Differential Interpretation of Information and the Post-Announcement Drift: A Story of Consensus Learning*

António Barbosa†
Instituto Universitário de Lisboa (ISCTE-IUL) and UNIDE

First version: March 17, 2009
This version: January 2, 2012

Abstract

I show how a post-announcement drift can be generated in a model with fully rational investors who interpret public information differently. Differential interpretation of information transforms public raw information into private interpreted information. If investors recognize their limited ability to interpret information, they will look for other investors' opinions in prices. Noise trading prevents investors from learning the market consensus interpretation of the announcement from the observation of a single price. But if noise trading follows a mean-reverting process, investors can gradually learn the market consensus from the observation of a series of prices. As investors become more confident about their interpretation of the announcement, they put more weight on it, and the announcement is gradually incorporated into prices, which generates a post-announcement drift. The model accounts for all salient empirical facts related to the post-announcement drift and delivers two new testable implications. If, in addition, investors make mistakes in extracting information from prices, the model also generates momentum.

JEL classification: G11; G12; G14
EFM classification: 350

---

*I am grateful to Zhaoyang Gu for sparking the idea for this paper and to Burton Hollifield for his support and helpful discussions. I thank Steffen Hitzemann, Rick Green, Stefano Gubellini, Bryan Routledge, Pedro Santa-Clara, Nitish Sinha, Chester Spatt, Qi Zhu, seminar participants at Carnegie Mellon University and Universidade Nova de Lisboa, and participants at the 2011 Midwest Finance Association, 2011 Eastern Finance Association and 2011 Financial Management Association European conferences for their helpful comments.

†ISCTE-IUL, Av. Forças Armadas, 1649-026 Lisbon, Portugal. Phone: +351 21 790 39 16. E-mail: antonio.barbosa@iscte.pt

1 Introduction

Under the market efficiency hypothesis, prices fully reflect public information. However, for the past 40 years, it has been shown that prices, specially those of small firms, drift following a vast array of announcements.[1] The traditional explanation for this phenomenon is that prices are slow to incorporate information. But the reason why prices are slow to incorporate information remains elusive. To a large extent, this elusiveness stems from a common misconception: that interpreting information is a trivial task everyone can perform in the exact same way. In reality, though, things are very different. Before investors can use new information to form trading decisions, they need to figure out its implications for the asset's fundamental value. Usually, this involves forecasting a stream of future cash-flows and discount factors, which makes the task far from trivial. But, above all, it makes it a subjective task. Different people, with different innate abilities, acquired skills, time available to do research, mood, among many other aspects, will naturally reach different conclusions based on the same information. This is evident, for example, from the dispersion of analyst forecasts.

In this paper, I develop a model that incorporates the idea of differential interpretation of public information, and I show how it can generate a post-announcement drift. I use a finite-horizon dynamic noisy rational expectations framework with a continuum of long-lived risk averse investors, similar to He and Wang (1995).[2] To obtain a post-announcement drift, I make three key assumptions: (i) there is differential interpretation of public announcements; (ii) investors recognize their limited ability to interpret information and so they do not hold dogmatically to their interpretations; (iii) the random net supply of the risky asset follows a mean-reverting process.

The assumption of differential interpretation of information essentially transforms the raw public information into private interpreted information. Assumption (ii) marks the departure of this model from the difference of opinions literature, where investors are assumed to hold their beliefs dogmatically (e.g. Harris and Raviv, 1993, Kandel and Pearson, 1995, Banerjee and Kremer 2010).[3] Instead, I assume that investors recognize their interpretations (or opinions) are imperfect, and so they have an incentive to look for others' opinions in order to improve their own interpretations. This can be accomplished by observing prices, since prices aggregate the opinions of all market participants. However, the information extracted from prices is noisy due to noise/liquidity trading, as in Grossman and Stiglitz (1980). Therefore, investors can learn something about the market consensus interpretation from a single observation of the price, but not everything. This is where assumption (iii) kicks in. As long as noise trading follows a mean-reverting process, the observa-

[1] These announcements include, but are not limited to, earnings surprises (Ball and Brown, 1968; Foster et al., 1984 and Bernard and Thomas, 1989, 1990), change in analyst recommendations (Michaely and Womack, 1999), share repurchases (Ikenberry et al., 1995), seasoned equity offerings (Loughran and Ritter, 1995), venture capital distributions (Gompers and Lerner, 1998); stock splits (Grinblatt, Masulis, and Titman, 1984), dividend initiations and omissions (Michaely et al., 1995). Daniel et al. (1998) provide a comprehensive list of references related to the post-announcement drift. More recently, Sinha (2011) uses news tonality to construct a measure of long-term qualitative information and show that prices underreact to news releases in general.

[2] The model is an extension of the model used by He and Wang (1995) to study trading volume.

[3] For example, Banerjee and Kremer (2010) assume that investors interpret public information differently but that each investor "believes that the other investor is wrong and so ignores his interpretation".

tion of subsequent prices provides additional information about the market consensus. As a result, investors can gradually correct their interpretation error by observing the series of prices following the announcement. In turn, as investors become more confident about their interpretation of the announcement, they put more weight on it relatively to their prior beliefs. Therefore, the information contained in the public announcement is gradually incorporated into posterior beliefs and prices, generating a post-announcement drift. The public information becomes fully incorporated into prices only after all market participants completely eliminate their interpretation errors.

The model delivers two new testable implications. The first is that the drift should last longer for stocks with less analyst coverage. The idea is that analysts help investors interpret information, thus speeding up the incorporation of information in prices. Hong et al. (2000) test this hypothesis but in the context of momentum, finding that analyst coverage and momentum are negatively associated.

The second is that the larger the interpretation uncertainty, the larger and longer the drift, and the smaller the price reaction to the announcement. This prediction can be tested in two different ways. The direct approach is to use a proxy for interpretation uncertainty and test whether interpretation uncertainty is negatively associated to the immediate price reaction to the announcement and positively associated to the magnitude and duration of the subsequent drift. The indirect approach is to test whether for stocks with similarly surprising news (e.g. SUE) the immediate price response to the announcement is inversely related to the magnitude and duration of the subsequent drift. There is already some evidence supporting this prediction, following the direct approach. Garfinkel and Sokobin (2006) find that the post-earnings announcement drift is larger when the unexpected volume at the announcement date, which proxies for differences of opinion, is larger. Zhang (2006) finds similar results for the post-analyst forecast revision drift using several proxies for information uncertainty. And Hirshleifer et al. (2009) find that the drift is larger in days with several announcements, likely because investors do not have enough time to carefully interpret the announcements.

In addition, the model is able to explain the main empirical facts related to the post-announcement drift. First, according to the model, every announcement whose interpretation is non-trivial will generate a drift, which explains why prices have been found to drift after such a vast and disparate array of announcements (see Daniel et al., 1998 for a comprehensive list with references).[4] Second, those drifts are long lasting because noise trading slows down the process of learning the market consensus from prices. Third, the drift is concentrated in the days following the announcement (Bernard and Thomas, 1989) because that is when investors are more uncertainty about their interpretations, and thus more willing to revise them. Fourth, when investors are positively(negatively) surprised by an announcement they tend to be positively(negatively) surprised by subsequent announcements, not because they do not understand the relation between those announcements (Bernard and Thomas, 1989), but because they do not have a full understanding of the first announcement. Thus, the drift

[4]In practice, the availability of data and the frequency of the announcement (i.e. number of observations) puts a limit on the type of announcements for which we can empirically detect a drift. This is likely the reason why the post-earnings announcement drift is more robust than drifts following other types of announcements (Fama 1998).

is also concentrated around subsequent announcement dates. Fifth, because arguably institutional investor make smaller interpretation errors and are less risk averse, they trade more aggressively to exploit the drift. As a result, the drift is attenuated although not completely eliminated when institutional investors trade more actively, which is associated to larger and more liquid stocks, and to positive news (Bartov et al., 2000 and Cohen et al., 2002).

Chan (2003) finds that momentum is an exclusive of stocks with news and Chordia and Shivakumar (2006) show that momentum is captured by the systematic component of post-earnings announcement drift. These findings suggest that post-announcement drift and momentum are two faces of the same coin. Given that the model successfully generates a post-announcement drift, it seems reasonable to assume that it generates momentum as well. But things are not so straightforward. Whereas interpretation errors cause beliefs conditional on the announcement to be biased toward the prior, the assumption that investors can flawlessly extract information from prices (a fundamental assumption in the REE framework) causes beliefs conditional on prices to be unbiased. Since the drift results from the gradual correction of beliefs biased toward the prior, we obtain a drift when conditioning on the announcement (post-announcement drift), but not when conditioning on the announcement date price reaction (momentum). Hence, for the model to generate momentum in addition to a post-announcement drift, it is necessary to allow for mistakes in the extraction of information from prices as well. With this modification, momentum and the post-announcement drift share the same root cause, differential interpretation of information, and have similar characteristics. Furthermore, the model can now rationalize why momentum and post-earnings announcement strategies are largely independent, as documented by Chan et al. (1996) and Brandt et al. (2008).

Several papers have analyzed models that can deliver some sort of momentum. Allen et al. (2006) show that, in a dynamic REE framework with overlapping generations of investors with heterogeneous beliefs, the law of iterated expectations does not always hold for average expectations. As a consequence, prices react sluggishly to changes in fundamental values and average expectations are closer to the fundamental value than prices. However, their model is bound to generate a drift since additional exogenous information about the fundamental value becomes available every period: each generation observes new private information in addition to the history of prices from which information about the private information of previous generations can be extracted. Cespa and Vives (2010) use a 2-period model similar to the one developed in this paper to study the circumstances where prices or average expectations are closer to fundamentals, but they fail to address how prices change over time in response to the arrival of private information. Furthermore, they always assume that additional private information from an exogenous source is available every period. Banerjee et al. (2009) show that higher-order difference of opinions generates momentum in a dynamic setting. Biais et al. (2010) consider an infinite horizon overlapping generations REE where a fraction of the investors receive a private signal every period. Momentum arises in their model but only because the cash-flow process exhibits positive serial correlation. Barberis et al. (1998) and Daniel et al. (1998) rely on different behavioral biases to generate short run momentum and long run reversal. Hong and Stein (1999) use boundedly rational investors to attain the same goal.

In their paper, momentum is generated by the gradual flow of information which is an exogenous assumption instead of an endogenous result as in the model of this paper. All these papers have in common the fact that they do not consider public announcements as a source of heterogeneous beliefs, and so cannot account for post-announcement drift.[5] In addition, many assume that new exogenous information is available every period. In contrast, in this paper I show that the release of exogenous information at a single date (the announcement date) is enough to generate a drift. All additional information is generated endogenously.

The remainder of the paper is organized as follows. I describe the model in Section 2 and solve for the equilibrium in Section 3. Section 4 is dedicated to the analysis of the post-announcement drift generated by the model. In Section 5 I discuss the results, confront them with the extant empirical evidence, and discuss the link between post-announcement drift and momentum. Section 6 concludes. All proofs are provided in the Appendix.

2 The Model

Consider an economy with one risky asset and one riskless asset traded at dates $1, 2, ..., T-1$. The riskless asset has a perfectly elastic supply and a zero net rate of return. The risky asset is liquidated at date T, paying $v \sim N\left(0, \sigma_v^2\right)$. The date t risky asset's per capita net supply (θ_t) is random, due to noise/liquidity driven demand, and follows the $AR\left(1\right)$ process

$$\theta_t = \rho\theta_{t-1} + \varepsilon_{\theta,t},\, \varepsilon_{\theta,t} \sim N\left(0, \sigma_\theta^2\right),\, t = 1, 2, ..., T-1 \tag{1}$$

with $0 \leq \rho \leq 1$ and $\theta_0 = 0$.[6]

There is a continuum of risk averse investors of measure 1, indexed by i. All investors have CARA preferences over their terminal wealth (W_T^i) with the same coefficient of risk aversion (α). Without loss of generality, the initial wealth level is normalized to 1. At every date t investor i chooses the risky asset demand (X_t^i) that maximizes his expected utility conditional on the information currently available to him $(\mathcal{F}_t^i)$ solving

$$\max_{X_t^i} \mathbb{E}\left(-e^{-\alpha W_T^i}\middle| \mathcal{F}_t^i\right) \quad s.t. \quad W_{t+1}^i = W_t^i + X_t^i\left(P_{t+1} - P_t\right), \tag{2}$$

where P_t is date t risky asset's price. Investors have homogeneous prior beliefs, $v \sim N\left(0, \sigma_v^2\right)$ and $\theta_0 = 0$, denoted by $\mathcal{F}_0$.

