Earnings Persistence, Earnings Informativeness, and Stock Return Regularities*

Kai Du
Smeal College of Business
Pennsylvania State University
kxd30@psu.edu

Steven Huddart
Smeal College of Business
Pennsylvania State University
huddart@psu.edu

April 2017

*We thank Jianhong Chen, Pingyang Gao, Robert Göx, Jeremiah Green, Shuoyuan He, Tina Huynh, Lawrence Jin, Rick Laux, Jonathan Lewellen, Jia Li, Pierre Liang, John Liechty, Hong Qu, Jalal Vafi Sani, Jack Stecher, Mark Tippett (discussant), Biqin Xie and workshop participants at Pennsylvania State University, University of Southern Denmark, Michigan State University, Carnegie Mellon University, University of Waterloo, University of Zurich, and the 10th MEAFA Research Meeting at University of Sydney for helpful comments. We especially acknowledge the input of Lingzhou Xue on an earlier project that is related to the paper. All errors are our own. Financial support from Penn State Smeal College of Business is gratefully acknowledged.

Reporting System, Investor Learning, and Stock Return Regularities

Abstract

We propose a unified framework for understanding the impacts of earnings persistence and earnings informativeness in explaining accounting-based regularities. The underlying states of a firm follow a Markov process and are only imperfectly revealed through signals from a reporting system. Based on the signals, a representative Bayesian investor forms beliefs about the hidden states. Our model delivers predictions that are consistent with several empirical findings on the relationships between accounting reports and stock returns, including market reactions to breaks of earnings strings, and the return predictability based on accruals and book-tax differences. We also conduct simulations to gauge how these regularities vary with two characteristics of the reporting system, informativeness and conservatism.

Keywords: Reporting system; Investor learning; Earnings strings; Accruals anomaly; Book-tax differences

1 Introduction

Empirical research in accounting and finance has identified a number of regularities on the relationship between stock prices and accounting information (for reviews, see Kothari 2001; Richardson et al. 2010). Even though researchers have attributed the regularities to various economic, psychological, and institutional factors, the explanations are still debated in the literature.

In particular, numerous explanations have based based on two important concepts that the empirical literature has been vague about: earnings persistence, i.e., the properties of underlying economic profitability process (via parameters a and b) and earnings informativeness, i.e., the properties of accounting information system (via parameters c and d). In general, observed earnings patterns are functions of both persistence and informativeness. While some empiricists are familiar with this statement, whether and how they recognize its application to explain empirical regularities is less clear. Furthermore, both concepts sometimes seem tangled up with the notion of earnings quality. An example of such confusion in the following question often heard in workshops: is earnings persistence or earnings quality an economic characteristics of the firm's fundamental business operations or a characteristics of the firm's reporting system?

In this paper, we propose a simple model of reporting system and investor learning. We show that the model yields predictions consistent with several empirical regularities, including market reactions to breaks in strings of earnings surprises, and the return predictability based on accruals and book-tax differences.

In the model, the underlying state of the firm evolves according to a Markov process. A reporting system generates accounting signals from states. A representative investor does not observe the underlying states, but has knowledge of the transition probabilities of the states and the state-contingent probabilities of the signals. Stock price is determined by the representative investor, who forms a belief in Bayesian fashion each period about the current state of the firm (and, hence, the firm's value) from the history of accounting signals. Because the accounting signals are imperfect indicators of the underlying state of the firm and the state can change from one period to the next, the stock price does not stabilize or settle down. Instead, there is typically a price change in response to each new signal.[1]

Departing from the world of the representative investor, much of our analysis seeks to imitate the way in which an empiricist could detect patterns based on a large number of idiosyncratic firms which collectively approximate the true data-generating process. We focus on two empirical regularities: (i) market reactions to breaks to earnings strings (e.g., Barth et al. 1999; Ke et al. 2003), and (ii) the return predictability based on the difference of two signals, such as accruals as the difference between earnings and cash flows (Sloan 1996) and the difference between book income and taxable income (e.g., Lev and Nissim 2004).

Regarding market reactions to breaks of earnings strings, our analysis shows that the average reaction to the break of a string of earnings increases is more negative and approaches a lower bound as the the string preceding the break lengthens. The intuition is that, as the earnings string grows, the investor's belief that the state is bad decreases. The break in the string triggers an upward revision in this belief, and the magnitude of this revision increases with the length of the string but is constrained by an upper bound implied by the nature

[1]More formally, the mean-reversion property of the Markov process ensures that the price never converges in the long run unless the reporting system is entirely uninformative.

of the belief updating process. This prediction is confirmed by empirical analysis using a sample of Compustat firms over the period 1963–2013. The finding suggests that investor learning may help explain the market valuation of patterns of earnings, providing insights different from existing explanations that link earnings patterns to growth and risk (e.g., Barth et al. 1999).

We then examine the implications of the model for return predictability. Because return predictability in our model is rooted in the mean-reversion of a Markov process, we interpret a "period" as an interval after which an underlying profitability shock may have reversed itself. We first show that, in the one-signal setup, accounting signal negatively predicts stock return in the subsequent period. This is because, from an empiricist's perspective, expected returns will be determined by distributional properties of all possible beliefs in the sample. The average subsequent return of a group of firms with a low signal will be different from a group with a high signal, because different signals imply different updating dynamics in individual investors' beliefs, which collectively cause systematic differences in the average returns. This argument is built on the distinction between the Bayesian perspective and the frequentist's view first noted in statistical decision theories (Berger 1985) and recently applied to asset pricing research (Lewellen and Shanken 2002).

We also explore a richer reporting system. Specifically, we suppose the firm issues two signals each period. The informativeness of the two signals need not be the same and the signals themselves may be concordant or discordant. We show that an empiricist can profit from a hedge strategy that takes a long position in firms for which the more informative signal is low while the less informative signal is high, and a short position in firms with the opposite combination of signals. This is because both signals in general negatively predict the stock returns in the next period; whenever the two signals disagree, an empiricist can still profit by relying on the more informative signal.

There are at least two real-world examples of the two-signal reporting system. The two signals could be interpreted as earnings and cash flows, so that their difference corresponds to accounting accruals. Alternatively, they could be interpreted as book income and taxable income. The two interpretations shed light on two accounting-based "anomalies," the accruals anomaly (Sloan 1996) and the return predictability of book-tax-differences (e.g., Lev and Nissim 2004), respectively. The two anomalies, stated in the language of our framework, suggest that larger accruals (earnings minus cash flows) and larger book-tax differences (book income minus taxable income) each imply lower future returns. An explanation based on our framework is that, earnings are more informative than cash flows, and book income is more informative than taxable income. This explanation for the accruals anomaly is consistent with the concurrent trends in the informativeness of GAAP earnings (e.g., Dichev and Tang 2008) and the strength of the accruals anomaly (e.g., Green et al. 2011) as well as cross-country variations in the magnitude of the accruals anomaly (e.g., Kaserer and Klingler 2008). The explanation of the return predictability of book-tax differences is consistent with the improved value relevance of book information provided by SFAS No. 109 (e.g., Ayers 1998) and the stronger "mispricing" of temporary book-tax differences after the same standard (e.g., Chi et al. 2014).

To better illustrate the empirical predictions as suggested by the framework, we simulate the model for a large number of idiosyncratic histories over a large number of periods, and conduct "empirical" tests based on simulated data. Through simulations, we are also able to study how a regularity varies with two prominent properties of the reporting system, informativeness and conservatism, although we acknowledge that the two concepts are intertwined.

The accounting theories on the reporting systems, following the contributions of Blackwell (1951) and Marschak and Miyasawa (1968), have primarily focused on the decision usefulness of accounting information in a statistical decision framework or an agency setting (e.g., Demski 1973; Gjesdal 1981). Our focus is different. We examine the implications of reporting systems for the dynamics of stock returns. Our findings shed light on the interaction between investor learning and financial reporting, and how the properties of reporting systems affect stock returns.

Our study provides a unified framework for relating the underlying fundamental value of the firm to investor learning from public reports. We offer a novel interpretation of earnings-return associations (for a review, see Kothari 2001), stock price reactions to breaks of earnings strings, and other accounting-based return regularities (for a review, see Richardson et al. 2010). By characterizing the interpretation of accounting reports as a filtering problem, our approach emphasizes the joint impact of financial reporting and investor learning on return dynamics. In this regard, our study is related to studies on various learning schemes such as regime-shifting beliefs and least-squares learning (e.g., Barberis et al., 1998; Evans and Honkapohja, 2001; Du 2016).

The remainder of the paper is organized as follows. Section 2 introduces a simple model of Bayesian learning by a representative investor in the presence of a reporting system, and characterizes investor's belief updating process and stock prices. Section 3 discusses the implications for empirical tests of return regularities. Section 4 conducts simulation analysis. Section 5 concludes.

2 Basic framework

2.1 The process of true earnings

A firm operates in a stochastic environment, and there are two possible states for the firm's fundamental profitability x_t in each period t: the bad state (B) and the good state (G), where $0 < B < G$, i.e., $x_t \in X = \{B, G\}$. The law of motion of the states is described by a Markov transition matrix,

$$P = \begin{bmatrix} \Pr(x_{t+1} = B | x_t = B) & \Pr(x_{t+1} = G | x_t = B) \\ \Pr(x_{t+1} = B | x_t = G) & \Pr(x_{t+1} = G | x_t = G) \end{bmatrix} = \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix}, \tag{2.1}$$

where $0 \leq a, b \leq 1$. The parameters a and b capture the persistence of the underlying earnings process. This assumption is motivated by observations that firm profitability is mean reverting (e.g., Fama and French 2000; Fairfield et al. 2009). We further impose the following condition which requires that the Markov process exhibits some level of persistence.

Condition 1 $a + b \geq 1$.

2.2 Reporting system

2.2.1 One-signal reporting system

We consider two alternative types of reporting systems. The first one is a one-signal system. Following Marschak and Miyasawa (1968) and Marschak (1971), we introduce a reporting system denoted by the 3-tuple $< X, Y, \eta >$, where $Y = \{L, H\}$ is the set of possible signals ("low" (L) and "high" (H)), and η is a measurement matrix which maps the states to accounting signals for each period. Even though the labeling of states and signals does not affect the properties of the reporting system,[2] we assume ordinal values for states and signals so that we may discuss certain properties of the reporting system, such as conservatism. The state-contingent probabilities of the signals are given by the matrix η,

$$\eta = \begin{bmatrix} \Pr(y_t = L|x_t = B) & \Pr(y_t = H|x_t = B) \\ \Pr(y_t = L|x_t = G) & \Pr(y_t = H|x_t = G) \end{bmatrix} = \begin{bmatrix} c & 1-c \\ 1-d & d \end{bmatrix}, \qquad (2.2)$$

where $0 \leq c, d \leq 1$. The structure of the model is illustrated in Figure 1.

The following condition ensures that the reporting system is sufficiently informative.

Condition 2 $c + d \geq 1$.

The reporting system essentially summarizes firm-specific technologies and practices of applying accounting rules to measure underlying economic transactions. The random mapping from states to signals is an abstraction of numerous accounting estimates which require accountants' assessment of uncertain economic positions. For example, accounting for loss contingencies (recognition vs. disclosure) requires accountants' subjective evaluation of whether a contingent loss is probable and estimable; impairment of long-lived assets involves a recoverability test based on expectations of future cash flows; accounting for bad debt provisions requires estimates of uncertain future collections.[3] The attributes of the mapping are influenced by a variety of factors, including accounting standards, industry norms, innate firm characteristics, and above all, the discretion of managers and accountants.

An important concept in theories of financial reporting systems is informativeness. Es-

[2]In the language of Marschak (1971), if η is of order 2×2 and W is a permutation matrix of order 2, then η and ηW are equivalent. In other words, labeling would not change the informativeness of a reporting system.

[3]Even though the law of large numbers stipulate that estimates based on large samples of certain transactions (such as warranty expenses) are predictable with certain precision, abundant other transactions are non-diversifiable in nature. Scott (1979) provides a theoretic account on the probabilistic nature of the reporting system.

sentially, informativeness is a concept based on the decision usefulness of the information system (e.g., Blackwell 1951). rigorous definitions of informativeness cannot typically be reduced to a complete ordering based on a subset of model parameters. Our setting abstracts from the decision-usefulness framework. However, we make the following observation.

Remark 1 *Consider two reporting systems $\eta = [c, 1-c; 1-d, d]$ and $\eta' = [c', 1-c'; 1-d', d']$. A sufficient condition for η to be more informative than η' is that $c \geq c'$ and $d \geq d'$.*

Some special cases of the reporting system are easy to define following prior literature (e.g., Marschak 1971). A reporting system η is said to be entirely uninformative if its rows are identical, i.e., $c + d = 1$; η is said to be perfect if $c = d = 1$.[4]

Our main analysis is based on this simple setup in which both the underlying states and accounting signals take binary values. However, the model can be extended to a general case in which both states and signals may take more than two values.

A Bayesian investor, whose beliefs reflect the consensus forecasts of heterogenous investors in the economy, updates her posterior about the underlying state given the growing history of the accounting signals. Denote investor's posterior at t by $\mu_t = \Pr(x_t = B|\boldsymbol{y}^t)$, where $\boldsymbol{y}^t = \{y_0, ..., y_t\}$ is the history of accounting reports. Using Bayes' Rule, it is easy to show that the updating process of investor beliefs is given by the following lemma.

