnings-related information. To understand the sources of these return anomalies, it is important to be parsimonious in assumptions about psychological biases, to avoid overfitting the data. We therefore explore the consequences of a single psychological constraint, limited attention, which we use to derive a rich set of untested empirical implications.

We assume that some investors neglect information contained in the latest earnings surprise, and that some investors further neglect the information contained in accruals and cash flow. In equilibrium, there is underreaction to earnings surprises, because some investors do not attend to the newly-arriving earnings news. Cash flows positively and accruals negatively predict future stock returns, because some investors do not impound the information contained in the division of earnings between cash flow and accruals. Since cash flow is more favorable than accruals as a forecaster of future earnings, high accruals are associated with overvaluation, and high cash flow with undervaluation. The model does not in general predict that either of the cash flow or the accruals effects on future returns will completely subsume the other. We also find that the level of earnings predicts future stock returns when investors are inattentive toward earnings news. This predictability arises when inattentive investors stick to their beliefs based on the prior expectation of short-term earnings (e.g., analyst forecasts), instead
of updating and improving the precision of their forecasts based on the realization of
short-term earnings. Finally, we show that when processing information is costly, in
equilibrium some investors may actively trade yet choose not to attend to some public
information. This may explain why inattentive investors can form incorrect expectations
that influence price in the long run.

The model provides a rich set of new empirical implications. When some investors
neglect earnings, and others attend to earnings but neglect accruals, price underreacts
to earnings, but overreacts to accruals relative to cash flow. In consequence, there is
stronger underreaction to cash flow than to earnings. Furthermore, if enough attention
is paid to earnings, there is overreaction to accruals. A lower quality of accruals as
an earnings forecaster (induced, for example, by earnings management) strengthens
the accrual anomaly, but has an ambiguous effect on the cash flow anomaly. The more
variable are accruals relative to cash flows, the stronger is the cash flow anomaly relative
to the accrual anomaly. The ratio of the cash flow effect to the accruals effect is stronger
than the ratio of accruals variance to cash flow variance. Higher correlation between cash
flows and accruals tends to weaken the accrual anomaly. In addition, owing to neglect of
earnings news, the profit anomaly becomes stronger when there is less attention toward
earnings news and expected earnings (such as analyst forecasts of short-term earnings)
becomes a less accurate predictor of future earnings compared to actual short-term
earnings.

Another appealing set of empirical implications is provided by events that shift atten-
tion to or away from the firm. For example, a greater number of distracting events such
as earnings announcements by other firms is predicted to intensify both post-earnings
announcement drift and the cash flow anomaly. The implications of shifts of attention
toward or away from a firm for the accrual anomaly is less clear-cut. Neglect of earn-
ings components (cash flow versus accruals) strengthens the accrual anomaly, whereas
neglect of the earnings innovation itself weakens the accrual anomaly.

Our reconciliation of the drift and accruals effects is potentially consistent with some
of the informal intuitions about investor ‘naïveté’ offered in the empirical literature. It
is of course reassuring that such insights can be captured within an equilibrium model.
However, formal analysis does not entirely confirm casual intuition. For example, in a
classic paper, Sloan (1996) argues that the accrual anomaly will be present if accruals
are a weaker forecaster than cash flows of future earnings, whereas we show that this is
a necessary but not sufficient condition. More importantly, our main contribution is to
develop a unified framework based upon a well-established psychological constraint that
offers a rich set of untested empirical predictions about earnings, cash flow, accruals, and stock market prices.