Investors observe public announcements, each based on a non-observable underlying signal for the liquidation value v. There are n of these underlying signals, defined by

$$\boldsymbol{s} = v\mathbf{1}_{(n\times 1)} + \boldsymbol{\varepsilon}_s,\, \boldsymbol{\varepsilon}_s \sim N\left(\mathbf{0}_{(n\times 1)}, \Sigma_s\right), \tag{3}$$

[5] The only one of these papers that can generate a post-announcement drift is Daniel et al. (1998). In their paper, post-announcement drift is a consequence of continuing overreaction to previous private information and requires that investors exhibit both overconfidence and biased self-attribution.

[6] The net supply corresponds to the asset's supply net of the noise/liquidity driven demand.

where $\mathbf{1}_{(n\times 1)}$ and $\mathbf{0}_{(n\times 1)}$ denote $(n\times 1)$-dimensional vectors of ones and zeros, respectively.[7] Investors need to interpret public announcements in order to figure out its implications for the liquidation value v and update their beliefs accordingly. The correct interpretation of a public announcement is its underlying signal. However, because investors have limited ability to interpret announcements, they make interpretation errors. A key assumption of the paper is that each investor makes a different interpretation

$$\tilde{\boldsymbol{s}}_t^i = \boldsymbol{s} + \boldsymbol{\varepsilon}_t^i,\ \boldsymbol{\varepsilon}_t^i \sim N\left(\mathbf{0}_{(n\times 1)}, \Sigma_{\tilde{s},t}\right), \tag{4}$$

where $\boldsymbol{\varepsilon}_t^i$ corresponds to the idiosyncratic interpretation error. In essence, the raw public information provided by the announcement is transformed into private interpreted information. Interpretation errors $\boldsymbol{\varepsilon}_t^i$ are assumed to be independent across i and t and with independent components (i.e. $\Sigma_{\tilde{s},t}$ is a diagonal matrix). The covariance matrix $\Sigma_{\tilde{s},t}$ determines not only the accuracy of investor's interpretations, but also the timing and underlying signal of public announcements: the j-th diagonal entry of $\Sigma_{\tilde{s},t}$ is equal to ∞ at every date except when a public announcement based on the j-th underlying signal is made.

The total information available to investor i and the common information available to all investors are defined as $\mathcal{F}_t^i = \{\mathcal{F}_0, \tilde{\boldsymbol{s}}_\tau^i, P_\tau : \tau = 1, ..., t\}$ and $\mathcal{F}_t^c = \{\mathcal{F}_0, P_\tau : \tau = 1, ..., t\}$, respectively. Unless it is explicitly stated otherwise, all random variables are independent of each other and across time.

Even though each investor forms his own interpretation, this model differs from difference of opinions models like Banerjee and Kremer (2010) in that investors are not dogmatic about their interpretation of public signals. Instead, investors recognize their limited ability to interpret public signals and would like to get rid of their interpretation error. Since interpretation errors are unbiased and independent across investors, and there is an infinite number of investors, the law of large numbers implies that the average interpretation of all market participants is the correct one, $\boldsymbol{s}$. From now on, I will refer to this average interpretation as the market consensus.

Investors can find information about the market consensus in equilibrium prices, which aggregate the information of all market participants. However, the existence of noise trading makes equilibrium prices only a noisy signal for the market consensus. But, as we will see, as long as future changes in net supply, and thus in prices, can be forecasted (i.e. $\rho < 1$), investors can always learn something more about the market consensus from the observation of each additional price, even though no additional exogenous information is released. Ultimately, it is the combination of differential interpretation of the public signals and the predictability of the changes in net supply that generates a post-announcement drift in this model.

[7] To distinguish vector-valued variables and parameters from scalar ones, I use bold-faced letters and numbers throughout the paper to denote the former.

3 Equilibrium

As usual in CARA-normal frameworks, there is an equilibrium where the price of the risky asset is a linear function of the state variables $\{v, \theta_\tau, \boldsymbol{s}, \tilde{\boldsymbol{s}}_\tau^i : \tau = 1, ..., t,\ 0 \le i \le 1\}$. I will only consider equilibria of this form.

The equilibrium price will depend on the whole history of signals and net supply realizations. However, as in He and Wang (1995), the history dependence of the date t equilibrium can be summarized by the posterior beliefs about the vector of signals conditional on the common information, $\mathbb{E}\left(\boldsymbol{s}|\mathcal{F}_t^c\right)$. This means that we can work with a low-dimensional state space and avoid the curse of dimensionality.

Theorem 1. *The price function in a linear equilibrium is given by*

$$P_t = \hat{\boldsymbol{p}}_t \mathbb{E}\left(\boldsymbol{s}|\mathcal{F}_t^c\right) + \boldsymbol{p}_t \boldsymbol{s} + p_{\theta,t}\theta_t = \hat{\boldsymbol{p}}_t \mathbb{E}\left(\boldsymbol{s}|\mathcal{F}_t^c\right) + \xi_t \tag{5}$$

where

$$\xi_t \equiv \boldsymbol{p}_t \boldsymbol{s} + p_{\theta,t}\theta_t. \tag{6}$$

Notice that ξ_t is observationally equivalent to P_t, since $\mathbb{E}\left(\boldsymbol{s}|\mathcal{F}_t^c\right)$ is known conditional on either $\mathcal{F}_t^c$ or $\mathcal{F}_t^i$.

In the remainder of this section I will prove Theorem 1 in three steps. First, I take the equilibrium price function as given and derive investors' conditional (on $\mathcal{F}_t^i$ and on $\mathcal{F}_t^c$) beliefs about the risky asset's terminal value and the n signals. Then, I use the equilibrium price function and investors' conditional beliefs and obtain the demand function by solving the utility maximization problem (2). Finally, given the demand function, I find the coefficients of the price function (5) that clear the market. All proofs are presented in Appendix A.

3.1 Posterior Beliefs

The assumption of a linear price function implies that the signal extracted from prices is linear in the state variables. Since all state variables are jointly normally distributed, posterior beliefs can be obtained using linear Gaussian filtering. The following lemma, drawing on the linear Gaussian filtering results in Liptser and Shiryaev (2001), gives the date t conditional expectations.

Lemma 2. *Given the linear price function, conditional expectations are defined by*

$$\mathbb{E}\left(\left.\begin{bmatrix} v \\ \boldsymbol{s} \\ \theta_t \end{bmatrix}\right| \mathcal{F}_t^i\right) = \mathbb{E}\left(\left.\begin{bmatrix} v \\ \boldsymbol{s} \\ \rho\theta_{t-1} \end{bmatrix}\right| \mathcal{F}_{t-1}^i\right) + K_t \left\{\begin{bmatrix} \xi_t \\ \tilde{\boldsymbol{s}}_t^i \end{bmatrix} - \mathbb{E}\left(\left.\begin{bmatrix} \xi_t \\ \boldsymbol{s} \end{bmatrix}\right| \mathcal{F}_{t-1}^i\right)\right\} \tag{7}$$

and

$$\mathbb{E}\left(\left.\begin{bmatrix} v \\ \boldsymbol{s} \\ \theta_t \end{bmatrix}\right| \mathcal{F}_t^c\right) = \mathbb{E}\left(\left.\begin{bmatrix} v \\ \boldsymbol{s} \\ \rho\theta_{t-1} \end{bmatrix}\right| \mathcal{F}_{t-1}^c\right) + K_t^c \left[\xi_t - \mathbb{E}\left(\xi_t | \mathcal{F}_{t-1}^c\right)\right] \tag{8}$$

with matrices K_t and K_t^c as defined in Appendix A.1.

Date t conditional expectations are a linear function of the previous period conditional expectations, the contemporaneous price (more precisely ξ_t, which is observationally equivalent to P_t), and the idiosyncratic interpretation of the contemporaneous public announcements ($\tilde{\boldsymbol{s}}_t^i$). Thus, for a given vector of underlying signals $\boldsymbol{s}$ and random net supply θ_t, conditional expectations are normally distributed across investors, due to the interpretation error. That is, investors who interpret one public announcement optimistically end up with optimistic expectations about that same announcement, despite their best efforts to correct their interpretation error.

3.2 Optimal Portfolio Allocation

Investors determine their optimal risky asset demand by solving the optimization problem (2). From (2) it is immediate that the optimal demand depends only on the beliefs about future excess returns per unit of risky asset. These beliefs can be pinned down from the price function and from the beliefs about the unobservable state variables of the economy, on which the equilibrium price depends. Let $\Delta P_{t+1} \equiv P_{t+1} - P_t$ denote the excess return per unit of risky asset from date t to date $t+1$. The next lemma gives the distribution of ΔP_{t+1} conditional on $\mathcal{F}_t^i$.

Lemma 3. *For every date $t = 1, 2, ..., T-1$,*

$$\Delta P_{t+1} = C_{t+1}\boldsymbol{\psi}_t + D_{t+1}\boldsymbol{\varepsilon}_{\Delta,t+1} \tag{9}$$

$$\boldsymbol{\psi}_{t+1} = F_{t+1}\boldsymbol{\psi}_t + G_{t+1}\boldsymbol{\varepsilon}_{\Delta,t+1} \tag{10}$$

where

$$\boldsymbol{\psi}_t \equiv \begin{bmatrix} 1 \\ \mathbb{E}\left(\boldsymbol{s}|\, \mathcal{F}_t^i\right) \\ \mathbb{E}\left(\boldsymbol{s}|\, \mathcal{F}_t^c\right) \\ \mathbb{E}\left(\theta_t|\, \mathcal{F}_t^i\right) \end{bmatrix}, \ \boldsymbol{\varepsilon}_{\Delta,t+1} \equiv \begin{bmatrix} v - \mathbb{E}\left(v|\, \mathcal{F}_t^i\right) \\ \boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}|\, \mathcal{F}_t^i\right) \\ \boldsymbol{\varepsilon}_{t+1}^i \\ \varepsilon_{\theta,t+1} \end{bmatrix},$$

$\boldsymbol{\varepsilon}_{\Delta,t+1}|\, \mathcal{F}_t^i \sim N\left(0, \Sigma_{\Delta,t+1}\right)$*, and C_{t+1}, D_{t+1}, F_{t+1}, G_{t+1} and $\Sigma_{\Delta,t+1}$ are constant matrices of appropriate order defined, respectively, by equations (25), (26), (20), (21) and (31) in Appendix A.2.*

The optimization problem (2) is solved through dynamic programming. Let $J\left(W_t^i, \boldsymbol{\psi}_t; t\right)$ denote the value function, which represents the optimized expected utility of current wealth given the future trading opportunities. The Bellman equation associated to problem (2) is

$$\begin{aligned} J\left(W_t^i, \boldsymbol{\psi}_t; t\right) &= \max_{X_t^i} \mathbb{E}\left[J\left(W_t^i + X_t^i \Delta P_{t+1}, \boldsymbol{\psi}_{t+1}; t+1\right) \Big|\, \mathcal{F}_t^i\right] \\ &s.t. \quad J\left(W_T^i, \boldsymbol{\psi}_T; T\right) = -e^{-\alpha W_T^i}. \end{aligned} \tag{11}$$

The normality of the conditional distribution of excess returns coupled to CARA preferences allows us to find the value and policy functions in closed form. The next lemma gives the policy function at date t, i.e. the optimal demand function, which is a linear function of the Gaussian vector $\boldsymbol{\psi}_t$.

Lemma 4. *Let ΔP_{t+1} follow the process defined in Lemma 3. Then the optimal demand function is*

$$X_t^i = \frac{1}{\alpha} Q_t \boldsymbol{\psi}_t,\ t = 1, 2, ..., T-1 \tag{12}$$

with matrix Q_t a function of $\{\boldsymbol{p}_{\tau_1}, \hat{\boldsymbol{p}}_{\tau_2}, p_{\theta,\tau_1} : \tau_1 = 1, ..., T-1,\ \tau_2 = t, ..., T-1\}$ as defined by equation (34) in Appendix A.3.

The demand of investor i is then a linear function of the expectations of $\boldsymbol{s}$ and θ_t conditional on his information set, and of the expectation of $\boldsymbol{s}$ conditional on the common information. Therefore, conditional on $\boldsymbol{s}$ and θ_t, the demand is normally distributed across investors because of interpretation errors and their effect on conditional expectations. More optimistic investors demand a larger quantity of the risky asset.