Lemma 1 *Given investor belief of the previous period (μ_{t-1}) and the accounting signal (y_t), the belief at the end of period t is given by*

$$\mu_t|_{y_t=L} = \frac{c((a+b-1)\mu_{t-1} + 1 - b)}{(c+d-1)(a+b-1)\mu_{t-1} + (1-d)b + c(1-b)}, \tag{2.3}$$

$$\mu_t|_{y_t=H} = \frac{(1-c)((a+b-1)\mu_{t-1} + 1 - b)}{(1-c-d)(a+b-1)\mu_{t-1} + db + (1-c)(1-b)}. \tag{2.4}$$

[4]In general, an information system is perfect if it is a permutation matrix, obtained by permuting the rows of an identify matrix, and contains exactly one entry of 1 in each row and each column and 0 elsewhere. Given the ordinal values of states and signals, we adopt a more restrictive definition.

To gain some insights on the belief updating process, consider two extreme cases—uninformative reporting system and perfect reporting system. With uninformative reporting system ($c + d = 1$), there is still updating in $\mu(t)$ due to the nature of Markov process,

$$\mu_t = (a + b - 1)\mu_{t-1} + 1 - b \tag{2.5}$$

As long as $0 < a < 1$ and $0 < b < 1$, the two-state Markov process dictates that investor beliefs will converge to the stationary distribution in the long run,

$$\lim_{t\to\infty} \mu_t = \pi_B = \frac{1-b}{2-a-b}. \tag{2.6}$$

With a perfect reporting system ($c = d = 1$), μ_t will oscillate between 1 and 0 depending on the current signal.

The recursive structure suggests that investor beliefs and stock prices are functions of the patterns of earnings sequences. The following proposition discusses the bounds to the beliefs implied by such recursion.

Proposition 1 *Suppose Conditions 1 and 2 hold. (i) If* $\mu_t \in [\underline{\mu}, \bar{\mu}]$*, then* $\mu_{t+s} \in [\underline{\mu}, \bar{\mu}]$ *for* $\forall s \geq 1$*; (ii) If the investor starts with prior* $\mu_0 \notin [\underline{\mu}, \bar{\mu}]$*, then with probability one,* μ_t *will eventually jump into* $[\underline{\mu}, \bar{\mu}]$ *and stay in this interval forever, where*

$$\underline{\mu} = \frac{1}{2}\left(\frac{a(c-1) + b(2c+d-2) - 2c + 2}{(a+b-1)(c+d-1)} - \sqrt{\frac{a^2(c-1)^2 - 2a(b-2)(c-1)d + d\,(b^2 d + 4b(c-1) - 4c + 4)}{(a+b-1)^2(c+d-1)^2}}\right),$$

and

$$\bar{\mu} = \frac{1}{2}\left(\sqrt{\frac{c\,(a^2 c + 4a(d-1) - 4d + 4) - 2(a-2)bc(d-1) + b^2(d-1)^2}{(a+b-1)^2(c+d-1)^2}} + \frac{(a-2)c + b(2c+d-1)}{(a+b-1)(c+d-1)}\right).$$

Intuitively, if the investor sees a long series of L signals, μ_t increases toward an upper bound $\bar{\mu} \leq 1$; if the investor sees a long series of H signals, μ_t decreases toward a lower bound $\underline{\mu} \geq 0$. In other words, if we focus on the long run when investor learning is no longer influenced by priors, it must be the case that investor beliefs are bounded by $\underline{\mu}$ and $\bar{\mu}$. The following corollary discusses the properties of the bounds.

Corollary 1 *Suppose Conditions 1 and 2 hold. The lower bound* $\underline{\mu}$ *and upper bound* $\bar{\mu}$ *exhibit the following relationships with respect to parameters of the reporting system: (i)* $\frac{\partial \underline{\mu}}{\partial a} \geq 0$, $\frac{\partial \underline{\mu}}{\partial b} \leq 0$, $\frac{\partial \underline{\mu}}{\partial c} \leq 0$, *and* $\frac{\partial \underline{\mu}}{\partial d} \leq 0$*; (ii)* $\frac{\partial \bar{\mu}}{\partial a} \geq 0$, $\frac{\partial \bar{\mu}}{\partial b} \leq 0$, $\frac{\partial \bar{\mu}}{\partial c} \geq 0$, *and* $\frac{\partial \bar{\mu}}{\partial d} \geq 0$.

Intuitively, ceteris paribus, if the B state is more persistent, investor's probabilistic assessment of the state being indeed B is higher (signified by higher lower and upper bounds); if the G state is more persistent, the opposite is true. A more informative reporting system will enlarge the range for possible investor beliefs, because more informative signals induce more substantive revisions.

2.2.2 Two-signal reporting system

In the preceding analysis we have defined a reporting system with one signal for each period and we could interpret the signal as earnings. In the real world, however, earnings are usually not the only indicator of firm profitability (e.g., Antle et al. 1994; Jiang 2016). There are various institutional contexts of financial reporting in which multiple channels of public information provide noisy information about the same underlying state. The most proverbial example of a two-signal reporting system is one that reports both earnings and cash flows. According to Christensen and Demski (2003), earnings and cash flows are "simply two different ways of doing the accounting, and both, in principle, are sources of information" (p. 128), and "the typical financial report contains accrual and cash basis renderings, two different partitions so to speak" (p. 115).

Another example of a two-signal reporting system is the coexistence of book income and taxable income. The rules and principles that govern the measurement of book income are sometimes different from those that govern income tax reporting (e.g., Graham et al. 2012). Therefore, book income and taxable income constitute alternative performance measures, supplementing each other so that together they provide more information about future earnings and firm valuation than each does by itself (e.g., Dhaliwal et al. 2017). Indeed, prior research documents that book-tax differences provide information for the persistence and future growth of pre-tax income (e.g., Hanlon 2005).

Therefore, it is of interest to augment the basic model with a second signal about the same state. Suppose the underlying state x_t is imperfectly revealed by a reporting system with two information channels, $< X, Y, Z, \eta^1, \eta^2 >$, where

$$\eta^1 = \begin{bmatrix} c_1 & 1-c_1 \\ 1-d_1 & d_1 \end{bmatrix}, \quad \eta^2 = \begin{bmatrix} c_2 & 1-c_2 \\ 1-d_2 & d_2 \end{bmatrix}. \tag{2.7}$$

For every state x_t, η^1 generates a signal y_t, and independently, η^2 generates a concurrent signal z_t. The structure of the two-signal reporting system is illustrated by Figure 2.

The following condition ensures that both signals are sufficiently informative.

Condition 3 $c_1 + d_1 \geq 1$, $c_2 + d_2 \geq 1$.

Analogous to the one-signal case, we can derive the beliefs of investor who observes two sequences of signals, y_t and z_t.

Lemma 2 *Given investor belief of the previous period (μ_{t-1}) and the signals (y_t and z_t),* *the belief at the end of period t is given by*

$$\mu_t|_{y_t=L,z_t=L} = \frac{c_1c_2((a+b-1)\mu_{t-1}+1-b)}{c_1c_2((a+b-1)\mu_{t-1}+1-b)+(1-d_1)(1-d_2)((1-a-b)\mu_{t-1}+b)}, \tag{2.8}$$

$$\mu_t|_{y_t=L,z_t=H} = \frac{c_1(1-c_2)((a+b-1)\mu_{t-1}+1-b)}{c_1(1-c_2)((a+b-1)\mu_{t-1}+1-b)+(1-d_1)d_2((1-a-b)\mu_{t-1}+b)}, \tag{2.9}$$

$$\mu_t|_{y_t=H,z_t=L} = \frac{(1-c_1)c_2((a+b-1)\mu_{t-1}+1-b)}{(1-c_1)c_2((a+b-1)\mu_{t-1}+1-b)+d_1(1-d_2)((1-a-b)\mu_{t-1}+b)}, \tag{2.10}$$

$$\mu_t|_{y_t=H,z_t=H} = \frac{(1-c_1)(1-c_2)((a+b-1)\mu_{t-1}+1-b)}{(1-c_1)(1-c_2)((a+b-1)\mu_{t-1}+1-b)+d_1d_2((1-a-b)\mu_{t-1}+b)}. \tag{2.11}$$

Similar to the one-signal case, we can provide bounds for the belief variable.

Proposition 2 *Suppose Conditions 1 and 3 hold. (i) If* $\mu_t \in [\mu^*, \mu^{**}]$*, then* $\mu_\tau \in [\mu^*, \mu^{**}]$ *for* $\forall \tau > t$*; (ii) If the investor starts with prior* $\mu_0 \notin [\mu^*, \mu^{**}]$*, then with probability one,* μ_t *will eventually jump into this interval and stay in this interval forever, where* μ^* *and* μ^{**} *are given by (A.16) and (A.14).*

The following corollary discusses the properties of the bounds. The intuitions behind the comparative statics are similar to those of Corollary 1.

Corollary 2 *Suppose Conditions 1 and 3 hold. The lower bound* μ^* *and upper bound* μ^{**} *exhibit the following relationships with parameters of the reporting system: (i)* $\frac{\partial \mu^*}{\partial a} \geq 0$*,* $\frac{\partial \mu^*}{\partial b} \leq 0$*,* $\frac{\partial \mu^*}{\partial c_j} \leq 0$*, and* $\frac{\partial \mu^*}{\partial d_j} \leq 0$ *for* $j = 1, 2$*; (ii)* $\frac{\partial \mu^{**}}{\partial a} \geq 0$*,* $\frac{\partial \mu^{**}}{\partial b} \leq 0$*,* $\frac{\partial \mu^{**}}{\partial c_j} \geq 0$*, and* $\frac{\partial \mu^{**}}{\partial d_j} \geq 0$ *for* $j = 1, 2$*.*

2.3 Stock prices

We interpret x_t as the underlying cash flows of period t, and assume the investor prices the firm based on the risk-neutral expectation of discounted future cash flows of all future periods. The post-accounting-signal price is the cum-dividend price, p_t, which as a function of investor beliefs is given by the following lemma.[5]

[5]We abstract from institutional contexts which may justify such pricing functions. The pricing function can be motivated by an economic setting in which investors trade on future claims to the entire discounted dividend stream of the company, without ever receiving (and thereby observing) the dividends. We make this

Lemma 3 *After the accounting signals becomes public in each period t, the cum-dividend stock price p_t is given by*

$$p_t = \frac{1+\delta}{\delta(2-a-b+\delta)}\big(G(1-a+\delta) + B(1-b) - \delta(G-B)\mu_t\big), \tag{2.12}$$

where δ is the discount rate.

It follows immediately that the price change from period t to period $t+1$ is given by

$$p_{t+1} - p_t = -\frac{1+\delta}{2-a-b+\delta}(G-B)(\mu_{t+1}-\mu_t), \tag{2.13}$$

where μ_{t+1} is given by (2.3) and (2.4) for the one-signal case and (2.8)–(2.11) for the two-signal case (with time indices updated by one period).[6]

Some discussion is in order about the nature of the profitability process and the role of learning in the pricing function. The Markov process of underlying states (dividends) dictates that learning does not in general lead the belief to a long-run mean. However, in the extreme case of entirely uninformative reporting system, the updating process reduces to (2.5) and price change is given by $p_{t+1} - p_t = \frac{1+\delta}{2-a-b+\delta}(G-B)((2-a-b)\mu_t - (1-b))$. In this case, even though the investor rationally ignores the signal, there is still updating due to the nature of Markov process, which eventually leads to the stationary distribution of beliefs and the price will be constant.

---

simplified assumption without formally modeling the institutions of dividends and trading in an overlapping generations model (e.g., De Long et al. 1990; Fischer et al. 2016), as our focus is to examine the implications of investor learning about the underlying profitability process.

[6] If we work instead with *ex-dividend* prices, the price function would be

$$p_t = \frac{G(1-a) + B(1-b+\delta) - (G-B)\delta\left((a+b-1)\mu_t - b\right)}{\delta(2-a-b+\delta)}, \tag{2.14}$$

and the price change would be

$$p_{t+1} - p_t = -\frac{a+b-1}{2-a-b+\delta}(G-B)(\mu_{t+1}-\mu_t). \tag{2.15}$$

The intuitions remain the same as the case of cum-dividend prices.

2.4 Issues in applying the model to empirical settings

2.4.1 Bayesian or frequentist?

The Bayesian investor in the model, who is assumed to be a rather passive actor, mechanically applies the recursive updating formula and prices the stock based on his expectation of future discounted dividends. Before discussing the empirical implications of the framework, we introduce a distinction between two perspectives: the Bayesian investor who updates her belief based on a history of earnings reports and a hypothetical frequentist-empiricist who attempts to detect return regularities in the data (i.e., a sample of earnings and stock prices).

The distinction between the Bayesian perspective and the frequentist perspective has been discussed by Berger (1985) in greater lengths. In asset pricing studies, Lewellen and Shanken (2002) use this distinction to explain why parameter uncertainty may give rise to return predictability but the agent within the model cannot perceive or exploit the predictability. Barberis et al. (1998) also implicitly adopt a frequentist perspective in examining the empirical implications of the regime-shifting beliefs. An important implication of the distinction is that the hedge strategies as perceived to be profitable by an empiricist cannot be exploited by the investor in the model, thereby perpetuating the existence of arbitrage opportunities.

Even though not a focus of the current study, we explicitly make the distinction between the two perspectives to avoid confusions. We use $E_t^i[\cdot]$ to denote the subjective expectation of the Bayesian investor, and use $E_t^f[\cdot]$ to denote the expectation of an empiricist who observes the entire sample of firms. For the investor, the probability that the underlying state is B is random and fluctuates over time as new signals arrive. However, from an empiricist's perspective, the average price change in the sample depends on the distribution of a large number of Bayesian investors. The following lemma derives the expectations from the Bayesian and frequentist perspectives.