Our model is, of course, very stylized. A natural further direction of extension is to consider finer components of accruals and cash flow. An analysis similar to that developed here would show that overvaluation will be increasing with the level of an earnings component if that component is a relatively strong incremental forecaster of future earnings, and will be decreasing with that component if it is a weak incremental forecaster of future earnings. Thus, our approach is consistent with the evidence on earnings components and returns of Richardson et al. (2005) and Dechow and Ge (2006). Furthermore, it would be simple to apply such a generalized analysis to normal versus abnormal levels of accruals (where ‘normal’ is evaluated relative to an industry- or other benchmark), in order to develop predictions about how managerial discretion affects misvaluation as studied empirically in Teoh, Welch and Wong (1998b, 1998a) and others. Also, by adding another time period, the relation between the forecasting power of variables for short-term versus long-term earnings with misvaluation could be considered. The framework can be applied to investment-related anomalies as well. These and other possible extensions suggest that the modeling approach offered here potentially has a wide range of applicability to circumstances in which market misvaluation derives from investor neglect of publicly available financial information.
Appendix

Proof of Proposition 6: By (34),

\[
\begin{align*}
\bar{b}_a^e &= f^e(\beta_1 - \beta_2) \frac{C^2 - V_a V_c}{V_a (V_a + V_c + 2C)} \\
&= f^e(\beta_1 - \beta_2) \frac{\rho^2 - 1}{V_a + 1 + 2\rho \sqrt{V_a \over V_c}}. \quad (53)
\end{align*}
\]

Similarly,

\[
\begin{align*}
\bar{b}_c^e &= f^e(\beta_1 - \beta_2) \frac{-C^2 + V_a V_c}{V_c (V_a + V_c + 2C)} \\
&= f^e(\beta_1 - \beta_2) \frac{1 - \rho^2}{V_a + 1 + 2\rho \sqrt{V_a \over V_c}}. \quad (54)
\end{align*}
\]

Differentiating \(\bar{b}_a^e\) in (53) with respect to \(\rho\):

\[
\frac{\partial (\bar{b}_a^e)}{\partial \rho} = f^e(\beta_1 - \beta_2) \frac{2\rho (V_a \over V_c) + 1 + 2\rho \sqrt{V_a \over V_c} - 2(\rho^2 - 1) \sqrt{V_a \over V_c}}{(V_a + 1 + 2\rho \sqrt{V_a \over V_c})^2}. \quad (55)
\]

Since \(\beta_1 > \beta_2\), It follows that \(\frac{\partial (\bar{b}_a^e)}{\partial \rho}\) is positive if and only if the term in the numerator is positive, i.e.,

\[
2\sqrt{V_a \over V_c} \left[ 1 + \rho^2 + \rho \left( \sqrt{V_a \over V_c} + \sqrt{V_c \over V_a} \right) \right] = 2\sqrt{V_a \over V_c} \left( \rho + \sqrt{V_a \over V_c} \right) \left( \rho + \sqrt{V_c \over V_a} \right) = 2(C + V_a)(C + V_c) \over V_c > 0. \quad (56)
\]

The second equality follows from the fact that \(\rho = C / \sqrt{V_a V_c}\), and the inequality holds under our assumptions in (16) that both accruals and cash flow are positively correlated with contemporaneous earnings.

Differentiating \(\bar{b}_c^e\) in (54) with respect to \(\rho\), similar reasoning shows that