3.3 Market Clearing

In the final step, I use the market clearing condition

$$\int_i X_t^i = \theta_t,\ t = 1, 2, ..., T-1$$

and the demand function previously determined to pin down the equilibrium price function parameters $\{\boldsymbol{p}_t, \hat{\boldsymbol{p}}_t, p_{\theta,t}\}_{t=1,...,T-1}$. The next lemma gives the system of nonlinear equations associated to the $T-1$ market clearing conditions whose solution determines the $(T-1)\times(2n+1)$ price function coefficients.

Lemma 5. *(i) At dates $t = 1, 2, ..., T-1$, the market for the risky asset clears if and only if*

$$Q_t = \begin{bmatrix} 0 & \frac{\alpha}{p_{\theta,t}} \boldsymbol{p}_t \left(I_n - \hat{\Gamma}_t^{-1}\right) & -\frac{\alpha}{p_{\theta,t}} \boldsymbol{p}_t \left(I_n - \hat{\Gamma}_t^{-1}\right) & \alpha \end{bmatrix} \tag{13}$$

with matrix $\hat{\Gamma}_t$ defined by equation (42) in Appendix A.4.

(ii) Moreover,

$$\boldsymbol{p}_t + \hat{\boldsymbol{p}}_t = \hat{K},\ 1 \le t \le T-1, \tag{14}$$

with $\hat{K}$ a vector independent of the price function coefficients as defined in Appendix A.2.

Corollary 6. *Since Q_t is a function of $\{\boldsymbol{p}_{\tau_1}, \hat{\boldsymbol{p}}_{\tau_2}, p_{\theta,\tau_1} : \tau_1 = 1, ..., T-1,\ \tau_2 = t, ..., T-1\}$ and $\hat{\boldsymbol{p}}_t$ can be obtained from $\boldsymbol{p}_t$, the $T-1$ market clear conditions defined above form a system of $(T-1)\times(n+1)$ nonlinear equations from which the $(T-1)\times(n+1)$ price function coefficients $\{\boldsymbol{p}_\tau, p_{\theta,\tau}\}_{\tau=1,...,T-1}$ can be determined.*

Details about the numerical algorithm used to solve the system of nonlinear equations are provided in Appendix B.

To the extent that this system of equations has a solution, Lemma 5 proves that the price function has the linear form hypothesized in Theorem 1. As He and Wang (1995) note, there exists a unique solution in the absence of residual uncertainty (Σ_s is a matrix of zeros, implying that $\boldsymbol{s} = v\mathbf{1}_{(n\times 1)}$), that is, when the market as whole has enough information to know the liquidation value with certainty. However, a proof of existence for the general case is difficult to construct. Using numerical methods I am able to solve for an equilibrium numerically for a wide range of parameters, which suggests that, in general, existence is not a problem. However, it is increasingly difficult to obtain a solution as the number of trading dates increase and/or ρ decreases. The large number of trading dates makes search methods hopelessly ineffective. And the iteration method outlined in Appendix B becomes increasingly unstable for small values of ρ.[8] Also, with residual uncertainty, the results of Grundy and Mcnichols (1989) suggest that there may exist up to three distinct equilibria, depending on the parameter values. Numerical results indicate that this is the case in this model as well.

4 Post-Announcement Drift

In this section I show how, on average, a post-announcement drift in price and in conditional expectations of the liquidation value arises as a consequence of the differential interpretation of information and the predictability of changes in net supply. I start with the case of a single announcement and no residual uncertainty, for which there is a closed form solution for the price function coefficients. Then, I introduce residual uncertainty and characterize the multiple equilibria that arise in this setting. I show that residual uncertainty has no impact on the qualitative results when we consider the equilibrium with the most desirable characteristics. Finally, I analyze the case of multiple public announcements.

The focus will be on what happens when averaging over the net supply realizations. Therefore, because prices are a linear function of the net supply, in what follows I fix the net supply to its average value of zero at all dates.

4.1 No Residual Uncertainty: A Closed Form Solution

When there is no residual uncertainty, every underlying signal is exactly equal to the liquidation value. Therefore, without loss of generality we can assume that there is a single underlying signal $s = v$. To avoid unnecessary complexity, I consider a 3-period model, which is the most parsimonious model that can deliver a drift. Additional periods do not change the qualitative results but make the algebraic expressions intractable. The solution is presented in Appendix C.

The case of interest is when investors receive exogenous information, i.e. observe a public announcement, only at date 1. This means that any additional information at date 2 has to be endogenously produced from the observation of prices. The next theorem establishes the necessary

[8] When the iteration method fails to find a solution, a perturbation of the starting point is usually helpful. This suggests that when the iteration method fails to find a solution it is likely because of a bad starting point and not because there is no solution.

and sufficient conditions for expectations ($\mathbb{E}\left(v|\mathcal{F}_t^i\right)$) and prices ($P_t$) to react to the announcement and drift in this setting.

Theorem 7. *Let $T=3$, $n=1$, $\Sigma_s=0$, $\Sigma_{\tilde{s},1}=\sigma_{\tilde{s},1}^2$ and $\Sigma_{\tilde{s},2}=\infty$.*

(i) The average (over supply shocks and investors) announcement date reaction in the expectation of the liquidation value conditional on the surprise in the liquidation value (s) is

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)-\mathbb{E}\left(v|\,\mathcal{F}_0\right)\right]=\left[1-\frac{Var\left(v|\,\mathcal{F}_1^i\right)}{\sigma_v^2}\right]s$$

and the drift is

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_2^i\right)-\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)\right]=\frac{Var\left(v|\,\mathcal{F}_1^i\right)-Var\left(v|\,\mathcal{F}_2^i\right)}{\sigma_v^2}s.$$

(ii) The average (over supply shocks and investors) announcement date price reaction conditional on the surprise in the liquidation value (s) is

$$\mathbb{E}_\theta\left(\Delta P_1\right)=\left[1-(1-p_1)\frac{Var\left(v|\,\mathcal{F}_1^c\right)}{\sigma_v^2}\right]s$$

and the drift is

$$\mathbb{E}_\theta\left(\Delta P_2\right)=\frac{Var\left(v|\,\mathcal{F}_2^i\right)\left(1-\rho+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\right)}{\sigma_v^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\left(\sigma_v^2+\sigma_{\tilde{s},1}^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\sigma_v^2\right)}\left(1-\rho\right)s.$$

(iii) There is a reaction in the expectation of the liquidation value and price at the announcement date if and only if the announcement is informative ($\sigma_{\tilde{s},1}^2<\infty$) and a surprise ($s\neq 0$), and the liquidation value is not known with certainty before the announcement ($\sigma_v^2>0$).

(iv) There is a drift in the expectation of the liquidation value and price if and only if $\left\{\sigma_v^2,\sigma_{\tilde{s},1}^2,\alpha^2\sigma_\theta^2\right\}\in(0,\infty)^3\wedge 0\leq\rho<1\wedge s\neq 0$, i.e.: there is a reaction to the announcement ($\sigma_{\tilde{s},1}^2<\infty$, $s\neq 0$ and $\sigma_v^2>0$); the liquidation value is not learned at the announcement date ($\sigma_{\tilde{s},1}^2>0$ and $\alpha^2\sigma_\theta^2>0$);[9] endogenous information about the liquidation value is produced at date 2 ($\rho<1$ and $\alpha^2\sigma_\theta^2<\infty$); and prior beliefs are informative ($\sigma_v^2<\infty$).

(v) The reaction to the announcement and drift in the expectation of the liquidation value and price, if they exist, have the same sign of the surprise in the liquidation value (s).

Part (iv) of the theorem establishes that two of the main assumptions of the model, differential interpretation of public information ($\sigma_{\tilde{s},1}^2>0$) and predictability of net supply changes ($\rho<1$), are necessary conditions for the existence of a drift. Differential interpretation transforms the public raw information into private interpreted information (exogenous information). In turn, the predictability of net supply changes allows the aggregation of that private information, the market consensus interpretation, to gradually leak through subsequent prices (endogenous information). Obviously,

[9] In the general model with residual uncertainty, the equivalent condition is that neither the liquidation value nor the signal underlying the announcement are learned at the announcement date.

the information leaked through prices only matters because investors do not hold dogmatically to their interpretations, the third main assumption of the model.

Before we look at the other necessary conditions to obtain a drift and, ultimately, explain how the drift arises, it is worthwhile to understand how endogenous information is produced. Looking at the price equation (5), it is clear that predictability in net supply changes implies predictability in prices. For example, if the current net supply is negative, the net supply is expected to increase in subsequent periods, which puts a downward pressure on future prices. Moreover, the farther away the current net supply is perceived to be from its unconditional mean, the larger the adjustment in net supply and price that is expected to occur in the future. Therefore, investors can forecast future prices based on their beliefs about the current level of net supply.

In subsequent periods, each investor compares his forecast with the realized price. Any forecast error can only be explained by (i) a contemporaneous supply shock, (ii) a wrong expectation about the previous period net supply, or (iii) a mix of both. For example, the observed price will exceed its forecast if the contemporaneous supply shock is negative, which has a positive impact on the current price; or if the previous period net supply was in fact larger than previously believed.[10] If investors believe (ii) or (iii) to be the case, the upward revision of the belief about the previous period net supply leads to a reassessment of the information contained in the previous period price and, consequently, to an upward revision of the belief on the signal underlying the public announcement (given ξ_t , θ_t and s are positively related since $p_{\theta,t}$ is negative). But since investors do not observe the net supply, they do not know for sure to what extent the discrepancy between the observed price and its forecast can be attributed to (i) or (ii). Instead, they have to use the information available to form a belief about what happened. Invariably, investors consider that the most likely scenario is that the discrepancy between observed and forecasted price is explained by a conjugation of the two factors, for this is the scenario that implies the smallest departure from their prior beliefs about the previous period net supply and contemporaneous supply shock.[11] Hence, in every period after the public announcement, some information is endogenously produced and beliefs are revised.[12]

The larger the ρ, the smaller the expected adjustment in the net supply and consequently the less predictable the price change is. Therefore, the less likely it is that discrepancies between the observed and forecasted price are due to (ii); and the less investors adjust their initial expectations about the previous period net supply and underlying signal. In the limit, when $\rho = 1$, net supply changes are unpredictable, and so are price changes. Consequently, eventual differences between

[10]The latter means that the net supply, if positive, decreased more than expected, implying that prices increased more than expected; and that the net supply, if negative, increased less than expected, implying that prices decreased less than expected. In both cases, the observed price beats its forecast.

[11]Consider that the positive forecast error at $t = 2$ is due exclusively to a wrong expectation about θ_1, that is, the realized scenario is $(\theta_1, \varepsilon_{\theta,2}) = \left(\mathbb{E}\left(\theta_1 | \mathcal{F}_1^i\right) + e, 0\right)$ where $e > 0$ is the error in beliefs about θ_1. This scenario implies a deviation of θ_1, but not of $\varepsilon_{\theta,2}$, from its prior mean. Both variables have a normal prior distribution and, as we know, the only point at which the first derivative of the probability density function of a normal distribution is zero is at its mean. Therefore, the alternative scenario $\left(\mathbb{E}\left(\theta_1 | \mathcal{F}_1^i\right) + e - \epsilon_1, -\epsilon_2\right)$, $\epsilon_1 \gtrapprox 0$, $\epsilon_2 \gtrapprox 0$ is marginally more likely than the realized scenario. That is, it is always more likely that the forecast error is due to the deviation of both variables from their prior means than only one.

[12]Note that when the observed price matches its forecast, endogenous information is still produced. Although there is no revision of the expectation about the underlying signal, the precision of that expectation increases.

observed and forecasted prices are attributed only to contemporaneous supply shocks, and there is no adjustment to the belief about the underlying signal, i.e. there is no endogenous production of information.

Most of the remaining conditions identified in the theorem as necessary to obtain a drift are straightforward to understand. We need some prior uncertainty about the liquidation value ($\sigma_v^2 > 0$), and an informative announcement ($\sigma_{\tilde{s},1}^2 < \infty$), otherwise there is nothing to be learned from the announcement and there is not even a price reaction to the announcement. This stresses the role of learning about the market consensus for the existence of a drift, and clearly shows that mean-reversion of net supply *per se* does not generate a drift. The role of mean-reversion of net supply is merely the endogenous production of information. A drift would still arise if, in the periods following the initial announcement, additional information became available from an exogenous source, instead of an endogenous one (see Theorem 8 below).