Lemma 4 *(i) From a Bayesian investor's perspective,*

$$E_t^i[p_{t+1} - p_t] = \frac{1+\delta}{2-a-b+\delta}(G-B)\big((2-a-b)\mu_t - 1 + b\big); \tag{2.16}$$

(ii) From a frequentist-empiricist's perspective,

$$E_t^f[p_{t+1} - p_t] = 0. \tag{2.17}$$

With changing dividend process, expected price change is positively related to lagged belief μ_t, and price does not revert to a long-run mean. This feature is unlike models with stationary dividend processes (e.g., Lewellen and Shanken 2002). As we can see from equation (3.1),

$$\text{sign}\left(E_t^i[p_{t+1} - p_t]\right) = \text{sign}\left(\mu_t - \pi_B\right). \tag{2.18}$$

The intuition behind equation (3.3) is as follows. If the current probability assessment of the state being B is greater than the stationary probability, price is temporarily deflated (relative to the price level implied by the stationary distribution) and will on average rise in the future. On the other hand, if the current probability assessment of the state being B is lower than the probability implied by the stationary distribution, price is temporarily inflated and will decline in the future. Expected price change is zero only when if $\mu_t = \pi_B$, i.e., when investor belief is consistent with the stationary distribution of the underlying states.

According to an empiricist, the expectation of next-period price revision is taken over all possible investor beliefs μ_t in the empirical sample, which is assumed to consist of a large number of idiosyncratic firms. The average belief perceived by the empiricist is the stationary probability of the B state. Therefore, unconditionally, the empiricist perceives no return predictability in the sample. However, as we shall show below, conditional on different signals, the empiricist does perceive predictability.

2.4.2 What is a "period"?

Making connections between the model and real world data requires the translation of one "period" in the model to time intervals in the real world, such as quarters and years. We note that the intuitions underlying the results may be most salient in different contexts and horizons. Therefore, we choose the real-world interpretation of time horizon that makes most sense in each context, considering the main forces at play.

For earnings strings, it is customary to define strings based on quarterly or annual earnings, because anecdotes and empirical evidence show that investors can make salient comparisons across adjacent quarters or years. Therefore, we interpret the model "period" as a quarter or a year when we study market reactions to earnings strings. However, for the evaluation of return predictability, we emphasize the impact of mean reversion of the fundamental states. Therefore, we interpret a model period as an interval after which the underlying profitability shock may have reversed itself, typically over a horizon longer than one fiscal year.

2.4.3 Event study or association study?

The plausibility of the connection between a model prediction and the finding of an empirical study also depends on the methodology employed by the empirical study. Traditionally, the earnings-return relationship has been investigated using either event studies or association studies (Kothari 2001). In a typical event study, the researcher examines whether an earnings announcement causes investors to revise their cash flow expectations, as manifested in cumulative price changes over a short window. In an association study, the researcher examines the relationship between returns over relatively long periods and unexpected earnings.

Even though such distinction is apparent in empirical implementations, there is no way to differentiate between short-term and long-term price behavior in the stylized model, in which the price change over one period can be regarded as both the market reaction and the buy-and-hold return over one fiscal period. Therefore, in our discussion of the market reaction to breaks of earnings strings, we compare the model to event studies to more confidently identify the capital market reaction to the "event" of a break to an earnings string. However, in discussions of return predictabilities, we adopt an association-study perspective to align with the nature of the trading strategies based on accounting signals.

3 Application 1: Difference between signals and future returns

3.1 Predictions

3.1.1 One-signal reporting system

This framework can also be used to address accounting-based return predictability, i.e., how a hedge portfolio formed based on current period's earnings signal could yield an abnormal return in the next period. The following proposition evaluates one such hedge strategy based on the one-signal case.

Proposition 3 *Earnings signal negatively predicts stock returns in the next period. In other words, a trading strategy that takes a long position in stocks with an L signal and a short position in stocks with an H signal and holds them for one period is profitable:*

$$E_t^f(p_{t+1} - p_t|y_t = L) - E_t^f(p_{t+1} - p_t|y_t = H) \geq 0 \tag{3.1}$$

where the expectation operator denotes the mean over the unconditional distribution of beliefs, from a frequentist's perspective.

The proof of Proposition 5 relies on arguments about the distributional knowledge which is not available to the investor within the model. The intuition that future price change conditional on a high signal is smaller than future price change conditional on a low signal involves two arguments. First, future price change is an increasing function of μ_t. Second, in the whole population, on average, μ_t with an L signal is greater than μ_t with an H signal. Intuitively, for firms that have just experienced an L signal, investors on average have a lower μ_t, but as we know from the previous section, lower μ_t will eventually rise to the stationary level π_B, therefore creating a higher return.

Note that this result cannot be literally interpreted as a negative relationship between earnings and one-year-ahead returns. Instead, the prediction of Proposition 5 is only relevant to the extent that fundamental profitability exhibits reversion over one "period." This is because Proposition 5 is driven by the mean reversion of the underlying profitability process: an H signal predicts a negative return to the extent that the H signal somewhat reveals that the underlying state is G, which has a nonzero probability of transitioning to B.

3.1.2 Two-signal reporting system

The two-signal reporting system could shed light on factors underlying the return predictability based on two signals. Examples of such return predictability includes the accruals anomaly and the association between book-tax differences and future returns.

The accruals anomaly refers to the finding that the accruals component of earnings negatively predicts future returns (Sloan 1996). Researchers have proposed various explanations for the accruals anomaly, including investor fixation on accruals (Sloan 1996), earnings persistence (Richardson et al. 2005), and investment (Wu et al. 2010). Let the two signals correspond to earnings ("Earn") and cash flows ("CF"), respectively, i.e., $< X, Y, Z, \eta^{\text{Earn}}, \eta^{\text{CF}} >$. For every state x_t, η^{Earn} generates an earnings signal y_t, and η^{CF} generates a cash flow signal z_t. Accruals can be defined as the difference between earnings and cash flows,

$$Accruals_t = y_t - z_t \in \{-(H - L), 0, H - L\}. \tag{3.2}$$

The accruals anomaly states that by taking a long position in stocks with $Accruals_t = -(H - L)$ and a short position in stocks with $Accruals_t = H - L$, we can make a positive hedge return over period $t + 1$.

Prior studies have also documented a return predictability based on book-tax differences (e.g., Lev and Nissim 2006; Chi et al. 2014). Specifically, the ratio of taxable income to book income positively predicts future stock returns. Let the two signals be book income (or earnings, denoted "Book") and taxable income ("Tax"), respectively, i.e., $< X, Y, Z, \eta^{\text{Book}}, \eta^{\text{Tax}} >$. For every state x_t, η^{Book} generates an earnings (book income) signal y_t, and η^{Tax} generates a cash flow signal z_t. Book-tax difference (BTD) can be defined as the difference between book income and taxable income,

$$BTD_t = y_t - z_t \in \{-(H - L), 0, H - L\}. \tag{3.3}$$

The empirical findings can be restated as a negative relationship between BTD_t and subsequent return.

The following proposition formalizes this notion and provides conditions for hedge strategies based on discordant signals to be profitable. For concreteness, we state the proposition in the context of earnings and cash flows. However, its implications are applicable to any two-signal reporting system.

Proposition 4 *Consider a trading strategy that takes a long position in stocks with with negative accruals ($y_t = L, z_t = H$) and a short position in stocks with positive accruals ($y_t = H, z_t = L$) and holds them for one period. A sufficient condition for the strategy to be profitable is, in addition to Conditions 1 and 3,*

$$\frac{c_2 d_2 (1 - d_1)}{d_1(1 - d_2) + c_2(d_2 - d_1)} \leq c_1 \leq 1. \tag{3.4}$$

Proposition 6 states that the hedge strategy that buys firms for which $y_t = L$ and $z_t = H$ and shorts firms for which $y_t = H$ and $z_t = L$ is only profitable if y_t is sufficiently informative. This is easy to understand given the one-signal return predictability result (Proposition 5), which states that earnings signals negatively predict the stock returns in the next period.

Therefore, whenever the two signals disagree, an empiricist can still make a profit through a hedge strategy that relies heavily on the more informative signal.

In the context of the accruals anomaly, Proposition 6 suggests that accruals negatively predict future returns if earnings are in general more informative than cash flows. This is consistent with the basic notion that net income is considered a better indicator of future operating cash flows than is current net operating cash flow (Spiceland et al. 2013, p.7). This explanation also finds support in the time-series trends and cross-country variations. Over the years, cash flows have become increasingly informative relative to earnings, whose quality has been shown to be deteriorating due to the standard setters' stated goal of moving away from matching (Dichev and Tang 2008) and large numbers of write-offs and other one-time items. At the same time, the accruals anomaly has become less significant (Green et al. 2011). In addition, across different reporting environments, Kaserer and Klingler (2008) show that the accrual anomaly only exists for German firms that use IFRS or U.S. GAAP, not for those that use German GAAP, which is less conservative (informative) than the other two accounting standards.

In the same vein, Proposition 6 also prescribes an explanation for the return predictability based on book-tax differences. Namely, book income is more informative than taxable income about profitability. This notion is consistent with prior studies which document that book income exhibits significantly greater incremental explanatory power than taxable income (e.g., Hanlon et al. 2005). In addition, it is consistent with the improved value relevance of book information provided by SFAS No. 109 (e.g., Ayers 1998) and the stronger "mispricing" of temporary book-tax differences after SFAS No. 109 (e.g., Chi et al. 2014).

3.2 Simulation

We conduct numerical simulations to further study the empirical regularities discussed in Section 3. Our focus is on patterns that cannot be examined by analytical characterizations in a tractable way.

We calibrate the three nonessential parameters as $(B, G, \delta) = (1, 2, .06)$, which will be maintained throughout the simulation analysis.[7] In order to determine μ_t, we will need to specify the initial belief, μ_0. The choice of initial belief will not affect the simulation analysis given that the time-series simulated are sufficiently long.

To fix ideas for the two-signal reporting system, we interpret y_t as earnings and z_t as cash flows, with their difference being accruals. We simulate N idiosyncratic firms for T periods, and form portfolios based on the level of accruals, $Accruals_t$. Specifically, for each period, we form a zero-investment portfolio that takes a long position in stocks with negative accruals ($Accruals_t = -(H-L)$) and shorting all stocks with positive accruals ($Accruals_{t-1} = H-L$). We use the equal-weighted returns as portfolio returns. The existence of accruals anomaly would be indicated by a positive return to the hedge portfolio over some forecasting horizon.

Figure 4 reports the simulation results. Panel A reports the buy-and-hold return to an accruals-based hedge portfolio, where the x-axis denotes the number of periods forward, for a baseline set of parameters, $(a, b, c^{\text{Earn}}, d^{\text{Earn}}, c^{\text{CF}}, d^{\text{CF}}) = (.75, .75, .75, .75, .60, .60)$. The return predictability of accruals exists for at least seven periods into the future.

Figure 4, Panel B reports the variations of the one-period ahead accruals-based portfolio returns with respect to the persistence of the state process, i.e., $a = b \in [.50, .95]$, given

[7] The qualitative patterns of the results are not sensitive the choice of these nonessential variables.

$(c^{\text{Earn}}, d^{\text{Earn}}, c^{\text{CF}}, d^{\text{CF}}) = (.75, .75, .60, .60)$. The magnitude of the accruals anomaly first increases but then decreases as the underlying state becomes more persistent. Panel C reports the variations of the accruals anomaly with respect to the relative informativeness of cash flows, i.e., $\Delta = c^{\text{CF}} - c^{\text{Earn}} = d^{\text{CF}} - d^{\text{Earn}} \in [-.25, .25]$, for $(a, b, c^{\text{Earn}}, d^{\text{Earn}}) = (.75, .75, .75, .75)$. The positive hedge return to the accruals strategy only obtains when cash flows are less informative than earnings, consistent with our discussion of the empirical findings on the accruals anomaly in Section 3. Panel D reports how the magnitude of the accruals anomaly varies with the relative conservatism of cash flows, i.e., $\Delta = c^{\text{CF}} - c^{\text{Earn}} \in [-.25, .25]$, for $(a, b, c^{\text{Earn}}, d^{\text{Earn}}, d^{\text{CF}}) = (.75, .75, .75, .75, .75)$. The pattern is similar to that of relative informativeness, which is unsurprising given the overlap between the two concepts.

3.3 Discussion

4 Application 2: Market reactions to breaks of earnings strings

4.1 Predictions

Before studying the market reactions to patterns in earnings, we first examine the earnings response coefficient (ERC) implied by the model. The following proposition characterizes the market reactions to different earnings signals.

Proposition 5 *The average market reaction to an H signal is higher than the average market reaction to an L signal, i.e.,*

$$E_t^f[p_t - p_{t-1} | y_t = H] - E_t^f[p_t - p_{t-1} | y_t = L] \geq 0 \tag{4.1}$$

Proposition 3 essentially states that the ERC is positive, consistent with insights from the vast literature (e.g., Collins and Kothari 1989).

We then examine the average market reactions to earnings strings. Prior studies have documented a strong negative market reaction to a break of a string of consecutive earnings increases (e.g., DeAngelo et al. 1996; Barth et al. 1999; Ke et al. 2003). For example, Ke et al. (2003) find that the mean abnormal return to a break of a string of quarterly earnings increases is -1.77% over the three-trading-day window [-2,1] and -4.29% over the 32-trading-day [-30,1] window relative to the earnings announcement. The findings on the market valuations of earnings strings and market reactions to breaks of strings have been associated with growth or risk. For example, Barth et al. (1999) show that firms with earnings strings have higher price-earnings multiples than other firms, and that earnings strings are positively associated with proxies for growth and negatively associated with proxies for risk. In this study, we intend to provide an alternative explanation based on investor learning.