\[
\frac{\partial (\bar{b}_c^e)}{\partial \rho} = -2f^e(\beta_1 - \beta_2) \sqrt{V_c \over V_a} \left( \rho + \sqrt{V_a \over V_c} \right) \left( \rho + \sqrt{V_c \over V_a} \right) \left( V_a + 1 + 2\rho \sqrt{V_a \over V_c} \right)^2 < 0.
\]
Proof of Proposition 9: It follows from equation (25) that
\[
E[S_2|c_1, a_1] = E[S_2] + (1 + \beta_1)(c_1 - \bar{e}_1) + (1 + \beta_2)(a_1 - \bar{a}_1).
\] (57)
Substituting \(E[S_2|e_1]\) and \(E[S_2|c_1, a_1]\) in (10) with (18) and (57), date 1 stock price can be written as
\[
S_1 = E[S_2] + [f^u(1 + \beta_1) + f^e(1 + \beta_{e|e_1})](c_1 - \bar{e}_1)
+ [f^a(1 + \beta_2) + f^e(1 + \beta_{e|e_1})](a_1 - \bar{a}_1) - A\text{var}(S_2)x_0.
\] (58)
From (58) and (25), the price change \(S_2 - S_1\) is
\[
S_2 - S_1 = [(f^u + f^e)(1 + \beta_1) - f^e(1 + \beta_{e|e_1})](c_1 - \bar{e}_1)
+ [(f^u + f^e)(1 + \beta_2) - f^e(1 + \beta_{e|e_1})](a_1 - \bar{a}_1) + \delta + A\text{var}(S_2)x_0. \tag{59}
\]
Since \(E[S_1] = E[S_2]\), the expected profit of an inattentive investor is
\[
E\left[\frac{E[S_2] - S_1}{A\text{var}(S_2)}(S_2 - S_1)\right] = \frac{\text{cov}(-S_1, S_2 - S_1)}{A\text{var}(S_2)}. \tag{60}
\]
Using \(S_1\) and \(S_2 - S_1\) given by (58) and (59) and simplifying some terms using the definition of \(\beta_{e|e_1}\) in (15), the covariance term can be written as follows.
\[
\text{cov}(-S_1, S_2 - S_1) = -f^u f^e(1 + \beta_{e|e_1})^2(V_a + V_e + 2C) - f^u f^a \text{var}(E[S_2|c_1, a_1])
- \frac{f^a f^e(1 + \beta_2)^2(V_e V_a - C^2)}{(V_a + V_e + 2C)} + A^2(\text{var}(S_2))^2V_x. \tag{61}
\]
Since \(V_a + V_e + 2C = V_e > 0\) and \(V_e V_a \geq C^2\), uninformed investors earn negative expected profits when there is no noise trading, i.e., \(V_x = 0\). However, the expected profit of an uninformed investor is positive when
\[
V_x > \frac{f^u f^e(1 + \beta_{e|e_1})^2(V_e)^2 + f^u f^a \text{var}(E[S_2|c_1, a_1])V_e + f^a f^e(1 + \beta_2)^2(V_e V_a - C^2)}{A^2(\text{var}(S_2))^2V_e}. \tag{62}
\]
The expected trading profit of an investor attending only to earnings is
\[
E\left[\frac{E[S_2|e_1] - S_1}{A\text{var}(S_2)}(S_2 - S_1)\right] = \frac{\text{cov}(E[S_2|e_1] - S_1, S_2 - S_1)}{A\text{var}(S_2)}. \tag{63}
\]
The equality follows from the fact that $E[E[S_2|e_1]] = E[S_1] = E[S_2]$. From (58) and (18),

$$E[S_2|e_1] - S_1 = [(f^a + f^e)(1 + \beta e e_1) - f^a(1 + \beta_1)](c_1 - \bar{c}_1) + [(f^u + f^a)(1 + \beta e e_1) - f^a(1 + \beta_2)](a_1 - \bar{a}_1) + A\text{var}(S_2)x_0. \quad (64)$$

From (59) and (64),

$$\text{cov}(E[S_2|e_1] - S_1, S_2 - S_1) = (f^u)^2(1 + \beta e e_1)^2V_e - \frac{f^a(f^u + f^e)(\beta_1 - \beta_2)^2(V_eV_a - C^2)}{V_e} + A^2(\text{var}(S_2))^2V_x. \quad (65)$$

Investors attending only to earnings can earn positive expected profits owing to uninformed investors (the first term) and noise trading (the last term). The expected profit of an investor who attends to earnings is positive if and only if

$$V_x > \frac{-(f^u)^2(1 + \beta e e_1)^2(V_e)^2 + f^a(f^u + f^e)(\beta_1 - \beta_2)^2(V_eV_a - C^2)}{A^2(\text{var}(S_2))^2V_e}. \quad (66)$$

Thus, investors attending only to earnings can earn positive expected profits if there is a large amount of noise trading ($V_x$ satisfying the above inequality).
References


44