In addition, to obtain a drift we need the equilibrium price at the announcement date to be a noisy, but informative, signal for the private information ($0 < \alpha^2\sigma_\theta^2 < \infty$), so that the market consensus is learned gradually.[13] But gradually learning about the market consensus, though necessary, is not a sufficient condition to generate a drift in expectations ($\mathbb{E}\left(v|\,\mathcal{F}_t^i\right)$) and prices ($P_t$). On average (across investors and net supply), expectations formed at the announcement date have to be systematically biased toward prior beliefs, which requires prior beliefs to be informative.[14] Only then does the gradual learning of the market consensus generate a drift in a predictable direction, as that bias is gradually corrected.

So, the question is, "Why are expectations about the underlying signal systematically biased toward the prior?" The answer is quite simply because investors only observe noisy versions of the announcement's underlying signal (s): their interpretation of the announcement ($\tilde{s}_t^i = s + \boldsymbol{\varepsilon}_t^i$); and the information they extract from equilibrium prices ($\xi_t = s + \frac{p_{\theta,t}}{p_t}\theta_t$).[15] Because of this, investors with informative prior beliefs ($\sigma_v^2 < \infty$) will always put some weight into their prior beliefs when forming posterior beliefs, as we can see from expression (49) derived in Appendix C and reproduced here with $\beta \to 1$ (i.e. only endogenous information at date 2):

$$\mathbb{E}\left(v|\,\mathcal{F}_1^i\right) = \frac{\frac{1}{\sigma_v^2}0 + \frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}\xi_1 + \frac{1}{\sigma_{\tilde{s},1}^2}\tilde{s}_1^i}{\frac{1}{\sigma_v^2} + \frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2} + \frac{1}{\sigma_{\tilde{s},1}^2}},\ \mathbb{E}\left(v|\,\mathcal{F}_2^i\right) = \frac{\frac{1}{\sigma_v^2}0 + \frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}\xi_1 + \frac{1}{\sigma_{\tilde{s},1}^2}\tilde{s}_1^i + \frac{(1-\rho)^2}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}\xi_2}{\frac{1}{\sigma_v^2} + \frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2} + \frac{1}{\sigma_{\tilde{s},1}^2} + \frac{(1-\rho)^2}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}}.$$

[13] If information at date 2 is from an exogenous source, it is not necessary to extract information from prices, and so we can have $\alpha^2\sigma_\theta^2 \to \infty$.

[14] The term biased is used loosely here. It does not mean that investors form biased beliefs given the information available at the time. Instead, it means that beliefs are biased when compared to the beliefs that would prevail if investors were able to correctly interpret the announcement, that is, if the underlying signal were observed directly.

[15] Here I'm using the alternative definition of ξ_t used in Appendix C.

Averaging over investors and net supply, posterior expectations become

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)\right]=\frac{\frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}+\frac{1}{\sigma_{\tilde{s},1}^2}}{\frac{1}{\sigma_v^2}+\frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}+\frac{1}{\sigma_{\tilde{s},1}^2}}s,\ \mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_2^i\right)\right]=\frac{\frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}+\frac{1}{\sigma_{\tilde{s},1}^2}+\frac{(1-\rho)^2}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}}{\frac{1}{\sigma_v^2}+\frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}+\frac{1}{\sigma_{\tilde{s},1}^2}+\frac{(1-\rho)^2}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}}s.$$

These expressions clearly show that there is a systematic bias towards the prior belief $\mathbb{E}\left(v|\,\mathcal{F}_0\right)=0$ as long as $\sigma_v^2<\infty$. Moreover, this bias shrinks as investors learn about the market consensus in date 2 ($\rho<1$); this is why expectations drift. If there were no bias to correct ($\sigma_v^2=\infty$), learning about the market consensus alone would not generate a drift. Naturally, if expectations drift, prices have to follow suit and drift in the same direction as expectations, otherwise the market would not clear.

A similar drift can be obtained when the source of additional information is exclusively exogenous, or a mixture of endogenous and exogenous source. In fact, the precision of exogenous signals can be chosen so that the average price drift is the same as in the case of endogenous information. This clearly illustrates that the drift results from the arrival of new information leading to a gradual learning of the market consensus and correction of the bias toward prior beliefs, and not from the dynamics of the net supply. The following theorem is the equivalent of Theorem 7 for the case where additional information at date 2 is exclusively exogenous ($\rho=1$).

Theorem 8. *Let $T=3$, $n=1$, $\Sigma_s=0$, $\Sigma_{\tilde{s},1}=\sigma_{\tilde{s},1}^2$, $\Sigma_{\tilde{s},2}=\frac{\beta}{1-\beta}\sigma_{\tilde{s},1}^2$, $0\leq\beta\leq1$ and $\rho=1$.*

(i) The average (over supply shocks and investors) announcement date reaction in the expectation of the liquidation value and price conditional on the surprise in the liquidation value (s) is

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)-\mathbb{E}\left(v|\,\mathcal{F}_0\right)\right]=\mathbb{E}_\theta\left(\Delta P_1\right)=\left(1-\frac{Var\left(v|\,\mathcal{F}_1^i\right)}{\sigma_v^2}\right)s$$

and the drift is

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_2^i\right)-\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)\right]=\mathbb{E}_\theta\left(\Delta P_2\right)=\frac{Var\left(v|\,\mathcal{F}_1^i\right)-Var\left(v|\,\mathcal{F}_2^i\right)}{\sigma_v^2}s.$$

(ii) There is a reaction in the expectation of the liquidation value and price at the announcement date if and only if the announcement is informative ($\sigma_{\tilde{s},1}^2<\infty$) and a surprise ($s\neq0$), and the liquidation value is not known with certainty before the announcement ($\sigma_v^2>0$).

(iii) There is a drift in the expectation of the liquidation value and price if and only if $\left\{\sigma_v^2,\sigma_{\tilde{s},1}^2\right\}\in(0,\infty)^2\wedge\alpha^2\sigma_\theta^2>0\wedge0\leq\beta<1\wedge v\neq0$, i.e.: there is a reaction to the announcement ($\sigma_{\tilde{s},1}^2<\infty$, $s\neq0$ and $\sigma_v^2>0$); the liquidation value is not learned at the announcement date ($\sigma_{\tilde{s},1}^2>0$ and $\alpha^2\sigma_\theta^2>0$); new information about the liquidation value is available at date 2 ($\beta<1$); prior beliefs are informative ($\sigma_v^2<\infty$).

(iv) The reaction to the announcement and drift in the expectation of the liquidation value and price, if they exist, have the same sign of the surprise in the liquidation value (s).

4.2 Calibration and Comparative Statics

As we just saw, the gradual learning of the market consensus interpretation of a public announcement generates a price drift. The question now is whether it can generate a quantitatively meaningful drift. I shed some light into this question by calibrating the model. The first thing to choose is the time interval between periods, which determines the frequency with which investors look for new information, revise beliefs and trade. This is actually a tricky choice. Fixing the net supply volatility over a given time frame, the shorter the time interval between net supply shocks, the less volatile these shocks have to be, resulting in more informative prices. In the limit, a day trader would be able to learn the market consensus immediately. Clearly, this is a limitation of the net supply process; a more realistic arrival process would obviate this problem. As a compromise, I set a weekly time interval. I consider 100 weeks (roughly 2 years) of trading until the asset is liquidated, which is more than enough for the drift to die out before liquidation.

The assumption of CARA preferences means that relative price changes, needed to calibrate the parameters associated to noise trading, depend on the price level. To deal with this free parameter, I follow Campbell, Grossman, and Wang (1993) and set the average liquidation value ($\mathbb{E}(v)$) to 1 and the average net supply ($\mathbb{E}(\theta_t)$) to 0, so that prices are relatively close to 1. In addition, I set the coefficient of absolute risk aversion (α) to 2 which, for wealth levels close to 1, is close to the coefficient of relative risk aversion.

To calibrate the parameters associated to noise trading, ρ and σ_θ^2, I match the return dynamics of the model to those of the S&P 500.[16] The autocorrelation in returns is almost exclusively driven by ρ, and setting $\rho = 0.5$ the model generates a first-order autocorrelation in weekly returns of -0.048, which closely matches that of the S&P 500 over the last 40 years (-0.043). The major driver of the unconditional volatility of returns is σ_θ^2, although σ_v^2 and $\sigma_{\tilde{s},1}^2$ have some influence. Setting $\sigma_\theta^2 = 0.15$, the volatility of weekly returns in the model is 2.7% compared to 2.3% of the S&P 500.

Finally, the variance of the prior belief about the liquidation value is set at $\sigma_v^2 = 0.25$. This implies a prior probability of bankruptcy in the next 2 years of about 2.5%. Since there is no guidance for the precision of interpretations, I set it to half of the precision of prior beliefs, $\sigma_{\tilde{s},1}^2 = 0.5$, in order to closely match the drift characteristics identified in Bernard and Thomas (1989).[17]

In Bernard and Thomas (1989), the average total price response (immediate response plus drift) to the announcement of extreme earnings surprises for small firms is 12.5%; 60% percent of the total price change occurs at or before the announcement date, with the remaining 40% spread over the next 9 months. In the calibrated model, a similar 12.5% surprise (calibrated by choosing s) results in a 69%/31% split. An econometrician measuring monthly drifts using 2500 observations

[16] There is no direct empirical data on the noise trading process. To a great extent this is due to the difficulty in separating informed trading from noise trading. Dow and Gorton (2008) provide a literature review on noise traders and note that the identity of noise traders is still an open question. Berkman and Koch (2008) use the dispersion of the daily net initiated order flow across brokers as a proxy for the level of noise trading. However, their measure of noise trading is only proportional to σ_θ^2 and they do not provide the time-series dynamics of their proxy.

[17] Interpretations might seem very diffuse but keep in mind that much of the uncertainty is resolved at the announcement date from the observation of the corresponding equilibrium price. In effect, the variance of the interpretation error drops from 0.5 to just 0.08 with the observation of the announcement day price.