The following proposition characterizes the comparative statics of the average market reaction with respect to breaks of earnings strings, where an earnings string is defined as a series of consecutive H signals.

Proposition 6 *In addition to Conditions 1 and 2, suppose either of the following two conditions is satisfied: (1)* $0 \leq b \leq b^*$*, where* $b^* = \frac{c\left((2c+d-1)-\sqrt{(1-d)(4c+3d-3)}\right)}{2(c+d-1)^2}$*; (2)* $b^* \leq b \leq 1$ *and* $0 \leq a \leq a^*$*, where* $a^* = \frac{b^2(c+d-1)^2-bc(2c+d-1)+c(c-d+1)}{c(1-d)}$*. (i) The average market reaction to a break of earnings strings is negative, and is more negative as the string of* H *signals becomes longer; (ii) As the earnings string of* H *signals becomes sufficiently long, the market reaction to the break approaches a lower bound:*

$$\begin{aligned} &\lim_{s\to\infty} E_t^f[p_t - p_{t-1}|y_t = L, y_{t-1} = ... = y_{t-s} = H] \\ &= -\frac{1+\delta}{2-a-b+\delta}(G-B)\left(\frac{c(a+b-1)\underline{\mu}+c(1-b)}{(c+d-1)(a+b-1)\underline{\mu}+(1-d)b+c(1-b)} - \underline{\mu}\right) \end{aligned} \tag{4.2}$$

It can be easily verified that the two conditions stated in Proposition 4 are fairly nonrestrictive. Notice that as the string of H signals up to $t-1$ becomes longer, μ_{t-1} is on average lower. Therefore, a sufficient condition for the pattern stated in the proposition is that $\mu_t|_{y_t=L}$ is a decreasing function of μ_{t-1}, which can be shown to hold when either of the two conditions is true. Intuitively, after observing a string of H signals, if the reporting system is sufficiently informative, μ_t will be walked down, so the price change upon the break of earnings string will be negative. The longer the string, the more negative the reaction to the break.

Analogously, if a earnings string is defined as consecutive low (L) signals, the average market reaction to breaks of earnings strings is positive, and the magnitude of the average market reaction becomes greater as the string of L signals becomes longer under certain conditions. In the limit, we have,

$$\begin{aligned}&\lim_{s\to\infty} E_t^f[p_t - p_{t-1}|y_t = H, y_{t-1} = ... = y_{t-s} = L]\\ =& -\frac{1+\delta}{2-a-b+\delta}(G-B)\Big(\frac{(1-c)(a+b-1)\bar{\mu} + (1-c)(1-b)}{(1-c-d)(a+b-1)\bar{\mu} + db + (1-c)(1-b)} - \bar{\mu}\Big) \qquad (4.3)\end{aligned}$$

Since this pattern has not been documented in the literature, we inspect the empirical data for a similar one. Specifically, we measure the actual market reaction as the average reaction to breaks of earnings strings of lengths 1 to 8, based on a sample of about one million firm-quarters over 1963–2013 on Compustat quarterly data, after excluding firms with fiscal year changes. We obtain stock return data from CRSP daily files. Earnings strings are defined as consecutive earnings increases, where an "earnings increase" is defined as higher earnings per share (EPS) before extraordinary items than the same quarter of the previous year. Earnings announcement return is measured as the cumulative abnormal return starting two trading days before to one trading day after the earnings announcement date.

Returns are adjusted by value-weighted market or beta, where beta is measured by market beta estimated using daily returns of the previous month. We find that the actual pattern, based on both market-adjusted and beta-adjusted cumulative announcement returns, bears a striking resemblance to the predicted pattern (Panel B of Figure 3).[8]

4.2 Simulation

We first examine how the length of an earnings string affects the market reaction to a break. We simulate the model over $T = 1,000$ periods[9] for $N = 10,000$ idiosyncratic histories ("firms"). For each period, we form an equal-weighted portfolio of all firms with an L signal following exactly s consecutive H signals, where s=1, 2, ..., 8 is the length of earnings strings. We plot the average of portfolio returns (and their difference) over all periods for different lengths of the earnings strings. Except for Panel C, we assume that $a = b = .75$.

Figure 3, Panel A shows that the market reactions to breaks of earnings strings are on average negative, regardless of the parameters of the reporting system, consistent with prior evidence (e.g., Ke et al. 2003). More importantly, Panel A also shows the market reaction becomes more negative for longer strings, but the effect of increased length on amplifying the market reaction diminishes as the length increases. Panel B shows that the simulated pattern resembles the pattern detected based on actual data, as detailed in Section 3.

We then examine how the market reaction to string breaks varies with properties of the

[8] In the pattern derived from actual data, there is a kink at $s = 4$, where s is the length of the earnings string. This is possibly due to the periodicity of financial reporting: one fiscal year is comprised of four quarters, and investors' perception of earnings strings may change abruptly around $s = 4$.

[9] The number of periods is not intended to be descriptive of real-world firm ages. Rather, we simulate long histories to ensure the existence of a large number of earnings strings of various lengths, because the consistency of the simulation results crucially relies on averaging over a large number of firms with the same type of earnings patterns.

profitability process and the reporting system. We only focus on three lengths of earnings strings, $s = 2, 3, 8$. We fix the reporting system at $c = d = .75$. When we vary the persistence of underlying state (captured by a and b, which are assumed to be equal in this exercise), the market reaction decreases (Figure 3, Panel C). This is due to the decreasing weight the investor assigns to the accounting signal as persistence increases from .50 to 1: When persistence is close to 1, the investor puts little weight on the break signal (L), which she attributes to noise embedded in the reporting system. Therefore, she does not revise her belief downward significantly.

In Panel D of Figure 3, we hold $a = b = .75$ and vary c and d (set to be equal) over the interval $[.50, 1]$, intended to capture the varying level of informativeness. In Panel E, we hold $(a, b, d) = (.75, .75, .50)$ and vary $c \in [.50, 1]$ intended to capture varying conservatism. The average reaction decreases with both informativeness and conservatism. This is intuitive because a more informative/conservative reporting system makes the investor react more strongly (and negatively) to an L signal which she believes to be less likely due to noise. This pattern is also evident in the contour plot in which both c and d vary from .50 to 1 for the case of $s = 3$ (Figure 3, Panel F).

Our "comparative statics" analysis based on simulated data is based on one-dimensional metrics of the properties of the reporting system. The implications of the simulation analysis is limited to the extent that both informativeness and conservatism of the reporting system are concepts that cannot typically be summarized by a complete ordering based on some scalars. In addition, it is difficult to separate informativeness from conservatism, because a higher c is associated with both greater informativeness and greater conservatism (see Remarks 1 and 2).

4.3 Discussion

5 Concluding remarks

This paper proposes a framework for understanding the role of reporting systems in accounting-based stock return regularities. The model generates predictions consistent with several empirical regularities on the relationship between accounting information and stock returns. We also discuss how stock return dynamics vary with the properties of the reporting system.

There are at least two limitations to the current study. First, the current framework features an exogenous reporting system. Therefore, we do not address how accounting reports could be influenced by accounting standards or managerial discretion (e.g., Chen et al. 2007). Second, we study a particular type of investor learning. However, other schemes of learning, including seemingly irrational beliefs such as limited memory and regime-shifting beliefs (e.g., Barberis et al. 1998), may also produce similar empirical regularities.

Appendix: Proofs

Proof of Lemma 1. By Bayes' Rule,

$$\begin{aligned}
&\mu_t|_{y_t=L} = \Pr(x_t = B|y_t = L, \boldsymbol{y}^{t-1})\\
=&\frac{\Pr(y_t = L|x_t = B)\Pr(x_t = B|\boldsymbol{y}^{t-1})}{\Pr(y_t = L|x_t = B, \boldsymbol{y}^{t-1})}\\
=&\frac{\sum_{x_{t-1}=B,G}\Pr(y_t = L|x_t = B)\Pr(x_t = B|x_{t-1})\Pr(x_{t-1}|\boldsymbol{y}^{t-1})}{\sum_{x_t=B,G}\sum_{x_{t-1}=B,G}\Pr(y_t = L|x_t)\Pr(x_t|x_{t-1})\Pr(x_{t-1}|\boldsymbol{y}^{t-1})}\\
=&\frac{ca\mu_{t-1} + c(1-b)(1-\mu_{t-1})}{(ca\mu_{t-1} + c(1-b)(1-\mu_{t-1})) + ((1-d)(1-a)\mu_{t-1} + (1-d)b(1-\mu_{t-1}))}\\
=&\frac{c((a+b-1)\mu_{t-1} + 1 - b)}{(c+d-1)(a+b-1)\mu_{t-1} + (1-d)b + c(1-b)}.
\end{aligned} \tag{A.1}$$

The expression for $\mu_t|_{y_t=H}$ can be derived analogously.

Proof of Proposition 1. For notational brevity, let $g_L(\mu)$ and $g_H(\mu)$ denote the belief updating functions given the belief from last period being $\mu_t = \mu$:

$$g_L(\mu) = \mu_{t+1}|_{y_{t+1}=L,\mu_t=\mu}, \tag{A.2}$$

$$g_H(\mu) = \mu_{t+1}|_{y_{t+1}=H,\mu_t=\mu}. \tag{A.3}$$

Suppose Conditions 1-2 hold, i.e., $a + b \geq 1$ and $c + d \geq 1$. It is straightforward to show that both $g_L(\mu)$ and $g_H(\mu)$ are increasing functions of μ on $[0, 1]$:

$$g_L' = \frac{c(1-d)(a+b-1)}{(c((a-1)\mu+1) + (a-1)(d-1)\mu + b(\mu-1)(c+d-1))^2} \geq 0, \tag{A.4}$$

$$g_H' = \frac{(1-c)d(a+b-1)}{(ac\mu + ad\mu - a\mu + b(\mu-1)(c+d-1) - c\mu + c - d\mu + \mu - 1)^2} \geq 0. \tag{A.5}$$

Additionally, we can show that $g^L(\mu)$ is a concave function on $[0, 1]$, i.e.,

$$g_L'' = \frac{2c(d-1)(a+b-1)^2(c+d-1)}{(c((a-1)\mu+1) + (a-1)(d-1)\mu + b(\mu-1)(c+d-1))^3} \leq 0; \tag{A.6}$$

and that $g^H(\mu)$ is a convex function for $\mu \in [0, 1]$, i.e.,

$$g_H''(\mu) = \frac{2(c-1)d(a+b-1)^2(c+d-1)}{(ac\mu + ad\mu - a\mu + b(\mu-1)(c+d-1) - c\mu + c - d\mu + \mu - 1)^3} \geq 0. \tag{A.7}$$

It is also easy to show that $g_L(\mu) \geq \mu$ is equivalent to

$$0 \leq \mu \leq \bar{\mu} \tag{A.8}$$

where $\bar{\mu} = \frac{1}{2}\left(\sqrt{\frac{c(a^2c+4a(d-1)-4d+4)-2(a-2)bc(d-1)+b^2(d-1)^2}{(a+b-1)^2(c+d-1)^2}} + \frac{(a-2)c+b(2c+d-1)}{(a+b-1)(c+d-1)}\right)$. Similarly, $g_H(\mu) \leq \mu$ is equivalent to

$$\underline{\mu} \leq \mu \leq 1 \tag{A.9}$$

where $\underline{\mu} = \frac{1}{2}\left(\frac{a(c-1)+b(2c+d-2)-2c+2}{(a+b-1)(c+d-1)} - \sqrt{\frac{a^2(c-1)^2-2a(b-2)(c-1)d+d(b^2d+4b(c-1)-4c+4)}{(a+b-1)^2(c+d-1)^2}}\right)$. It is also easy to verify that $\underline{\mu} \leq \bar{\mu}$ holds. Therefore, for $\forall \mu \in [\underline{\mu}, \bar{\mu}]$, we have $\mu \leq g_L(\mu) \leq \bar{\mu}$ and $\underline{\mu} \leq g_H(\mu) \leq \mu$. In other words, if we start from $\mu_t \in [\underline{\mu}, \bar{\mu}]$, μ_{t+1} will always stay within the range $[\underline{\mu}, \bar{\mu}]$. For $\forall \mu \in [0, \underline{\mu})$, we have $g_L(\mu) > \mu$ and $g_H(\mu) > \mu$. In other words, if we start from $\mu_t \in [0, \underline{\mu})$, μ_{t+1} will eventually rise to the range $[\underline{\mu}, \bar{\mu}]$. For $\forall \mu \in (\bar{\mu}, 1]$, we have $g_L(\mu) < \mu$ and $g_H(\mu) < \mu$. In other words, if we start from $\mu_t \in (\bar{\mu}, 1]$, μ_{t+1} will eventually fall to the interval $[\underline{\mu}, \bar{\mu}]$.

Proof of Corollary 1. It can be easily verified by checking the signs of the derivatives given Conditions 1 and 2.