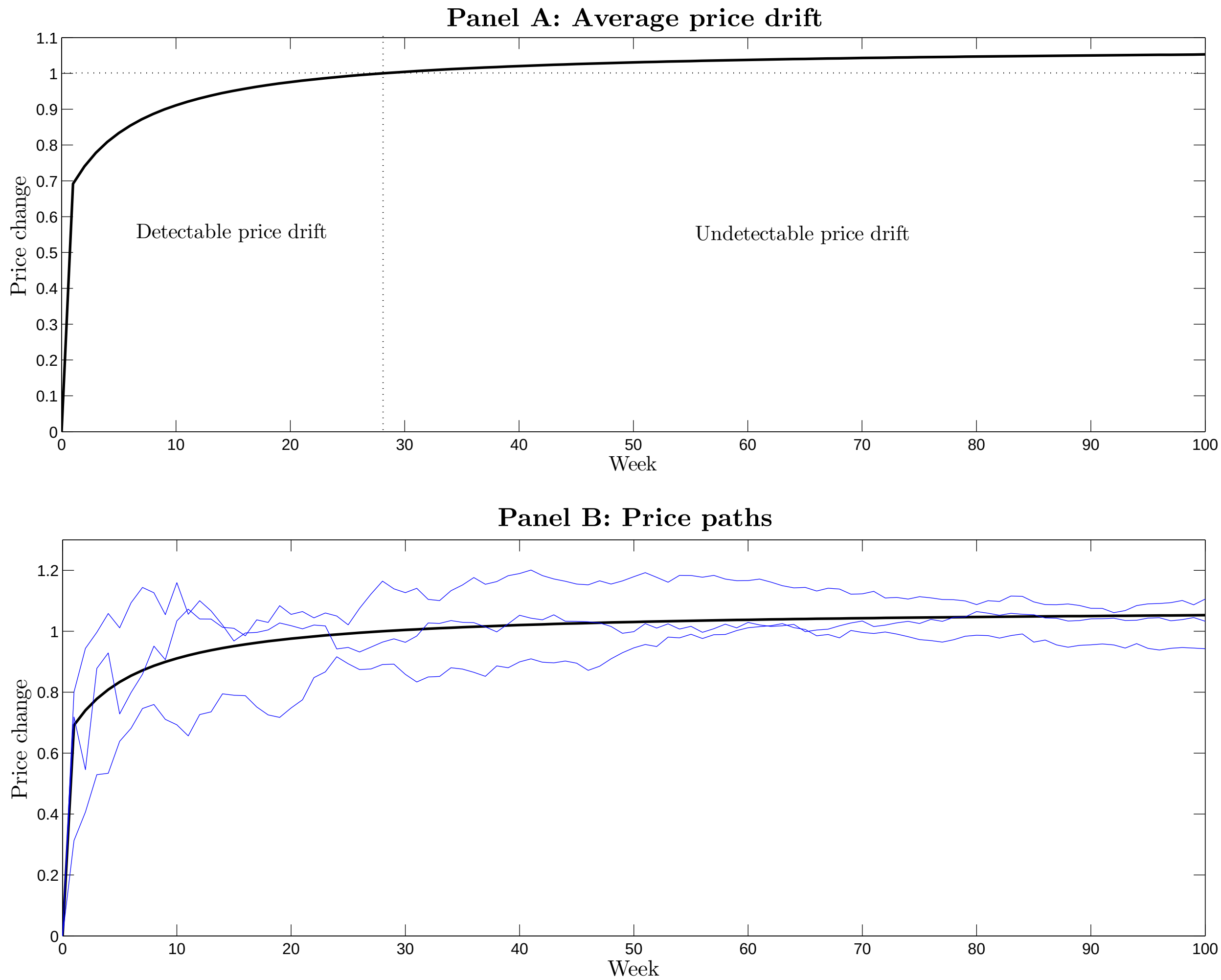

Figure 1: Price drift. Panel A plots the cross-sectional average of prices over the 100 weeks following and including the announcement, illustrating the average price drift. Panel B plots the average price path (thick line) along with three of its realizations (thin lines). The parametrization corresponds to the calibration discussed in the main text: $T = 101$, $n = 1$, $\sigma_v^2 = 0.25$, $\Sigma_s = 0$ (no residual uncertainty), $\Sigma_{\tilde{s},1} = \sigma^2_{\tilde{s},1} = 0.5$, $\Sigma_{\tilde{s},t} = 10^{10}$ $(1 < t \leq T-1)$, $\sigma_\theta^2 = 0.15$, $\rho = 0.5$, $\alpha = 2$, $s = 1.132$, $\mathbb{E}(v) = 1$. The parameter s was chosen so that the measurable price change in reaction to the announcement (initial reaction plus measurable drift) corresponds to 12.5%. An econometrician with 2,500 observations measuring drift month by month will detect a drift up to week 28 (roughly 7 months). In Panel A this is identified by the dotted lines. In both Panels, price changes are normalized to the total measurable price reaction to the announcement by such an econometrician. The initial price response to the announcement is 69% of the total measurable price reaction, with the remaining 31% spread over the next 28 weeks. In the last 72 weeks prices increase by an additional 5%, corresponding to the undetectable drift.

(Bernard and Thomas 1989 have around 2700), would find a significant drift up to the seventh month following the announcement. The model thus seems to be able to generate a realistic drift for plausible parameter values. Figure 1 shows the average price drift in the 100 periods following the announcement of positive news (the case of negative news is symmetric) for the calibrated model.

Taking the calibrated model as the benchmark, Figure 2 shows the impact of the main model parameters in the way public announcements are incorporated into prices. Note that with alternative

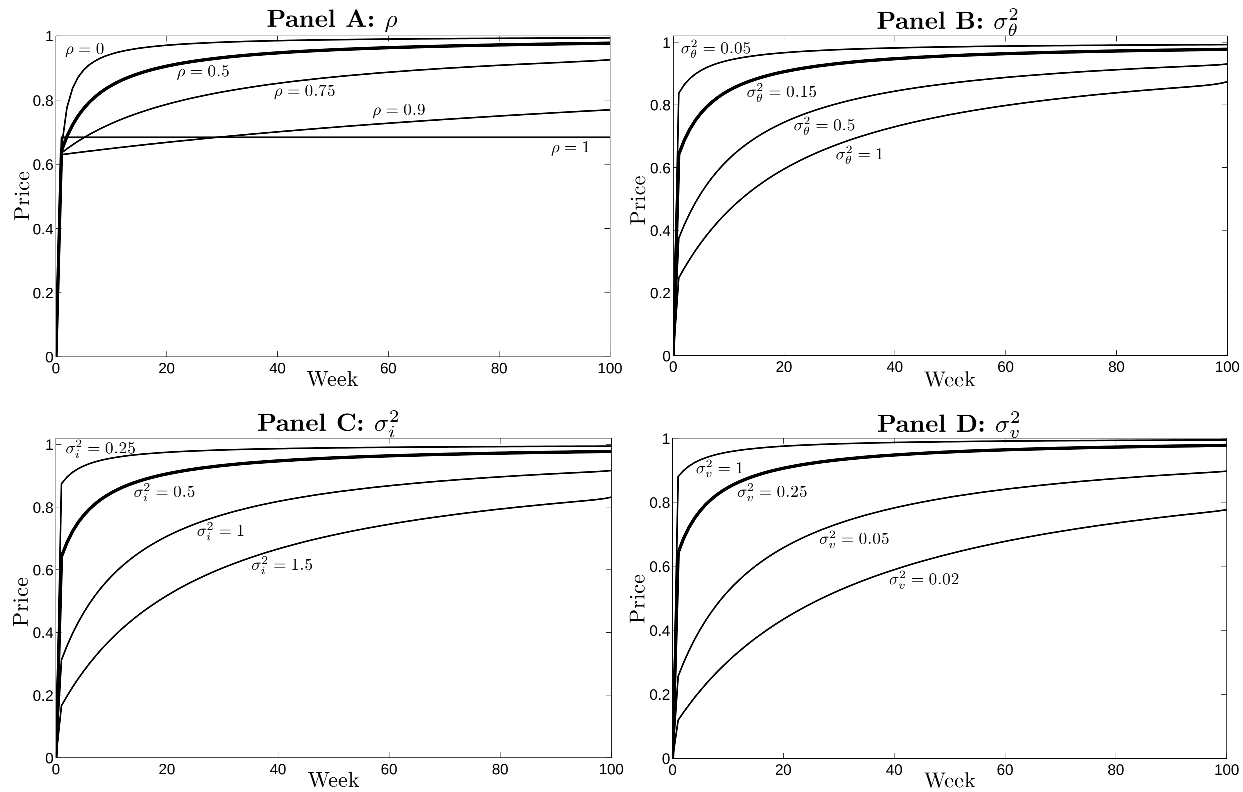

Figure 2: Price drift comparative statics. Panels A to D plot the price drift for different values of ρ (Panel A), σ_θ^2 (Panel B), $\sigma_{\tilde{s},1}^2$ (Panel C) and σ_v^2 (Panel D). The base case (thick lines) corresponds to the parametrization in Figure 1.

parametrizations 100 weeks may not be long enough for most of information to be incorporated into prices; in those cases, and as long as consensus learning takes place ($\rho < 1$ and $\sigma_\theta^2 < \infty$), the drifts would continue beyond the 100th week until prices converge to 1.

The speed of mean-reversion of net supply (ρ) and the volatility of net supply (σ_θ^2) determine how much can be learned from the observation of prices; both low ρ and low σ_θ^2 speed up the incorporation of information into prices.[18] But whereas σ_θ^2 determines how much can be learned from the observation of a single price, ρ determines how much more can be learned from the observation of additional prices. Therefore, a lower ρ speeds up the incorporation of information in the periods following the announcement, but has no impact on the immediate price reaction to the announcement.[19] This results in a drift that is more concentrated around the announcement date. In turn, a lower σ_θ^2 increases the immediate price reaction to the announcement, leaving less information to be incorporated in subsequent periods leading to a smaller drift, both in magnitude and duration. The same results are obtained with a smaller volatility of the interpretation errors ($\sigma_{\tilde{s},1}^2$) since there is less uncertainty about the market consensus and thus less to be learned about it. Finally, the prior uncertainty about the liquidation value (σ_v^2) has the opposite impact of $\sigma_{\tilde{s},1}^2$ and

[18] Actually, what matters is the product $\alpha^2\sigma_\theta^2$, which represents the ratio of supply noise to risk tolerance. I focus only on σ_θ^2 for simplicity.

[19] There is an exception to the latter when ρ is very close to 1. In that case, the immediate price reaction increases because investors realize that they will not come close to learn the market consensus before the liquidation date.

σ_θ^2. If investors are confident about prior beliefs (low σ_v^2), they put a lot of weight on them and just a little on their announcement interpretation, which implies that prices provides little information about the market consensus. Therefore, the immediate price reaction to the announcement is small and the drift his large and long lasting.

A pattern seems to emerge from this comparative statics exercise: for a given market consensus, the smaller the immediate price response to the announcement, the larger and longer lasting the drift is. In particular, this is the case when the announcement is more difficult to interpret or, equivalently, the more disagreement there is in its interpretation. This suggests a novel testable implication: for stocks with similarly surprising news (e.g. earnings), those with a smaller initial price reaction to the announcement should have a larger and longer lasting drift.

4.3 Residual Uncertainty: Multiple Equilibria

Introducing residual uncertainty has two consequences: first, the equilibrium price function can no longer be determined in closed form; second, there are up to 3 distinct equilibria as in Grundy and Mcnichols (1989).[20] Numerical results suggest that only one of these equilibria exists regardless of the parametrization chosen; I call it the primary equilibrium. The other equilibria exist only when there is residual uncertainty, but not if it is too large, and the net supply is persistent (ρ is large) and not too volatile; these are the secondary equilibria.

The primary equilibrium corresponds to the equilibrium we analyzed in the absence of residual uncertainty. The introduction of residual uncertainty does not change its properties and so the mechanism behind the price drift is exactly the same as before. The only effect of residual uncertainty is the decreased relevance of the announcement; the underlying signal of the announcement is now a noisy, not a noiseless, signal for the liquidation value. Hence, both the immediate and delayed price response to the announcement are smaller.

The secondary equilibria also deliver a post-announcement drift in expectations and prices. Like in the primary equilibrium, the drift results from the gradual learning of the market consensus and consequent correction of the bias toward prior beliefs. However, the mechanism by which the information about the market consensus is endogenously generated is completely different. It relies on the price being considerably more sensitive to the underlying signal (s) than to the net supply (θ_1) at the announcement date (date 1), and the opposite to happen in the next period (date 2). This means that investors can learn a lot about s, but not much about θ_1, from the announcement date price; and that investors can learn a lot about θ_2, but not so much about s, directly from the observation of the subsequent price. However, as long as the net supply is persistent (ρ close to 1) and supply shocks are not too volatile (σ_θ^2 is low), investors can learn a big deal about θ_1 from θ_2. Investors can then use the improved knowledge about θ_1 to revisit the information provided by the announcement date price and further refine their beliefs about s. Therefore, the date 2 price ends up providing a reasonable amount of information about s.[21] In this scenario, a large θ_2 leads to an

[20]They analyze a special case of the model in this paper, with $\rho = 1$, $T = 3$ and a supply shock only at date 1, finding a non-revealing equilibrium and two revealing equilibria.

[21]Consider the following example, based on Grundy and Mcnichols (1989): $\theta_t = \theta \,\forall t$, i.e. $\rho = 1$, $\sigma_{\theta,1}^2 > 0$ and

upward revision of the θ_1 and, consequently, to an upward revision of the expectation of s (for the market to absorb a higher net supply at a given price, the asset has to become more attractive). As a consequence, the negative price impact of a larger net supply is more than compensated for by the respective upward revision in the expectation of the liquidation value and date 2 prices become *positively* correlated with net supply. At the same time, because the two effects of net supply on the price partially offset each other, the date 2 price exhibits very little volatility. Finally, almost all the learning takes place at dates 1 and 2. As a result, there is a meaningful drift for just one period after the announcement.

Because secondary equilibria have these undesirable properties, do not exist for some parametrizations, and are very hard to find numerically, I focus on the primary equilibrium in the remainder of the paper.[22]

4.4 Additional Public Announcements

The effect of additional public announcements on the price dynamics depends on their impact on the bias in beliefs created by previous announcements which, as we saw, is a key element of the post-announcement drift. To simplify the exposition, let us focus on what happens when a second public announcement is released.

When analyzing the effect of an additional public announcement on the bias toward prior beliefs there is an important distinction between the case where the underlying signal of the second announcement is the same of the first announcement and the case it is not. In the first case, the second announcement provides the same information (the underlying signal) of the first previous announcement but in a different way, helping investors to improve their interpretation of the first announcement.[23] One example is an analyst's opinion about a previous earnings announcement; another is the clarification of some aspects pertaining to a previous announcement. In the second case, the new announcement provides novel information. An example is the announcement of earnings for a different quarter.