Proof of Lemma 2. We only prove the expression for $\mu_t|_{y_t=L,z_t=L}$. By Bayes' Rule,

$$\begin{aligned}
&\mu_t|_{y_t=L,z_t=L} \\
=&\Pr(x_t = B|y_t = L, z_t = L, \boldsymbol{y}^{t-1}, \boldsymbol{z}^{t-1}) \\
=&\frac{\Pr(y_t = L, z_t = L|x_t = B)\Pr(x_t = B|\boldsymbol{y}^{t-1}, \boldsymbol{z}^{t-1})}{\Pr(y_t = L, z_t = L|x_t = B, \boldsymbol{y}^{t-1}, \boldsymbol{z}^{t-1})} \\
=&\frac{\sum_{x_{t-1}=B,G}\Pr(y_t = L, z_t = L|x_t = B)\Pr(x_t = B|x_{t-1})\Pr(x_{t-1}|\boldsymbol{y}^{t-1})}{\sum_{x_t=B,G}\sum_{x_{t-1}=B,G}\Pr(y_t = L, z_t = L|x_t)\Pr(x_t|x_{t-1})\Pr(x_{t-1}|\boldsymbol{y}^{t-1})} \\
=&\frac{c_1c_2a\mu_{t-1} + c_1c_2(1-b)(1-\mu_{t-1})}{(c_1c_2a\mu_{t-1} + c_1c_2(1-b)(1-\mu_{t-1})) + ((1-d_1)(1-d_2)(1-a)\mu_{t-1} + (1-d_1)(1-d_2)b(1-\mu_{t-1}))} \\
=&\frac{c_1c_2((a+b-1)\mu_{t-1} + 1 - b)}{c_1c_2((a+b-1)\mu_{t-1} + 1 - b) + (1-d_1)(1-d_2)((1-a-b)\mu_{t-1} + b)}.
\end{aligned} \tag{A.10}$$

The other three scenarios, $\mu_t|_{y_t=L,z_t=H}$, $\mu_t|_{y_t=H,z_t=L}$, and $\mu_t|_{y_t=H,z_t=H}$, can be proved analogously.

Proof of Proposition 2. Define $g_{ij}(\mu) \equiv \mu_{t+1}|_{y_{t+1}=i,z_{t+1}=j,\mu_t=\mu}$, where $i, j \in \{L, H\}$. It is easy to show that given Conditions 1 and 3, i.e., for $(a, b, c_1, d_1, c_2, d_2)$ such that $a, b, c_1, d_1, c_2, d_2 \in (0, 1)$, $a + b > 1$, $c_1 + d_1 > 1$ and $c_2 + d_2 > 1$, the following two inequalities always hold:

$$g_{HH}(\mu) < g_{LH}(\mu) < g_{LL}(\mu), \tag{A.11}$$

$$g_{HH}(\mu) < g_{HL}(\mu) < g_{LL}(\mu). \tag{A.12}$$

Therefore, in order to bound the value of μ_t, we only need to consider the bounds of $g_{LL}(\mu)$ and

$g_{HH}(\mu)$. We can show that $g_{LL}(\mu) \geq \mu$ is equivalent to

$$0 \leq \mu \leq \mu^{**} \tag{A.13}$$

where μ^{**} is given by

$$\mu^{**} = \frac{1}{2}\left(\sqrt{\frac{c_1c_2\left(a^2c_1c_2 + 4a(d_1(-d_2) + d_1 + d_2 - 1) + 4(d_1 - 1)(d_2 - 1)\right) + 2(a-2)bc_1c_2(d_1 - 1)(d_2 - 1) + b^2(d_1 - 1)^2(d_2 - 1)^2}{(a+b-1)^2(c_1c_2 - d_1d_2 + d_1 + d_2 - 1)^2}} + \frac{(a-2)c_1c_2 + b(2c_1c_2 - d_1d_2 + d_1 + d_2 - 1)}{(a+b-1)(c_1c_2 - d_1d_2 + d_1 + d_2 - 1)}\right), \tag{A.14}$$

and that $g_{HH}(\mu) \leq \mu$ is equivalent to

$$\mu^* \leq \mu \leq 1 \tag{A.15}$$

where μ^* is given by

$$\mu^* = \frac{1}{2}\left(\frac{a(c_1 - 1)(c_2 - 1) + b(2c_1(c_2 - 1) - 2c_2 - d_1d_2 + 2) + 2(c_1(-c_2) + c_1 + c_2 - 1)}{(a+b-1)(c_1(c_2 - 1) - c_2 - d_1d_2 + 1)} - \sqrt{\frac{a^2(c_1 - 1)^2(c_2 - 1)^2 + 2a(b-2)(c_1 - 1)(c_2 - 1)d_1d_2 + d_1d_2\left(b^2d_1d_2 + 4b(c_1(-c_2) + c_1 + c_2 - 1) + 4(c_1 - 1)(c_2 - 1)\right)}{(a+b-1)^2(c_1(-c_2) + c_1 + c_2 + d_1d_2 - 1)^2}}\right). \tag{A.16}$$

The conditions $g_{LL}(\mu) \geq \mu$ and $g_{HH}(\mu) \leq \mu$ ensure that if $y_{t+1} = L$ and $z_{t+1} = L$, then $\mu_{t+1} \geq \mu_t$; if $y_{t+1} = H$ and $z_{t+1} = H$, then $\mu_{t+1} \leq \mu_t$. It is also easy to show that the second order conditions for concavity/convexity are met. For $\forall a, b, c, d$ such that Conditions 1 and 3 hold, we have

$$g''_{LL}(\mu) = -\frac{2c_1c_2(d_1 - 1)(d_2 - 1)(a+b-1)^2(c_1c_2 - d_1d_2 + d_1 + d_2 - 1)}{(c_1c_2((a-1)\mu + 1) - (a-1)(d_1 - 1)(d_2 - 1)\mu + b(p-1)(c_1c_2 - d_1d_2 + d_1 + d_2 - 1))^3} \leq 0, \tag{A.17}$$

$$g''_{HH}(\mu) = \frac{2(c_1 - 1)(c_2 - 1)d_1d_2(a+b-1)^2(c_1(-c_2) + c_1 + c_2 + d_1d_2 - 1)}{(c_1(c_2 - 1)((a-1)\mu + 1) - ac_2\mu - ad_1d_2\mu + ap + b(\mu - 1)(c_1(c_2 - 1) - c_2 - d_1d_2 + 1) + c_2\mu - c_2 + d_1d_2\mu - \mu + 1)^3} \geq 0 \tag{A.18}$$

The rest of the proof is analogous to that of Proposition 1.

Proof of Corollary 2. It can be easily verified by checking the signs of the derivatives given Conditions 1 and 3.

Proof of Lemma 3. The risk-neutral cum-dividend price is equal to the discounted expected future cash flows. Note that $\frac{1}{1+\delta} < 1$ ensures the convergence of $I + \frac{1}{1+\delta}P + \frac{1}{(1+\delta)^2}P^2 + ... = (I - \frac{1}{1+\delta}P)^{-1}$. Therefore, we have

$$p_t = (\mu_t, 1 - \mu_t)\left(I - \frac{1}{1+\delta}P\right)^{-1}\bar{X} \tag{A.19}$$

where $\bar{X} = (B, G)'$. This can be easily rewritten as

$$p_t = \frac{1+\delta}{\delta(2-a-b+\delta)}\big(G(1-a+\delta) + B(1-b) - \delta(G-B)\mu_t\big). \tag{A.20}$$

Proof of Lemma 4. To prove equation (3.1), we need to derive an expression for $E_t^i[\mu_{t+1}|\mu_t]$. Note that

$$E_t^i[\mu_{t+1}|\mu_t] = \Pr(y_{t+1} = L|\mu_t) \cdot \mu_{t+1}|_{y_{t+1}=L} + \Pr(y_{t+1} = H|\mu_t) \cdot \mu_{t+1}|_{y_{t+1}=H} \tag{A.21}$$

where

$$\begin{aligned}
\Pr(y_{t+1} = L|\mu_t) &= \sum_{x_{t+1}\in\{B,G\}} \sum_{x_t\in\{B,G\}} \Pr(y_{t+1} = L|x_{t+1})\Pr(x_{t+1}|x_t)\Pr(x_t|\mu_t) \\
&= \Big(\mu_t a + (1-\mu_t)(1-b)\Big)c + \Big(\mu_t(1-a) + (1-\mu_t)b\Big)(1-d), && \text{(A.22)} \\
\Pr(y_{t+1} = H|\mu_t) &= \sum_{x_{t+1}\in\{B,G\}} \sum_{x_t\in\{B,G\}} \Pr(y_{t+1} = L|x_{t+1})\Pr(x_{t+1}|x_t)\Pr(x_t|\mu_t) \\
&= \Big(\mu_t a + (1-\mu_t)(1-b)\Big)(1-c) + \Big(\mu_t(1-a) + (1-\mu_t)b\Big)d, && \text{(A.23)} \\
\mu_{t+1}|_{y_{t+1}=L} &= \frac{c(a+b-1)\mu_t + c(1-b)}{(c+d-1)(a+b-1)\mu_t + (1-d)b + c(1-b)}, && \text{(A.24)} \\
\mu_{t+1}|_{y_{t+1}=H} &= \frac{(1-c)(a+b-1)\mu_t + (1-c)(1-b)}{(1-c-d)(a+b-1)\mu_t + db + (1-c)(1-b)}. && \text{(A.25)}
\end{aligned}$$

It follows that

$$E_t^i[\mu_{t+1}|\mu_t] = a\mu_t + (1-b)(1-\mu_t). \tag{A.26}$$

Based on the pricing function, it is then straightforward to show that the expected price change from an investor's perspective is

$$E_t^i[p_{t+1} - p_t] = \frac{1+\delta}{2-a-b+\delta}(G-B)\big((2-a-b)\mu_t - 1 + b\big). \tag{A.27}$$

For an empiricist, the expectation of current period belief reflects the sample mean. Therefore,

$$E_t^f[p_{t+1} - p_t] = \frac{1+\delta}{2-a-b+\delta}(G-B)\big((2-a-b)\pi_B - 1 + b\big) = 0. \tag{A.28}$$

Proof of Proposition 3. The difference in average market reactions to difference signals is given by

$$\begin{aligned}
&E_t^f[p_t - p_{t-1}|y_t = H] - E_t^f[p_t - p_{t-1}|y_t = L] \\
=&- \frac{1+\delta}{2-a-b+\delta}(G-B)\Big(E_t^f[\mu_t - \mu_{t-1}|y_t = H] - E_t^f[\mu_t - \mu_{t-1}|y_t = L]\Big) \\
=&- \frac{1+\delta}{2-a-b+\delta}(G-B)\int_{\mu_{t-1}}\Big(g_H(\mu_{t-1}) - g_L(\mu_{t-1})\Big)dF(\mu_{t-1}) \\
\geq& 0. \qquad \text{(A.29)}
\end{aligned}$$

The last inequality obtains because the integrand is negative for $\mu_{t-1} \in (\underline{\mu}, \bar{\mu})$.

Proof of Proposition 4. Suppose the earnings string is of length $s = 1$, i.e., $y_t = L, y_{t-1} = H$. The conditional expectation of price change for an empiricist is

$$\begin{aligned}
&E_t^f[p_t - p_{t-1}|y_t = L, y_{t-1} = H] \\
=&- \frac{1+\delta}{2-a-b+\delta}(G-B)E_t^f[\mu_t - \mu_{t-1}|y_t = L, y_{t-1} = H] \\
=&- \frac{1+\delta}{2-a-b+\delta}(G-B)\int_{\mu_{t-1}}(g_L(\mu_{t-1}) - \mu_{t-1})\, dF(\mu_{t-1}|y_{t-1} = H) \\
\leq& 0, \qquad \text{(A.30)}
\end{aligned}$$

because $g_L(\mu_{t-1}) \geq \mu_{t-1}$ for $\mu_{t-1} \in [\underline{\mu}, \bar{\mu}]$. Suppose the earnings string is of length $s = 2$, i.e., $y_t = L, y_{t-1} = y_{t-2} = H$. The conditional expectation of price change for an empiricist is

$$\begin{aligned}
&E_t^f[p_t - p_{t-1}|y_t = L, y_{t-1} = y_{t-2} = H] \\
=&- \frac{1+\delta}{2-a-b+\delta}(G-B)E_t^f[\mu_t - \mu_{t-1}|y_t = L, y_{t-1} = y_{t-2} = H] \\
=&- \frac{1+\delta}{2-a-b+\delta}(G-B)\int_{\mu_{t-1}}(g_L(\mu_{t-1}) - \mu_{t-1})\, dF(\mu_{t-1}|y_{t-1} = y_{t-2} = H) \qquad \text{(A.31)} \\
\leq& 0. \qquad \text{(A.32)}
\end{aligned}$$

Now we would like to find conditions for

$$E_t^f[p_t - p_{t-1}|y_t = L, y_{t-1} = H] \geq E_t^f[p_t - p_{t-1}|y_t = L, y_{t-1} = y_{t-2} = H] \qquad \text{(A.33)}$$

This is equivalent to

$$\int_{\mu_{t-1}}(g_L(\mu_{t-1}) - \mu_{t-1})\, dF(\mu_{t-1}|y_{t-1} = H) \leq \int_{\mu_{t-1}}(g_L(\mu_{t-1}) - \mu_{t-1})\, dF(\mu_{t-1}|y_{t-1} = y_{t-2} = H) \qquad \text{(A.34)}$$

Define $F^H(\mu_{t-1}) = F(\mu_{t-1}|y_{t-1} = H)$ and $F^{HH}(\mu_{t-1}) = F(\mu_{t-1}|y_{t-1} = y_{t-2} = H)$. By integration by parts, the left hand side of (B.34) is

$$
\begin{aligned}
&\int_{\mu_{t-1}} \left(g_L(\mu_{t-1}) - \mu_{t-1}\right) dF(\mu_{t-1}|y_{t-1} = H) \\
=& \underbrace{\left(g_L(\bar{\mu}) - \bar{\mu}\right)}_{=0} \underbrace{F^H(\bar{\mu})}_{=1} - \left(g_L(\underline{\mu}) - \underline{\mu}\right) \underbrace{F^H(\underline{\mu})}_{=0} - \int_{\underline{\mu}}^{\bar{\mu}} F^H(\mu_{t-1}) \left(g'_L(\mu_{t-1}) - 1\right) d\mu_{t-1} \\
=& - \int_{\underline{\mu}}^{\bar{\mu}} \left(g'_L(\mu_{t-1}) - 1\right) F^H(\mu_{t-1}) d\mu_{t-1}. \qquad \text{(A.35)}
\end{aligned}
$$