When both announcements have the same underlying signal, the second announcement is, in essence, an exogenous signal for the market consensus interpretation of the first announcement.

$\sigma^2_{\theta,t} = 0$, $t > 1$. Writing the conditional expectations as a function of prior beliefs, the underlying signal and the series of net supply levels, it is straightforward to see that we can write $P_t = \varphi_0 + \varphi_{s,t}s + \varphi_{\theta,t}\theta$ for some coefficients φ_0, $\varphi_{s,t}$ and $\varphi_{\theta,t}$. Supposing that there is an equilibrium in which $\varphi_{s,1}/\varphi_{\theta,1} \neq \varphi_{s,2}/\varphi_{\theta,2}$, we can solve the system of equations for s and θ, and resolve all uncertainty, by date 2. In our general model, $\sigma^2_{\theta,t} = \sigma^2_{\theta,1} > 0$, $t > 1$, and so $\theta_t \neq \theta_{t-1}$. This implies that $P_t = \varphi_0 + \varphi_{s,t}s + \sum_{\tau=1}^{t} \varphi_{\theta_\tau,t}\theta_\tau$. Therefore, we always have one more variable than equations and we cannot solve the system of equations for s. But we can get an accurate estimate of s at date 2 if θ_t is persistent and not very volatile, and if the sensitivity of the price function to s and θ_1 is different enough at dates 1 and 2. The role of residual uncertainty is to allow the latter to subsist in equilibrium.

[22]Although not explicitly stated, He and Wang (1995) and Cespa and Vives (2010) analyze only the primary equilibrium, most likely because it was the only equilibrium they were able to compute numerically.

[23]Technically, given the assumption that interpretation errors are independent across time (i.e. investors are not systematically pessimistic or optimistic in their interpretations), the observation of an additional public announcement based on the same underlying signal generates an additional independent signal for the market consensus interpretation (the underlying signal) of both announcements. This unequivocally reduces the uncertainty in the interpretation of these initial announcement.

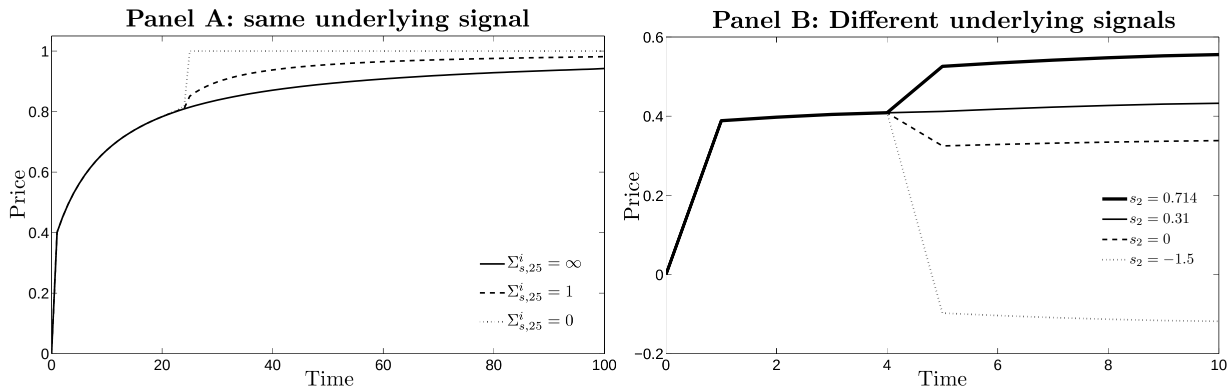

Figure 3: Price drift with two announcements. This figure plots the price drift when there is a second announcement based on the same underlying signal (Panel A) and based on independent underlying signals (Panel B). The parametrization used on Panel A is: $T = 101$, $n = 1$, $\sigma_v^2 = 0.25$, $\Sigma_s = 0$, $\Sigma_{\tilde{s},1} = 0.5$, $\Sigma_{\tilde{s},t} = 10^{10}\, \forall t \in [2, 24] \cup [26, 100]$, $\sigma_\theta^2 = 0.1$, $\rho = 0.5$, $\alpha = 2$, $s = 1$. The parametrization used in Panel B is: $T = 11$, $n = 2$, $\sigma_v^2 = 0.25$, $\Sigma_s = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$, $\Sigma_{\tilde{s},1} = \begin{bmatrix} 1 & 0 \\ 0 & 10^3 \end{bmatrix}$, $\Sigma_{\tilde{s},5} = \begin{bmatrix} 10^{10} & 0 \\ 0 & 1 \end{bmatrix}$ $\Sigma_{\tilde{s},t} = \begin{bmatrix} 10^{10} & 0 \\ 0 & 10^{10} \end{bmatrix}$ $(t > 1 \wedge t \neq 5)$, $\sigma_\theta^2 = 0.1$, $\rho = 0.75$, $\alpha = 2$, $s = 1$. In Panel B it was considered a smaller number of time periods due to the difficulty in solving the model numerically when there is residual uncertainty.

Therefore, the second announcement always decreases the bias in beliefs inherited from the first announcement. As a result, at the second announcement date the price jumps in the same direction it was drifting. Afterward, prices will continue to drift in the same direction as long as some uncertainty about the market consensus subsists. However, since posterior beliefs are now less biased, that drift will be smaller. Panel A of Figure 3 illustrates this case.

Now, suppose that the first and second announcement have different and independent underlying signals, s_1 and s_2.[24] In this case, the second announcement does not resolve any uncertainty about s_1; it just creates uncertainty about s_2.[25] Naturally, the price will jump up (down) at the second announcement date if most investors are positively (negatively) surprised by the announcement. However, this price jump is not necessarily followed by a drift in the same direction, as we can see from Panel B of Figure 3. Suppose that the second announcement is a positive surprise. In this case beliefs about the underlying signal s_2 will be biased downward. But what matters for the drift is the bias in beliefs about the liquidation value (v), which depends on the bias in beliefs about both underlying signals. If the first announcement was also a positive surprise, then the beliefs about both underlying signals are biased downward and clearly the beliefs about the v will remain biased downward following the second announcement. As a result, the upward jump in price at the second announcement date is followed by a positive drift. However, if the first announcement was a negative surprise, the beliefs about each underlying signal are biased in opposite directions. In this case, the beliefs about v can be biased either upward or downward, depending on which announcement dominates. If the second announcement is not surprising enough, if its underlying

[24] The case of correlated underlying signals is a mixture between this and the case of a single underlying signal.

[25] This is not entirely true, since both underlying signals are linked through the liquidation value v. Therefore, when investors learn about v from the second announcement, they also indirectly learn a little bit about the first underlying signal. If the two underlying signals were correlated, then investors could also learn about the first underlying signal directly from the second announcement.

signal is not precise enough, or its interpretation is precise enough (implying a small bias in the beliefs about s_2), beliefs about v will remain biased upward, and the upward jump in price at the second announcement date is followed by a negative drift.

As we can see from Panel A of Figure 3, when there is a single underlying signal the drift is concentrated in the days following the first announcement and in the day of the second announcement. This is precisely the kind of price reaction to earnings announcements that Bernard and Thomas (1989) document. However, different earnings announcement are more likely to be described as announcements with different underlying signals. In this case, investors are not necessarily positively (negatively) surprised by the announcement following a positive (negative) announcement. But on average, they are. To see why, consider the parametrization used in Panel B of Figure 3. If investors made no interpretation errors and were able to observe directly s_1, they would form beliefs $\mathbb{E}\left(v|\, s_1\right) = \mathbb{E}\left(s_2|\, s_1\right) = 0.714$ upon the observation of the first announcement. Thus, on average there would be no reaction to the observation of the second announcement. But since investors make interpretation errors, their beliefs about v and s_2 are biased toward the prior belief 0. As a result, the average investor has beliefs $\mathbb{E}\left(v|\, \mathcal{F}_4^i\right) = \mathbb{E}\left(s_2|\, \mathcal{F}_4^i\right) < 0.714$ which means that investors will tend to be positively surprised again by the second announcement.

5 Discussion

5.1 Why Prices Are Slow to Incorporate Information

There are two broad categories of explanations for the post-announcement drift and momentum phenomena: failure to completely account for a priced risk factor; and slow incorporation of information on prices. Empirical evidence seems to dismiss the risk-based explanation (e.g. Bernard and Thomas, 1989; Fama and French, 1996 and Hong et al., 2000), which leaves the slow incorporation of information on prices as the plausible explanation for these two anomalies. Various studies have identified ease of trading, analyst coverage and institutional ownership as factors that facilitate the incorporation of information on prices (e.g. Bernard and Thomas, 1989; Bartov et al., 2000; Hong et al., 2000; Cohen et al., 2002 and Chan, 2003). Yet, the reason why the average investor is slow to act on public information has either been overlooked (in Hong and Stein, 1999 slow dissemination of private information needed to interpret public information is an exogenous assumption); or explained by behavioral biases (e.g. Daniel et al., 1998 and Barberis et al., 1998).

Perhaps the greatest obstacle in understanding why prices are slow to fully reflect public information comes from the usual assumption that public information has a one-time impact on expectations and prices. After all, why would anything change once everyone observes the same information? What is usually missed is that information needs to be interpreted before it can be used to form a trading decision, and each individual will obviously do it differently. Differences in innate ability, in acquired skills, in mood, in the time available to do research (e.g., reading the executive summary of the earnings report vs. the full report) all contribute to differential interpretation of public information. If investors hold to their beliefs dogmatically, the public information still has a one-time impact on beliefs and prices. But if individuals recognize their limited ability to process information, they will look for others' opinions, the market consensus, to revise their own opinions. When the appropriate conditions are in place (mean-reversion of noise trading), the release of public information triggers a stream of information about its market consensus interpretation. Thus, prices are slow to fully reflect the public information because investors learn gradually about the market consensus. Chan (2003) reaches a similar conclusion: "[...] I find that investors are slow to respond to public news. In other words, the underreaction appears to result not from barriers to knowing news, but to barriers to understanding it. One possibility is that analysts do not give investors private information, but help them digest public news."

Another potential explanation for the slow incorporation of information is limited attention. The model developed in this paper incorporates some aspects of limited attention. Limited attention, understood as limited time to process information, is one source of differential interpretation. Different investors will look at the same piece of information from different angles, and focus on different aspects of the information released. The lack of time to perform an exhaustive analysis contributes to more heterogeneous interpretations. Also, a subsequent observation of the same basic information presented in a different way gives investors the opportunity to look at the information from a different perspective, thus refining their initial interpretation of the information.[26] But limited attention is unlikely to be the only source of differential interpretation. Dedicated financial analysts are less likely to suffer from limited attention; yet their forecasts still vary significantly.

The model, however, does not incorporate limited attention understood as selective attention, where some investors are not even aware that the public information exists (see Hirshleifer and Teoh, 2003 and DellaVigna and Pollet, 2009). Investors gradually becoming aware of the information certainly contributes to slow incorporation of information in prices. But it is unlikely that a fraction of the market large enough to cause a significant price impact stays unaware of public news for periods long enough to explain the post-announcement drift, which lasts for almost 1 year and in some cases even longer.

5.2 Why the Drift Survives, Even with Sophisticated Institutional Investors

As we previously saw, on average prices increase (decrease) following a good (bad) announcement because the market consensus is gradually learned. This means that an automated trading strategy that takes a long (short) position following a good (bad) announcement, yields a positive expected return. This is true even if the strategy is conditional on the interpretation of the signal instead of being conditional on the underlying signal itself.[27] Every investor knows this. Yet, investors

[26] Gilbert, Kogan, Lochstoer, and Ozyildirim (2011) show that prices respond to changes in the Leading Economic Index, which can be computed in advance from previously released information. Huberman and Regev (2001) document a large price reaction to the re-release of news for a bio-tech firm, several months after the initial release of information. However, the extent to which the price reaction to the repackaging of previously available information is due to reassessment of initial interpretations or to the information finally reaching previously inattentive investors, is unclear.

[27] Banerjee et al., 2009 critic Allen et al. (2006) for obtaining a drift conditional on the final realization of the liquidation value, arguing that conditioning on the final price one can always find drifts even when prices follow

rationally choose to trade in a way that does not seem to fully exploit this predictability in returns, allowing the drift to persist.

The reason for this apparently inconsistent behavior lies in the fact that investors only observe noisy signals for the market consensus, and not the market consensus itself. To illustrate this point, consider first the simpler case where the source of information is exclusively exogenous which, as Theorem 8 shows, also produces a post-announcement drift. In this case we obtain (see Appendix C for derivations)

$$
\begin{aligned}
\mathbb{E}\left(\Delta P_2 | \, s\right) &= \frac{Var\left(v | \, \mathcal{F}_1^i\right) - Var\left(v | \, \mathcal{F}_2^i\right)}{\sigma_v^2} s, & \mathbb{E}\left(\Delta P_3 | \, s\right) &= \frac{Var\left(v | \, \mathcal{F}_2^i\right)}{\sigma_v^2} s, \\
\mathbb{E}_i\left[\mathbb{E}\left(\Delta P_2 | \, \tilde{s}_1^i\right)\right] &= \frac{Var\left(v | \, \mathcal{F}_1^i\right) - Var\left(v | \, \mathcal{F}_2^i\right)}{\sigma_v^2 + \sigma_{\tilde{s},1}^2} s, & \mathbb{E}_i\left[\mathbb{E}\left(\Delta P_3 | \, \tilde{s}_1^i\right)\right] &= \frac{Var\left(v | \, \mathcal{F}_2^i\right)}{\sigma_v^2 + \sigma_{\tilde{s},1}^2} s, \\
\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\Delta P_2 | \, \mathcal{F}_1^i\right)\right] &= 0, & \mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\Delta P_3 | \, \mathcal{F}_1^i\right)\right] &= 0.
\end{aligned}
$$

Conditioning on market consensus interpretation (s), prices are expected to drift in the direction of the surprise in s following the announcement, as long as the market consensus is gradually learned (i.e. $Var\left(v | \, \mathcal{F}_2^i\right) < Var\left(v | \, \mathcal{F}_1^i\right)$). Similarly, an investor who conditions only on his announcement interpretation ($\tilde{s}_1^i$), expects prices to drift in the direction of the surprise in $\tilde{s}_1^i$. Since $\tilde{s}_1^i$ is an unbiased signal for s, it follows that the average market participant gets it right and expects prices to drift in the same direction of s. However, when conditioning on *all information available* at date 1 ($\mathcal{F}_1^i$), on average (over net supply shocks) the average market participant expects prices to remain unchanged after the announcement.[28] This means that investors use their information fully to exploit the drift. The existence of a drift is only detected *ex-post* or by someone with *superior information*.

In the model with endogenous information the conclusions are essentially the same, even though the predictability of net supply changes gives rise to hedging demand, which complicates things a little bit (a 3-period example is provided in Appendix C). The only difference is that, on average, prices are expected to change even when conditioning on $\mathcal{F}_1^i$. As I explained before, on average investors form expectations of the liquidation value that are biased toward the prior, since they only receive noisy signals for the market consensus. Consequently, investors attribute part of the price change at the announcement date to a supply shock. To illustrate, consider a positive announcement ($s > 0$), which implies a positive average price response at the announcement date. Also, focus on the average case, where the net supply is zero at all dates. Most investors will believe that a negative supply shock occurred at the announcement date,

$$
\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\theta_1 | \, \mathcal{F}_1^i\right)\right] = -\frac{Var\left(v | \, \mathcal{F}_1^i\right)}{\alpha \sigma_{\tilde{s},1}^2 \sigma_v^2} s < 0,
$$

martingale processes. This criticism does not apply to the model in this paper, because there is a public announcement to condition upon. Moreover, it is not necessary to condition on the unobservable underlying signal to expect a drift, since on average the interpretation of the announcement is unbiased.

[28]Following the announcement of a positive surprise, on average investors with optimistic (pessimistic) interpretations will expect a drift (reversal) in the next period. The opposite applies following a negative surprise. However, on average, market participants do not expect further price changes.

even though on average $\theta_1 = 0$. This shock is expected to dissipate in the periods following the announcement, and so prices are expected to revert during that period, moving in the opposite direction of the announcement surprise,

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\Delta P_2 | \mathcal{F}_1^i\right)\right] = -\frac{Var\left(v|\mathcal{F}_2^i\right)Var\left(v|\mathcal{F}_1^i\right)\left(1-\rho+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\right)\alpha^2\sigma_\theta^2}{\sigma_v^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\left(\sigma_v^2+\sigma_{\tilde{s},1}^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\sigma_v^2\right)}\left(1-\rho\right)s < 0.$$

But, at the same time, investors also expect prices to drift in the direction of the surprise in s. Thus, once the net supply level is close enough to its mean and the adjustment in the net supply level becomes small, prices are expected to revert again and drift in the same direction of s,

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\Delta P_3 | \mathcal{F}_1^i\right)\right] = \frac{Var\left(v|\mathcal{F}_2^i\right)Var\left(v|\mathcal{F}_1^i\right)}{\sigma_{\tilde{s},1}^2\sigma_v^2}\left(1-\rho\right)s > 0.$$

This price change predictability survives in equilibrium because the hedging demand offsets the short-run demand: prices are expected to decrease in the short-run, but they are expected to increase in the long-run. As a result, the positive long-run hedging demand offsets the negative short-run demand, clearing the market (recall that on average the net supply is zero). This means that the expected short-run reversal prevents investors from exploiting the drift as fully as in the model with exogenous information. That is, over the long run investors expect prices to drift,

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\Delta P_2 + \Delta P_3 | \mathcal{F}_1^i\right)\right] = \frac{Var\left(v|\mathcal{F}_2^i\right)Var\left(v|\mathcal{F}_1^i\right)\left[\sigma_v^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\left(\sigma_{\tilde{s},1}^2+\rho\sigma_v^2\right)\right]}{\sigma_{\tilde{s},1}^2\sigma_v^2\left[\sigma_v^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\left(\sigma_v^2+\sigma_{\tilde{s},1}^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\sigma_v^2\right)\right]}\left(1-\rho\right)s > 0.$$

The average investor correctly expects prices to drift when conditioning only on his interpretation ($\tilde{s}_1^i$). However, he cannot do better by exploiting this expectation. Any strategy based only on a subset of the information available will be suboptimal, since investors are rational and fully understand how the economy works. Nonetheless, it is interesting to compare how this and other simple strategies designed to take advantage of the drift compare to the optimal trading strategy. I consider two types of strategies, buy&hold and myopic; and two different information sets, all information available ($\mathcal{F}^i$) and only the announcement interpretation ($\tilde{s}^i$). In both types of strategies, investors trade as in the static model.[29] The difference is that in the buy&hold strategy they trade only at the announcement date, whereas in the myopic strategy they trade every period, adjusting their demands to reflect price changes and updated beliefs (the latter only if using information set $\mathcal{F}^i$).

Table 1 provides the unconditional average, standard deviation, Sharpe ratio and certainty equivalent of the terminal wealth obtained with each of these four strategies plus the optimal strategy, based on the 100,000 simulations of the economy using the calibration described in Section 4.2 and $W_0 = 0$. The numbers in parenthesis correspond to the values conditional on the randomly chosen $P_1 = 0.474$ and $\tilde{s}_1^i = 0.127$. It is clear that, although profitable, the alternative strategies

[29] Demand is given by $X_t^i = \frac{E(v|\mathcal{F}_t)}{\alpha Var(v|\mathcal{F}_t)}$ for some information set $\mathcal{F}_t$.

	Optimal	Buy&Hold ($\mathcal{F}^i$)	Myopic ($\mathcal{F}^i$)	Buy&Hold ($\tilde{s}^i$)	Myopic ($\tilde{s}^i$)
$\mathbb{E}(W_T)$	0.128	0.094	0.105	0.102	0.121
	(0.069)	(0.027)	(0.041)	(0.085)	(0.118)
$Std(W_T)$	0.246	0.253	0.229	0.358	0.336
	(0.169)	(0.117)	(0.111)	(0.365)	(0.357)
$\frac{\mathbb{E}(W_T)}{Std(W_T)}$	0.521	0.370	0.460	0.285	0.360
	(0.410)	(0.233)	(0.370)	(0.233)	(0.331)
$CE(W_T)$	0.0670	0.0429	0.0567	-0.0048	-0.0482
	(0.0363)	(0.0135)	(0.0268)	(-0.0487)	(-0.1021)
Rel $CE(W_T)$	100%	64%	85%	-7%	-72%
	(100%)	(37%)	(74%)	(-134%)	(-281%)

Table 1: Performance of alternative strategies: The performance of the optimal trading strategy is compared with that of 4 alternative strategies that attempt to exploit the price drift. Investors following buy&hold strategies trade as in the static model, but only at the announcement date, and hold their portfolio until the liquidation date. Investors following myopic strategies trade every period as in the static model (i.e. as if they were to hold the portfolio until the liquidation date), adjusting for price changes and, possibly, revisions in beliefs. For each type of strategy, investors can condition their decisions on all information available (announcement interpretation and price, denoted by $\mathcal{F}^i$) or only on their announcement interpretation ($\tilde{s}^i$). The table reports the unconditional average, standard deviation, Sharpe ratio and certainty equivalent of the terminal wealth based on 100,000 simulations of the risky asset, using the calibration detailed in Section 4.2 and $W_0 = 0$. The numbers in parenthesis correspond to the same values conditional on the randomly chosen values $P_1 = 0.474$ and $\tilde{s}_1^i = 0.127$.

are suboptimal. The more the strategy departs from the optimal one, the smaller the Sharpe ratio and the certainty equivalent of wealth: myopic strategies outperform buy&hold strategies, whereas strategies conditioning on all information beat those conditioning only on the announcement interpretation. In particular, ignoring information contained in prices results in a loss of utility compared to not trading at all.[30] These results are similar to those of Biais et al., 2010, who show that in a dynamic REE model with private information momentum strategies may be more profitable than price-contingent strategies (similar to the optimal strategy in this paper) but are, nonetheless, suboptimal.

But what if a fraction of the market participants do not make interpretation errors and directly observe s? We can think of these agents as institutional investors, who have time and resources to correctly interpret the announcement.[31] These institutional investors will expect prices to drift, even when conditioning on all information available to them, and will trade accordingly, buying from (selling to) individual investors when news are good (bad). As a result, prices will incorporate information faster. However, as long as there is residual uncertainty, institutional investors face some risk, since neither prices at future trading dates nor the liquidation value are known with certainty; the former due to future supply shocks, the latter due to residual uncertainty.[32] Risk

[30] It can be formally demonstrated for the 3-period example with endogenous information that a buy&hold strategy conditioning on all available information is the one that provides higher expected utility among all buy&hold strategies, but is outperformed by the optimal dynamic strategy.

[31] Institutional investors do not have to be especially gifted. Instead, they may employ several individuals that form their own interpretation, and then pool those interpretations to form a more precise interpretation. In this sense, what distinguishes institutional investors from individual investors is the ability to directly use the interpretations of several individuals, whereas individual investors have to rely on prices to indirectly observe the interpretations of others.

[32] The risk in exploiting the drift decreases with the distance to the liquidation date. The more trading periods there are until the liquidation date, the smaller chance that one cannot unwind the speculative position with profit

averse institutional investors are then forced to trade conservatively, limiting the price impact and information spillover. Therefore, the drift, albeit attenuated, can still persist in the presence of investors who can interpret the announcement accurately. (An extension of the model to two types of investors with different interpretation precisions and risk aversion is provided in Appendix D.)

In reality, although institutional investors seem to be able to interpret information more precisely than individual investors they cannot do it perfectly; moreover, they face barriers that prevent them to make full use of their informational advantage (Cohen et al., 2002). This reinforces the idea that institutional investors facilitate the incorporation of information in prices, but only to a limited extent. Hence, a drift can persist even in the presence of institutional investors.

5.3 Empirical Evidence on the Post-Announcement Drift

Decades of empirical research on the post-announcement drift phenomenon make it hard to find new testable implications. But, on the bright side, we have a wealth of empirical facts which can be used to validate the model. The most salient empirical facts on the post-announcement drift are:

1. Prices are slow to incorporate information following almost all kinds of announcement: earnings surprises (Ball and Brown, 1968; Foster et al., 1984 and Bernard and Thomas, 1989, 1990); change in analyst recommendations (Michaely and Womack, 1999); share repurchases (Ikenberry et al., 1995); seasoned equity offerings (Loughran and Ritter, 1995); venture capital distributions (Gompers and Lerner, 1998); stock splits (Grinblatt et al., 1984); dividend initiations and omissions (Michaely et al., 1995).[33]

2. The post-announcement drift is long lasting: 3 quarters for earnings surprises (Bernard and Thomas, 1989); 4 to 5 years for tender offers and seasoned equity offerings (Ikenberry et al., 1995 and Loughran and Ritter, 1995).