Similarly, the right hand side is

$$
\int_{\mu_{t-1}} \left(g_L(\mu_{t-1}) - \mu_{t-1}\right) dF(\mu_{t-1}|y_{t-1} = y_{t-2} = H) = - \int_{\underline{\mu}}^{\bar{\mu}} \left(g'_L(\mu_{t-1}) - 1\right) F^{HH}(\mu_{t-1}) d\mu_{t-1}. \qquad \text{(A.36)}
$$

Therefore, inequality (B.34) becomes

$$
\int_{\underline{\mu}}^{\bar{\mu}} \left(g'_L(\mu_{t-1}) - 1\right) F^H(\mu_{t-1}) d\mu_{t-1} \geq \int_{\underline{\mu}}^{\bar{\mu}} \left(g'_L(\mu_{t-1}) - 1\right) F^{HH}(\mu_{t-1}) d\mu_{t-1}. \qquad \text{(A.37)}
$$

A sufficient condition for (B.39) is

$$
\left(g'_L(\mu_{t-1}) - 1\right) F^H(\mu_{t-1}) \geq \left(g'_L(\mu_{t-1}) - 1\right) F^{HH}(\mu_{t-1}) \qquad \text{(A.38)}
$$

for $\forall \mu_{t-1}$. In the following, we first prove

$$
F^H(\mu_{t-1}) \leq F^{HH}(\mu_{t-1}), \qquad \text{(A.39)}
$$

and derive a sufficient condition for the above inequality to hold. By definition, for a given μ_{t-3},

$$
\begin{aligned}
F^H(\mu_{t-1}) =& F(\mu_{t-1}|y_{t-1} = H) = \Pr(g_H(\mu_{t-2}) \leq \mu_{t-1}) = \Pr(\mu_{t-2} \leq g_H^{-1}(\mu_{t-1})) \\
=& \gamma \Pr(g_L(\mu_{t-3}) \leq g_H^{-1}(\mu_{t-1})) + (1-\gamma) \Pr(g_H(\mu_{t-3}) \leq g_H^{-1}(\mu_{t-1})) \qquad \text{(A.40)}
\end{aligned}
$$

where $\gamma = \Pr(y_{t-2} = L)$;

$$
\begin{aligned}
F^{HH}(\mu_{t-1}) =& F(\mu_{t-1}|y_{t-1} = y_{t-2} = H) = \Pr(g_H(g_H(\mu_{t-3})) \leq \mu_{t-1}) \\
=& \Pr(g_H(\mu_{t-3}) \leq g_H^{-1}(\mu_{t-1})). \qquad \text{(A.41)}
\end{aligned}
$$

Because $g_L(\mu_{t-3}) \geq g_H(\mu_{t-3})$ for $\forall \mu_{t-3}$, it follows that

$$
\Pr(g_L(\mu_{t-3}) \leq g_H^{-1}(\mu_{t-1})) \leq \Pr(g_H(\mu_{t-3}) \leq g_H^{-1}(\mu_{t-1})), \qquad \text{(A.42)}
$$

and therefore

$$F^H(\mu_{t-1}) \leq F^{HH}(\mu_{t-1}) \tag{A.43}$$

for $\forall \mu_{t-1} \in [\underline{\mu}, \bar{\mu}]$. In other words, F^H first-order stochastically dominates (f.o.s.d.) F^H. Given (B.43), a sufficient condition for (B.38) to hold is

$$g'_L(\mu_{t-1}) - 1 \leq 0. \tag{A.44}$$

for $\forall \mu_{t-1}$. It is easy to show that (B.44) is met when either of the following two conditions (in addition to Conditions 1 and 2) is true: (i) $0 \leq b \leq b^*$, where $b^* = \frac{c\left((2c+d-1)-\sqrt{(1-d)(4c+3d-3)}\right)}{2(c+d-1)^2}$; (ii) $b^* \leq b \leq 1$ and $0 \leq a \leq a^*$, where $a^* = \frac{b^2(c+d-1)^2-bc(2c+d-1)+c(c-d+1)}{c(1-d)}$.

Now let us consider longer earnings strings. Note that based on the proof of (), analogous arguments for longer sequences of H signals can be proved by induction. For example, F^{HH} first-order stochastically dominates (f.o.s.d.) F^{HHH}, and so forth. As the earnings string becomes infinitely long, we have

$$\lim_{s\to\infty} \mu_{t-1} = \underline{\mu} \tag{A.45}$$

and

$$\lim_{s\to\infty} F^{HH...H}(\mu_{t-1}) = 1. \tag{A.46}$$

As a result,

$$\begin{aligned}
&\lim_{s\to\infty} E_t^f[p_t - p_{t-1}|y_t = L, y_{t-1} = ... = y_{t-s} = H] = E_t^f[p_t - p_{t-1}|y_t = L, \mu_{t-1} = \underline{\mu}] \\
&= -\frac{1+\delta}{2-a-b+\delta}(G-B)\Big(\frac{c(a+b-1)\underline{\mu} + c(1-b)}{(c+d-1)(a+b-1)\underline{\mu} + (1-d)b + c(1-b)} - \underline{\mu}\Big).
\end{aligned} \tag{A.47}$$

Analogously, if the investor observes a sequence of L signals before seeing a break (a G signal), the market reaction will be

$$\begin{aligned}
&\lim_{s\to\infty} E_t^f[p_t - p_{t-1}|y_t = H, y_{t-1} = ... = y_{t-s} = L] = E_t^f[p_t - p_{t-1}|y_t = H, \mu_{t-1} = \bar{\mu}] \\
&= -\frac{1+\delta}{2-a-b+\delta}(G-B)\Big(\frac{(1-c)(a+b-1)\bar{\mu} + (1-c)(1-b)}{(1-c-d)(a+b-1)\bar{\mu} + db + (1-c)(1-b)} - \bar{\mu}\Big).
\end{aligned} \tag{A.48}$$

Proof of Proposition 5. Recall the price change is given by

$$p_{t+1} - p_t = -\frac{1+\delta}{2-a-b+\delta}(G-B)(\mu_{t+1} - \mu_t). \tag{A.49}$$

where μ_{t+1} is given by (2.3) and (2.4). We also know from (B.26) that $E_t^i[\mu_{t+1}|\mu_t] = a\mu_t + (1 -$

$b)(1-\mu_t)$. Therefore,

$$E_t^i(p_{t+1}-p_t|y_t=L,\mu_t=\mu)=\frac{1+\delta}{2-a-b+\delta}(G-B)\Big((2-a-b)\mu-1+b\Big), \tag{A.50}$$

$$E_t^i(p_{t+1}-p_t|y_t=H,\mu_t=\mu)=\frac{1+\delta}{2-a-b+\delta}(G-B)\Big((2-a-b)\mu-1+b\Big). \tag{A.51}$$

Let $\phi(\mu)=\frac{1+\delta}{2-a-b+\delta}(G-B)\Big((2-a-b)\mu-1+b\Big)$. Note that $\phi(\mu)$ is a linear increasing function of μ,

$$\phi'(\mu)=\frac{(1+\delta)(2-a-b)}{2-a-b+\delta}(G-B)>0. \tag{A.52}$$

The expected price changes given current-period signals can be expressed as

$$E_t^f(p_{t+1}-p_t|y_t=L)=\int_{\mu_t}\phi(\mu_t)dF(\mu_t|y_t=L), \tag{A.53}$$

$$E_t^f(p_{t+1}-p_t|y_t=H)=\int_{\mu_t}\phi(\mu_t)dF(\mu_t|y_t=H). \tag{A.54}$$

Based on the fact that $\phi(\mu)$ is a linear increasing function of μ, the inequality $E_t^f(p_{t+1}-p_t|y_t=L)\geq E_t^f(p_{t+1}-p_t|y_t=H)$ is equivalent to

$$\int_{\mu_t}\mu_t dF^L(\mu_t)\geq\int_{\mu_t}\mu_t dF^H(\mu_t). \tag{A.55}$$

Note that the distribution of μ_t conditional on signals are different between L and H. Also, with probability 1, $\mu_t\in[\underline{\mu},\bar{\mu}]$. Therefore, evaluating the integration over the full support of $[0,1]$ is equivalent to evaluating the integration over $[\underline{\mu},\bar{\mu}]$. Denote $F^H(\mu_t)=F(\mu_t|y_t=H)$ and $F^L(\mu_t)=F(\mu_t|y_t=L)$. We know, from Proposition 1, that $F^H(\underline{\mu})=F^L(\underline{\mu})=0$ and $F^H(\bar{\mu})=F^L(\bar{\mu})=1$. Therefore, we have

$$\begin{aligned}\int_{\mu_t}\mu_t dF^j(\mu_t)=\int_{\underline{\mu}}^{\bar{\mu}}\mu_t dF^j(\mu_t)&=\bar{\mu}\cdot F^j(\bar{\mu})-\underline{\mu}\cdot F^j(\underline{\mu})-\int_{\underline{\mu}}^{\bar{\mu}}F^j(\mu_t)d\mu_t\\&=\bar{\mu}-\int_{\underline{\mu}}^{\bar{\mu}}F^j(\mu_t)d\mu_t.\end{aligned} \tag{A.56}$$

for $j\in\{L,H\}$. (B.57) is equivalent to $\int_{\underline{\mu}}^{\bar{\mu}}F^L(\mu_t)d\mu_t\leq\int_{\underline{\mu}}^{\bar{\mu}}F^H(\mu_t)d\mu_t$. In the following, we prove a sufficient condition of the above inequality, namely, $F^L(\mu_t)\leq F^H(\mu_t)$ for $\forall\mu_t$. By definition, for a given μ_{t-1},

$$F^L(\mu)=F(\mu_t|y_t=L)=\Pr(g_L(\mu_{t-1})\leq\mu), \tag{A.57}$$

$$F^H(\mu)=F(\mu_t|y_t=H)=\Pr(g_H(\mu_{t-1})\leq\mu). \tag{A.58}$$

Given that $g_L(\mu_{t-1}) \geq g_H(\mu_{t-1})$ for $\forall \mu_{t-1}$, it follows that $F^L(\mu) \leq F^H(\mu)$ for $\forall \mu$.

Proof of Proposition 6. The hedge return obtained by trading on the two signals is

$$E_t^f(p_{t+1} - p_t | y_t = H, z_t = L) - E_t^f(p_{t+1} - p_t | y_t = L, z_t = H). \tag{A.59}$$

Note that

$$E_t^f(p_{t+1} - p_t | \mu_t = \mu) = \frac{1+\delta}{2-a-b+\delta}(G-B)\Big((2-a-b)\mu - 1 + b\Big) \equiv \phi(\mu). \tag{A.60}$$

Therefore,

$$E_t^f(p_{t+1} - p_t | y_t = H, z_t = L) = \int_\mu \phi(\mu) dF(\mu | y_t = H, z_t = L), \tag{A.61}$$

$$E_t^f(p_{t+1} - p_t | y_t = L, z_t = H) = \int_\mu \phi(\mu) dF(\mu | y_t = L, z_t = H). \tag{A.62}$$

Recall that

$$\mu_t|_{y_t=H,z_t=L} = \frac{(1-c_1)c_2((a+b-1)\mu_{t-1} + 1 - b)}{(1-c_1)c_2((a+b-1)\mu_{t-1} + 1 - b) + d_1(1-d_2)((1-a-b)\mu_{t-1} + b)}, \tag{A.63}$$

$$\mu_t|_{y_t=L,z_t=H} = \frac{c_1(1-c_2)((a+b-1)\mu_{t-1} + 1 - b)}{c_1(1-c_2)((a+b-1)\mu_{t-1} + 1 - b) + (1-d_1)d_2((1-a-b)\mu_{t-1} + b)}. \tag{A.64}$$

Note that with probability 1, $\mu_t \in [\mu^*, \mu^{**}]$. Therefore, evaluating the integration over the full support of $[0,1]$ is equivalent to evaluating the integration over $[\mu^*, \mu^{**}]$. Denote $F^{HL}(\mu_t) = F(\mu_t | y_t = H, z_t = L)$ and $F^{LH}(\mu_t) = F(\mu_t | y_t = L, z_t = H)$, we know that $F^{HL}(\mu^*) = F^{LH}(\mu^*) = 0$ and $F^{HL}(\mu^{**}) = F^{LH}(\mu^{**}) = 1$. Therefore, analogous to equation (B.58), we have

$$\int_{\mu_t} \mu_t dF^{HL}(\mu_t) = \mu^{**} - \int_{\mu^*}^{\mu^{**}} F^{HL}(\mu_t) d\mu_t. \tag{A.65}$$

The inequality $\int_{\mu_t} \mu_t dF^{HL}(\mu_t) \geq \int_{\mu_t} \mu_t dF^{LH}(\mu_t)$, is equivalent to $\int_{\mu^*}^{\mu^{**}} F^{HL}(\mu_t) d\mu_t \leq \int_{\mu^*}^{\mu^{**}} F^{LH}(\mu_t) d\mu_t$. In the following, we prove a sufficient condition of the above inequality, namely, $F^{HL}(\mu_t) \leq F^{LH}(\mu_t)$ for $\forall \mu_t$. By definition, for a given μ_{t-1},

$$F^{HL}(\mu) = \Pr(\mu^{HL}(\mu_{t-1}) \leq \mu). \tag{A.66}$$

Assume $a, b, c_1, d_1, c_2, d_2 \in [0,1]$, $a + b \geq 1$, $c_1 + d_1 \geq 1$, and $c_2 + d_2 \geq 1$. It is easy to show that

$$\mu_t|_{y_t=H,z_t=L\mu_{t-1}=\mu} \geq \mu_t|_{y_t=L,z_t=H,\mu_{t-1}=\mu} \tag{A.67}$$

iff

$$1 - d_1 \leq c_1 \leq \frac{c_2 d_2 (1 - d_1)}{d_1(1 - d_2) + c_2(d_2 - d_1)} \tag{A.68}$$

But we know that $\forall \mu_{t-1}$, $\mu^{HL}(\mu_{t-1}) \geq \mu^{LH}(\mu_{t-1})$ given the assumptions. Therefore,

$$\Pr(\mu^{HL}(\mu_{t-1}) \leq \mu) \leq \Pr(\mu^{LH}(\mu_{t-1}) \leq \mu)$$

Given that $\phi(\mu)$ is a linear increasing function of μ, we have

$$E_t^f(p_{t+1} - p_t | y_t = H, z_t = L) - E_t^f(p_{t+1} - p_t | y_t = L, z_t = H) \geq 0 \tag{A.69}$$

iff

$$1 - d_1 \leq c_1 \leq \frac{c_2 d_2 (1 - d_1)}{d_1(1 - d_2) + c_2(d_2 - d_1)}. \tag{A.70}$$

where it can be shown that $1 - d_1 \leq \frac{c_2 d_2 (1 - d_1)}{d_1(1-d_2)+c_2(d_2-d_1)}$ always holds.

References

[1] Antle, R., J. Demski, and S. Ryan. 1994. Multiple sources of information, valuation, and accounting earnings. *Journal of Accounting, Auditing and Finance* 9, 675–696.

[2] Antle, R., and R. Lambert. 1988. Accountants' loss functions and induced incentives for conservatism. In *Economic Analysis of Information and Contracts: Essays in Honor of John Butterworth*, edited by G. Feltham, A. Amershi, and W. Ziemba. Boston, MA: Kluwer Academic Publishers.

[3] Ayers, B. 1998. Deferred tax accounting under SFAS No. 109: An empirical investigation of its incremental value-relevance relative to APB No. 11. *The Accounting Review* 73, 195–212.

[4] Barberis, N., A. Shleifer, and R. Vishny. 1998. A model of investor sentiment. *Journal of Financial Economics* 49, 307–345.

[5] Barth, M., J. Elliott, and M. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37, 387–413.

[6] Berger, J. 1985. *Statistical Decision Theory and Bayesian Analysis.* New York: Springer-Verlag.

[7] Blackwell, D. 1951. Comparison of experiments. In *Proceedings of the Second Berkeley Symposinm on Mathematical Statistics and Probability*, edited by J. Neyman. Berkeley: University of California Press.

[8] Chen, Q., T. Hemmer, and Y. Zhang. 2007. On the relation between conservatism in accounting standards and incentives for earnings management. *Journal of Accounting Research* 45, 541–565.

[9] Chi, S., M. Pincus, and S. Teoh. 2014. Mispricing of book-tax differences and the trading behavior of short sellers and insiders. *The Accounting Review* 89, 511–543.

[10] Christensen, J., and J. Demski. 2003. *Accounting Theory: An Information Content Perspective*. McGraw-Hill/Irwin.

[11] Collins, D., and S. Kothari. 1989. An analysis of inter-temporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics* 11, 143–181.

[12] DeAngelo, H., DeAngelo, L., and D. Skinner. 1996. Reversal of fortune: Dividend signaling and the disappearance of sustained earnings growth. *Journal of Financial Economics* 40, 341–371.

[13] Demski, J. 1973. The general impossibility of normative accounting standards. *The Accounting Review* 48, 718–723.

[14] Dhaliwal, D., H. Lee, M. Pincus, and L. Steele. 2017. Taxable income and firm risk. *Journal of the American Taxation Association* 39, 1–24.

[15] Dichev, I., and V. Tang. 2008. Matching and the changing properties of accounting earnings over the last 40 years. *The Accounting Review* 83, 1425–1460.

[16] Du, K. 2016. Investor expectations, earnings management, and asset prices. Working paper, Pennsylvania State University.

[17] Evans, G., and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics.* Princeton, NJ: Princeton University Press.

[18] Fairfield, P., Ramnath, S., and T. Yohn. 2009. Do industry-level analyses improve forecast of financial performance? *Journal of Accounting Research* 47, 147–178.

[19] Fama, E., and K. French. 2000. Forecasting profitability and earnings. *Journal of Business* 73, 161–175.

[20] Fischer, P. M. Heinle, and R. Verrecchia. 2016. Beliefs-driven price association. *Journal of Accounting and Economics* 61, 563–583.

[21] Gjesdal, F. 1981. Accounting for stewardship. *Journal of Accounting Research* 19, 208–231.

[22] Gigler, F., C. Kanodia, H. Sapra, and R. Venugopalan. 2009 Accounting conservatism and the efficiency of debt contracts. *Journal of Accounting Research* 47, 767–797.

[23] Green, J., J. Hand, M. Soliman. 2011. Going, going, gone? The apparent demise of the accruals anomaly. *Management Science* 57, 797–816.

[24] Hanlon, M. 2005. The persistence and pricing of earnings, accruals, and cash flows when firms have large book-tax differences. *The Accounting Review* 80, 137–166.

[25] Hanlon, M., S. Laplante, and T. Shevlin. 2005. Evidence for the possible information loss of conforming book income and taxable income. *Journal of Law and Economics* 48, 407–442.

[26] Jiang, X. 2016. Biases in accounting and nonaccounting information: Substitutes or complements? *Journal of Accounting Research* 54, 1297–1330.

[27] Kaserer, C. and C. Klingler. 2008. The accrual anomaly under different accounting standards: Lessons learned from the German experiment. *Journal of Business Finance and Accounting* 35: 837–859.

[28] Ke, B., S. Huddart, and K. Petroni. 2003. What insiders know about future earnings and howthey use it: Evidence from insider trades. *Journal of Accounting and Economics* 35, 315–346.

[29] Kothari, S. 2001. Capital market research in accounting. *Journal of Accounting Economics* 31, 105–231.

[30] Lev, B., and D. Nissim. 2004. Taxable income, future earnings, and equity values. *The Accounting Review* 79, 1039–1074.

[31] Lewellen, J., and J. Shanken. 2002. Learning, asset-pricing tests, and market efficiency. *Journal of Finance* 57, 1113–1145.

[32] Marschak, J. 1971. Economics of information systems. *Journal of the American Statistical Association* 66, 192–219.

[33] Marschak, J., and K. Miyasawa. 1968. Economic comparability of information systems. *International Economic Review* 9, 137–174.

[34] Richardson, S., R. Sloan, M. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39, 437–485.

[35] Richardson, S., I. Tuna, and P. Wysocki. 2010. Accounting anomalies and fundamental analysis: A review of recent research advances. *Journal of Accounting and Economics* 50, 410–454.

[36] Scott, W. 1979. Scoring rules for probabilistic reporting. *Journal of Accounting Research* 17, 156–178.

[37] Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71, 289–315.

[38] Spiceland, J., J. Sepe, and M. Nelson. 2013. *Intermediate Accounting*. 7th ed. New York, NY: McGraw-Hill Irwin.

[39] Wu, J., L. Zhang, and X. Zhang. 2010. The q-theory approach to understanding the accrual anomaly. *Journal of Accounting Research* 48, 177–223.

[40] Chen, Q., J. Francis, and W. Jiang. 2005. Investor learning about analyst predictive ability. *Journal of Accounting and Economics* 39, 3–24.

[41] Dechow, P., and W. Ge. 2006. The persistence of earnings and cash flows and the role of special items: Implications for the accrual anomaly. *Review of Accounting studies* 11, 253–296.

[42] Dichev, I. 2008. On the balance sheet-based model of financial reporting. *Accounting Horizons* 22, 453–470.

[43] Hansen, L. 2007. Beliefs, doubts and learning: Valuing macroeconomic risk. *American Economic Review* 97, 1–30.

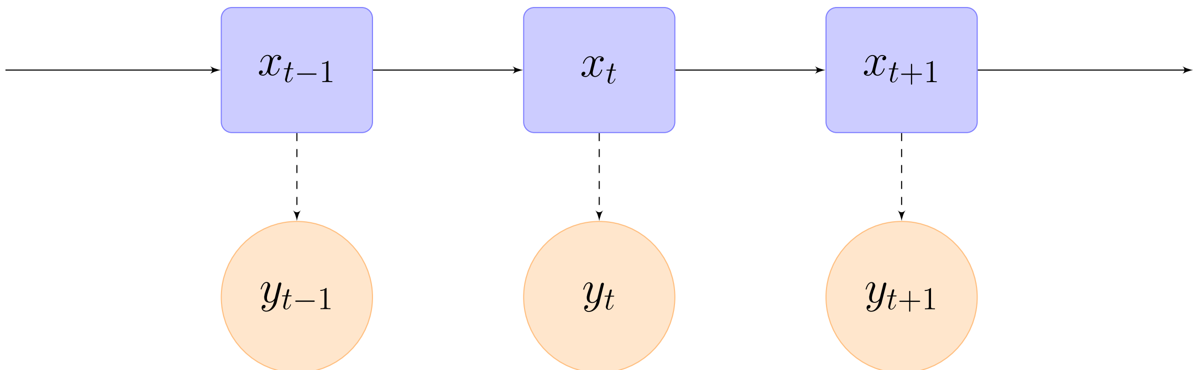
[44] Li, Feng. 2011. Earnings quality based on corporate investment decisions. *Journal of Accounting Research* 49, 721–752.

[45] Mashruwala, C., S. Rajgopal, and T. Shevlin. 2006. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. *Journal of Accounting and Economics* 42, 3–33.

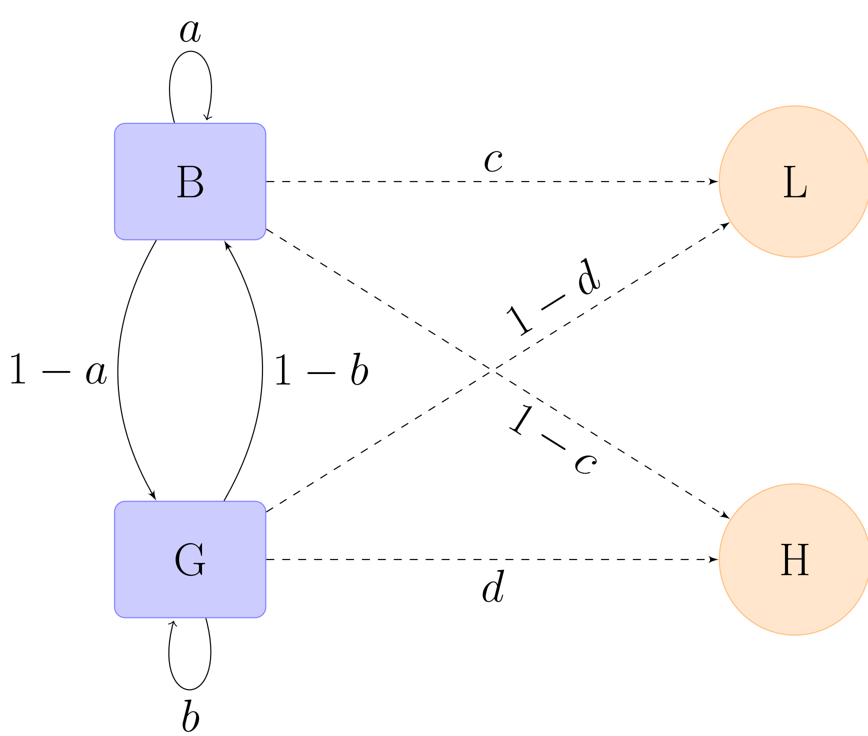
[46] Zhang, X. 2007. Accruals, investment, and the accrual anomaly. *The Accounting Review* 82, 1333–1363.

Figure 1: One-signal reporting system

Panel A illustrates the time-series structure of the model. Panel B illustrates the law of motion (a Markov process) for the underlying state x_t and the reporting system which maps the state to an accounting signal y_t.



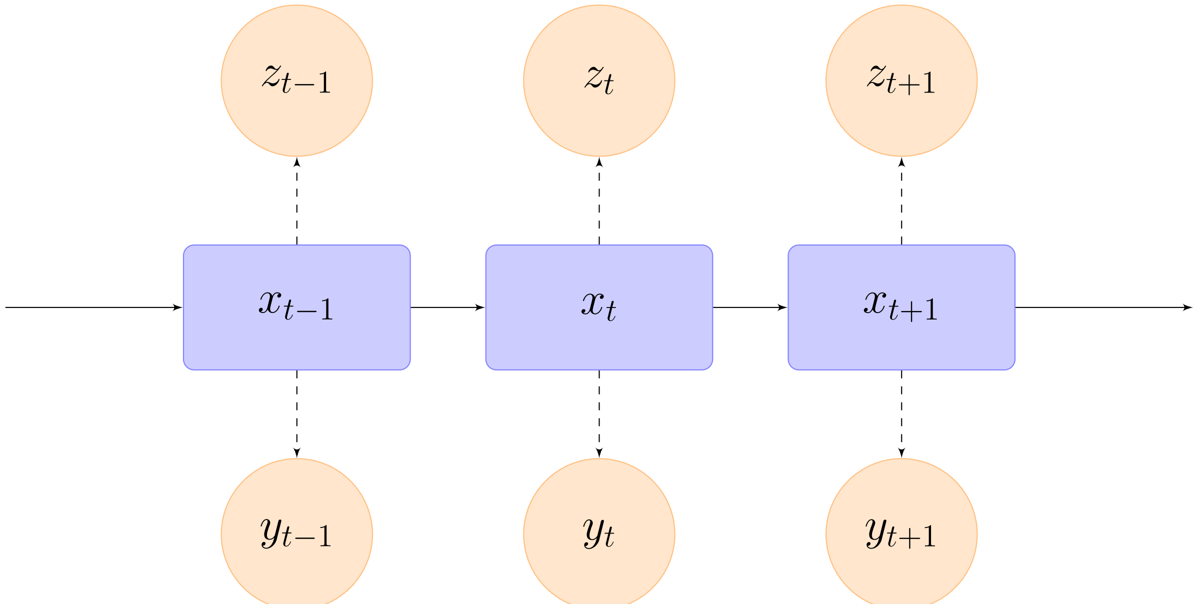
Panel A: The time-series structure



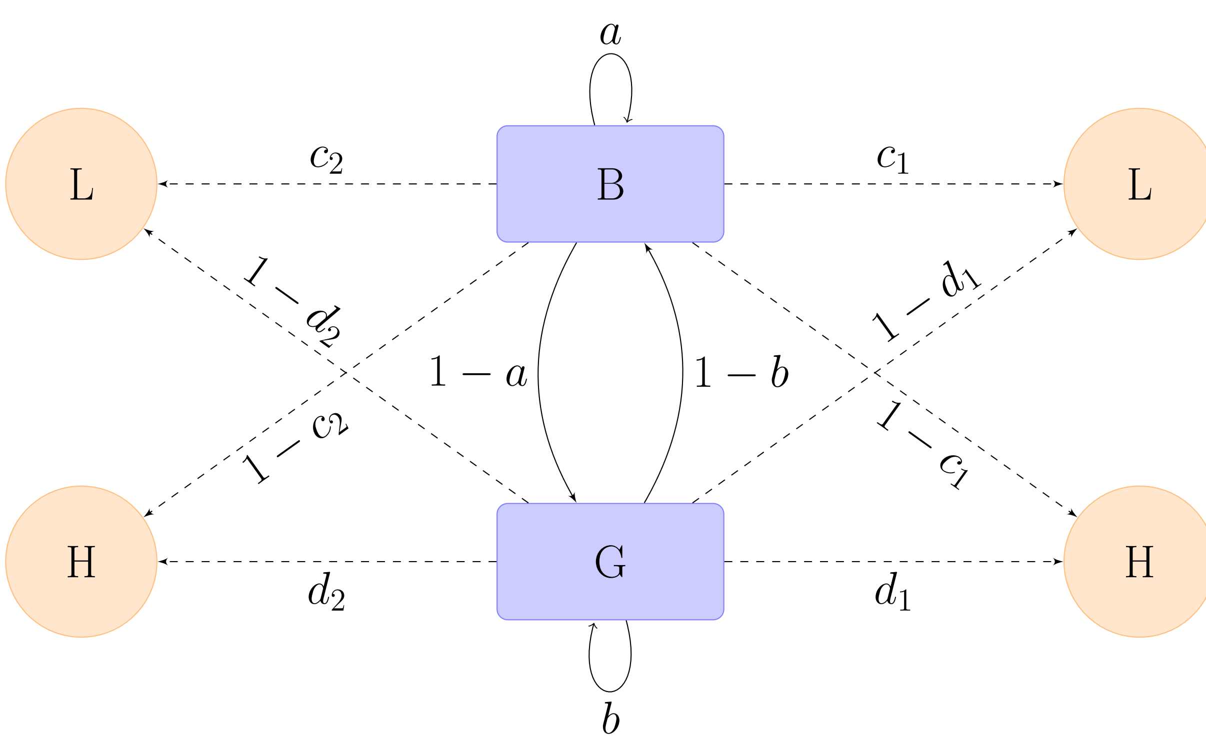
Panel B: The law of motion and the reporting system

Figure 2: Two-signal reporting system

Panel A illustrates the time-series structure of a two-signal reporting system. Panel B illustrates the law of motion (a Markov process) for the underlying state x_t and the reporting system which maps the state to an accounting signals y_t and z_t.



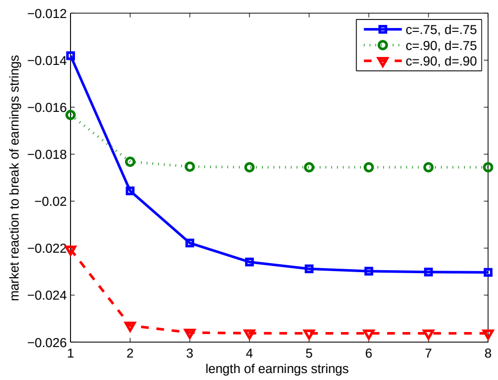
Panel A: The time-series structure



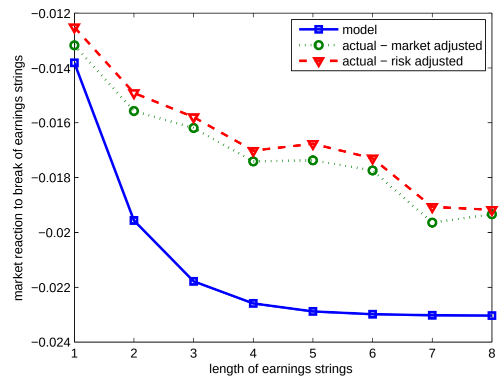
Panel B: The law of motion and the reporting system

Figure 3: Market reaction to breaks of earnings strings

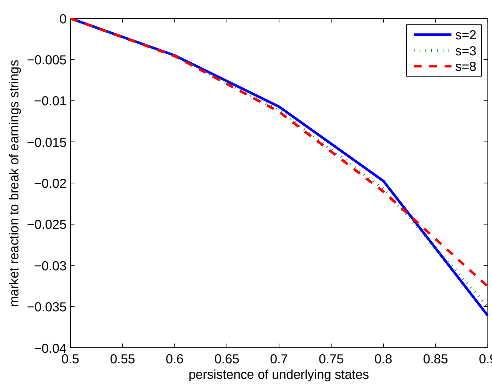
Panel A presents the simulated market reaction to breaks of earnings strings. To calculate the market reaction to breaks, we simulate the model over T=1,000 periods for N=10,000 idiosyncratic histories ("firms"); for each period, we form an equal-weighted portfolio of all firms with an L signal following exactly s=1,2,...,8 consecutive H signals, where s is the length of earnings strings. We plot the average of portfolio returns (and their difference) over all periods for different lengths of the earnings strings. Panel B contrasts the simulated pattern with actual pattern based on empirical data. "Actual" reaction is the average reaction to breaks of earnings strings of lengths 1 to 8, based on Compustat firm-quarters over 1963–2013 after excluding firms with fiscal year changes. Earnings strings are defined as consecutive earnings increases, where an earnings increase is defined as earnings per share (EPS) before extraordinary items in the observation quarter being higher than EPS for the same quarter of the previous year. Panel C varies $a = b \in [.50, 1]$, holding $c = d = .75$; Panel D varies $c = d \in [.50, 1]$, holding $a = b = .75$; Panel E varies $c \in [.50, 1]$, holding $a = b = .75$ and $d = .50$; Panel F plots the contour of equal-differential-returns for combinations of different values of c and d for $c, d \in [.5, 1]$.



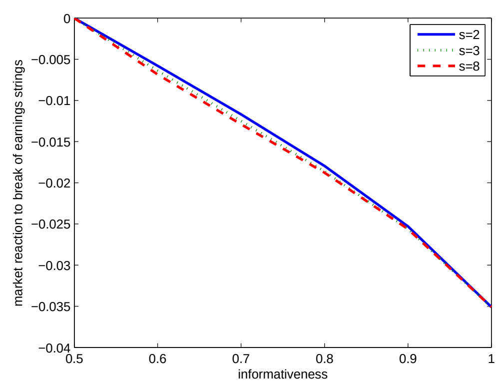
Panel A: Alternative calibrations



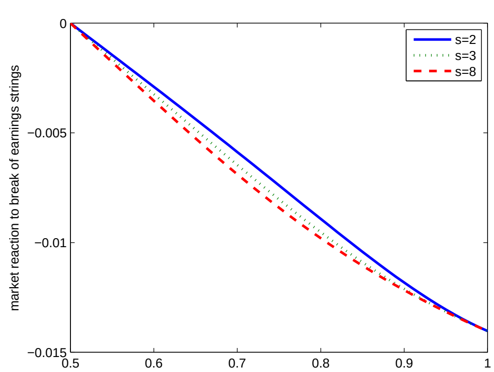
Panel B: Model vs. actual market reaction to breaks



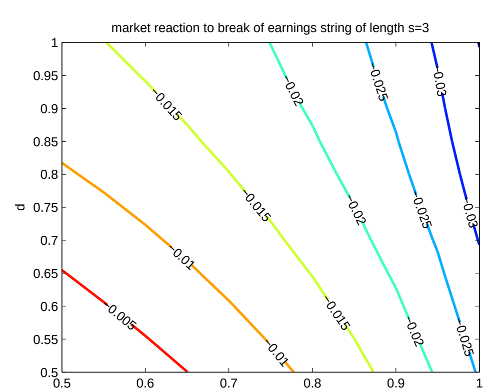
Panel C: Varying persistence



Panel D: Varying informativeness



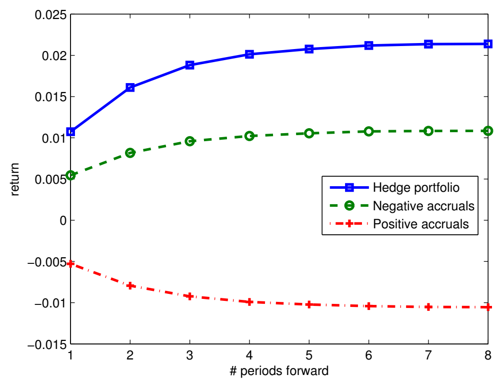
Panel E: Varying conservatism



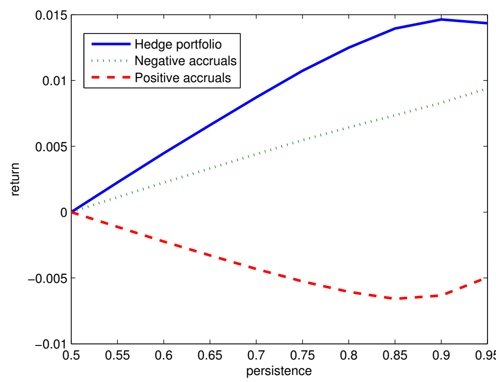
Panel F: Contour of equal reaction (s=3)

Figure 4: Accruals anomaly

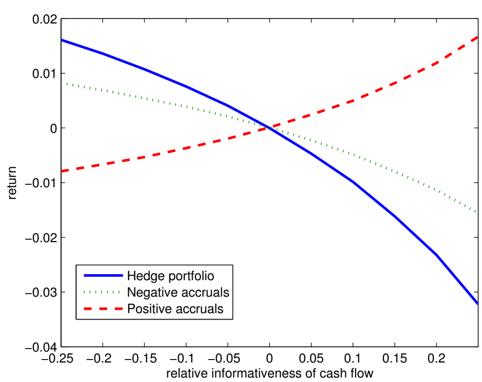
We simulate the model over T=1,000 periods for N=10,000 idiosyncratic histories ("firms"). For each period, we form a zero-investment portfolio that takes a long position in firms with negative accruals ($Accruals_t = -(H-L)$) and shorting all firms with positive accruals ($Accruals_{t-1} = H - L$). We use the equal-weighted returns as portfolio returns. Panel A reports the buy-and-hold return to an accruals-based hedge portfolio, where the x-axis denotes the number of periods forward, for a generic set of parameters, $(a, b, c^{\text{Earn}}, d^{\text{Earn}}, c^{\text{CF}}, d^{\text{CF}}) = (.75, .75, .75, .75, .60, .60)$. Panels B, C, and D report the variations of the one-period ahead accruals-based portfolio returns with respect to: (B) the persistence of the state process, i.e., $a = b \in [.50, .95]$, for $(c^{\text{E}}, d^{\text{E}}, c^{\text{CF}}, d^{\text{CF}}) = (.75, .75, .60, .60)$; (C) the relative informativeness of cash flows, i.e., $\Delta = c^{\text{CF}} - c^{\text{E}} = d^{\text{CF}} - d^{\text{E}} \in [-.25, .25]$, for $(a, b, c^{\text{E}}, d^{\text{E}}) = (.75, .75, .75, .75)$; and (D) the relative conservatism of cash flows, i.e., $\Delta = c^{\text{CF}} - c^{\text{E}} = \in [-.25, .25]$, for $(a, b, c^{\text{E}}, d^{\text{E}}, d^{\text{CF}}) = (.75, .75, .75, .75, .75)$.



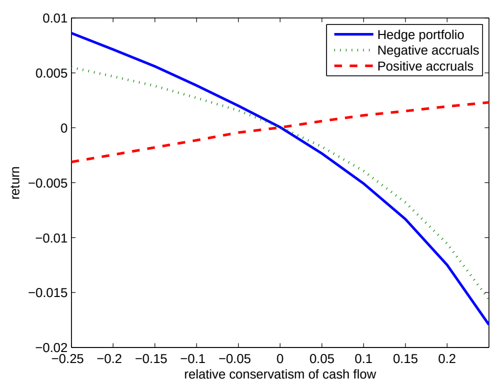
Panel A: Buy-and-hold returns based on accruals



Panel B: Varying persistence



Panel C: Varying relative informativeness



Panel D: Varying relative conservatism