3. The post-announcement drift is larger for smaller firms, stocks with larger transaction costs, stocks with smaller institutional ownership (Bartov et al., 2000 and Cohen et al., 2002), and for negative news (Cohen et al., 2002).

4. The post-earnings announcement drift is concentrated in the week following the announcement, and in the days surrounding the subsequent quarterly earnings announcements (Bernard and Thomas, 1989).

All of these empirical facts can be explained by the model. According to the model, all public announcements whose information is open to interpretation will originate a drift. This includes the vast majority of public announcements, which explains why the post-announcement drift is so pervasive. One possible exception is the announcement of a price target, since a price target can be immediately used by an investor to make a trading decision. However, it can still be argued that price targets are open to subjective interpretation, because of credibility concerns.

before the liquidation date and has to face the residual uncertainty associated to the liquidation value.

[33] See Appendix A of Daniel et al. (1998) for a more complete list of references on stock price patterns.

In the model the drift results from learning the market consensus interpretation from the observation of subsequent prices. Each price provides just a little extra information about the market consensus, which makes learning process a gradual one. This explains why in general the drift lasts for so long. In particular, the larger the interpretation uncertainty, the longer lasting the drift. It can also explain why the drift is concentrated in the periods immediately following the announcement: investors revise their beliefs more aggressively when the uncertainty about the market consensus is larger. On the other hand, the drift is concentrated around the date of the following announcement because by then investors' still lack a full understanding of the initial announcement; not because they do not understand the process governing the signals, as suggested by Bernard and Thomas (1989). As a result, investors are systematically positively (negatively) surprised by announcements following positive (negative) announcements. This explanation is in line with the evidence provided by Ball and Bartov (1996), who show that investors do recognize the earnings process (unlike Bernard and Thomas 1989 claim), but instead underestimate its parameters.

Finally, the model also predicts that the drift should be larger for stocks with less institutional trading, since institutional investors are assumed to have superior ability to interpret the announcement. Even though the model does not produce predictions directly related to the drift for small stocks, stocks with large transactions costs or for the asymmetry between the drift for good and bad news, Bartov et al. (2000) and Cohen et al. (2002) show that institutional investors are less active in small cap stocks, stocks with large transactions costs and face restrictions on short-selling.[34] In fact, Bartov et al. (2000) show that controlling for institutional ownership, transaction costs and firm size have no impact on the drift.

Another prediction, that has not yet been tested in the context of the post-announcement drift, is that the drift should last longer for stocks with less analyst coverage. Analysts' recommendations and forecasts help to decrease interpretation errors and speed up the incorporation of information on prices, as Chan (2003) suggests. Hong et al. (2000) show that momentum is stronger for stocks with low analyst coverage, which suggests that the same might be true for the post-announcement drift.

But the main, still untested, model prediction is that the larger the interpretation uncertainty, the larger and longer the drift, and the smaller the price reaction to the announcement (see Panel C of Figure 2). This prediction can be tested in two different ways. The direct approach is to use a proxy for interpretation uncertainty and test whether interpretation uncertainty is negatively associated to the immediate price reaction to the announcement and positively associated to the magnitude and duration of the subsequent drift. The indirect approach is to test whether for stocks with similarly surprising news (e.g. earnings) the immediate price response to the announcement is inversely related to the magnitude and duration of the subsequent drift.

There is already some evidence supporting this prediction using the direct approach. Garfinkel and Sokobin (2006) and Choi and Kim (2001) use unexpected volume at earnings announcement date as a proxy for investor opinion divergence. They find a positive association between unexpected

[34]Cohen et al. (2002) provide a comprehensive discussion on the frictions that keep institutional investors from trading more aggressively on their information.

volume and the size of the drift, as the model predicts, but only for positive earnings surprises. Choi and Kim (2001) suggest that short-selling constraints may be an explanation for why the result stands only for positive earnings. Short-selling constraints bar institutional investors from taking advantage of the underreaction to negative news, which has a negative impact on unexpected volume. Controlling for institutional trading may yield the predicted results. In addition, Zhang (2006) finds that the post-analyst forecast revision drift is positively associated with several proxies for information uncertainty.

In turn, Hirshleifer et al. (2009) find a weaker price reaction to earnings announcement and a stronger post-announcement drift in days where there are many earnings announcements made by different firms. This supports the idea that investors have limited time to process information, which leads to a less precise interpretation of information. However, it is unclear to which extent the results are driven by investors rushing their analysis and then looking for others' opinions, as predicted by the model, or by investors missing the initial earnings announcement and catching up with the news later on. One way to distinguish between these two hypotheses is to look at the drift in subsequent calm days. If investors are missing news and trying to catch up when they can, prices should adjust faster in subsequent calm days, when investors have time to go through past news. Also, if investors use weekends to catch up with past news they missed, these drifts should be similar after the first few weekends following the announcements. However, DellaVigna and Pollet (2009) show that the drift is considerably stronger for Friday announcements (investors are likely distracted by the weekend proximity) than for other weekday announcements in the first 60 days following the announcement. This evidence suggests that investors catching up with news they previously missed is not what drives the post-announcement drift.

5.4 Post-Announcement Drift, Hence Momentum?

Chan (2003) shows that momentum is associated to news releases, supporting the view that post-announcement drift and momentum are two faces of the same coin. Chordia and Shivakumar (2006) show that momentum (price momentum) is captured by the systematic component of post-earnings announcement drift (earnings momentum), leading them to conclude that "... the price momentum anomaly is a manifestation of the earnings momentum anomaly...". In face of this evidence, it is tempting to jump to the conclusion that if the model generates post-announcement drift, then it generates momentum as well. However, the relationship between these two anomalies is not mechanical.

The problem is that conditioning on the announcement date return (conditioning on a return over a period that includes the announcement date gives the same results), we capture the price response not only to news but also to the supply shock. For example, a portfolio of stocks with a large announcement day return will include both stocks with positive news and stocks with a negative supply shock (and not so positive news). The price of the former will drift upwards, contributing to momentum. But the price of the latter will revert, and for two reasons. First, because the negative net supply will revert to its mean, pushing prices down. Second, because

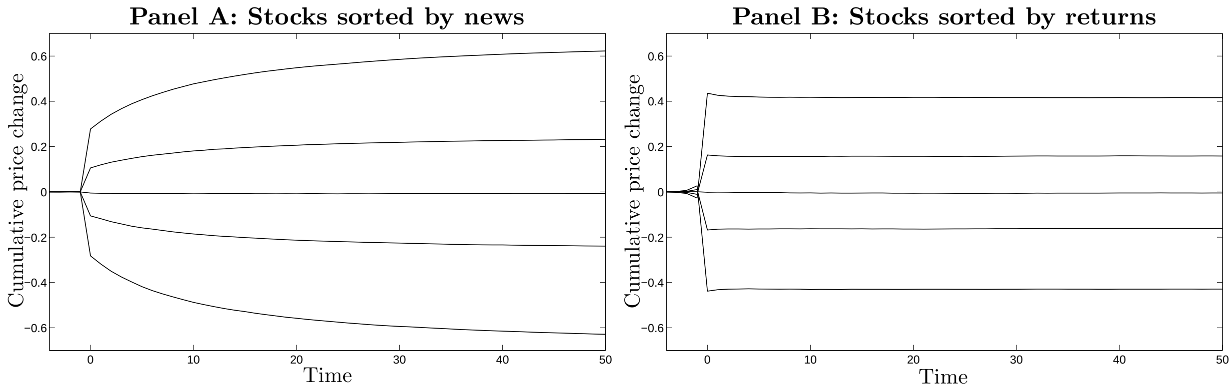

Figure 4: Drift conditional on news vs. drift conditional on announcement date return. This figure plots the cumulative price change from dates -5 to 50 for quintile portfolios formed by sorting stocks by news (Panel A) and by the announcement date price change (Panel B). The announcement date is normalized to date 0. Each of the 100,000 stocks is priced by the price function (5), and corresponds to one realization of the vector of random variables $\{s, \varepsilon_{\theta,t} : 1 \le t \le 106\}$. The parametrization used is: $T = 107$, $n = 1$, $\sigma_v^2 = 0.25$, $\Sigma_s = 0$, $\Sigma_{\tilde{s},6} = 1$, $\Sigma_{\tilde{s},t} = 10^{10}$ $(1 < t \le T-1 \wedge t \neq 6)$, $\sigma_\theta^2 = 0.1$, $\rho = 0.5$, $\alpha = 2$.

over-optimistic expectations of the liquidation value, caused by the negative supply shock, will be gradually corrected. Since investors are able to flawlessly extract information from prices (although not from announcements), the expectation of the liquidation value conditional on prices is, on average, unbiased.[35] This means that, if not for the mean reversion of the, on average, negative supply shock, there would be neither momentum nor reversal when conditioning on returns. The slight reversal the model produces (see Panel B of Figure 4) is due to the mean-reversion of net supply.

Even though this result seems disappointing, Foster et al. (1984) find very similar empirical results to those illustrated in Figure 4. There is a drift when conditioning on unexpected earnings, but no drift with a very slight reversal when conditioning on the abnormal return around the announcement date. Their results are strikingly similar to what we obtain from the model. What we can take from this is that it matters whether we condition on the event, or on the price reaction to the event. Researchers that use the latter as a proxy for the former, should keep this in mind.

The apparent contradiction in the findings of Foster et al. (1984) and Chan (2003) is likely to be explained by differences in methodology (event-time vs. calendar-time) and by the fact that Foster et al. (1984) focus on the first 60 days following the announcement, whereas Chan (2003) considers a longer horizon. Chan (2003) finds a slight reversal in the first month after portfolio formation, and only a slight drift in the first month when excluding the first week after portfolio formation. In addition, portfolio formation based on monthly returns, as used by Chan (2003), means that the subsequent performance analysis misses anything between 1 to 5 weeks of returns after the news event. This means that the drift document by Chan (2003) starts close to the end of the horizon considered by Foster et al. (1984), which reconciles both studies.

The question now is how can we modify the model to generate a drift conditional on the an-

[35] Banerjee et al. (2009) also point out that there is no drift conditional on prices in REE models because investors form unbiased expectations from prices.

nouncement day return, i.e. momentum, in addition to a drift conditional on the announcement. The latter is obtained by assuming that investors cannot fully understand the announcement, thus making imperfect interpretation of the announcement that lead them to form beliefs biased toward the prior. Yet, it is assumed that investors fully understand how the equilibrium price is formed, allowing them to flawlessly extract information from prices. As a result, beliefs conditional on prices are unbiased. Since there is no bias to be corrected, the gradual learning of the market consensus does not generate momentum. Although the assumption that investors can flawlessly extract information from prices is useful to derive the equilibrium in a rational expectations framework, it is hardly a realistic one. Investors need to take in consideration the actions of all other market participants in order to correctly determine the coefficients of the price equation, which is anything but straightforward.

So, let us entertain the idea that investors make unbiased and independent mistakes when extracting information from prices. The additional source of noise decreases the precision of the information extracted from prices, making beliefs conditional on prices become biased toward the prior. As investors learn more about the market consensus from subsequent prices (because supply shocks, extraction errors, or both, revert to their mean), this bias is gradually corrected, generating momentum.

Mistakes in the extraction of information from prices can be proxied within the current model by assuming that net supply shocks are more volatile than they really are (see Panel A of Figure 5).[36] That is, being σ_θ the true volatility of net supply shocks, $\hat{\sigma}_\theta > \sigma_\theta$ is used to determine the coefficients of the price equation, whereas σ_θ is used to simulate net supply shocks. This means that the volatility of net supply shocks within the model ($\hat{\sigma}_\theta$) now lumps together the true volatility of net supply shocks (σ_θ) and the volatility of the information extraction error. In this case, the parameter ρ will be a function of the speed of mean reversion of the two sources of noise in prices. Extraction errors more (less) persistent than supply shocks are proxied by a $\hat{\rho}$ larger (smaller) than its real value ρ (see Panel B of Figure 5).

With this ad hoc modification to the model, momentum and post-announcement drift have the same underlying cause: differential interpretation of the public information and gradual learning of the market consensus. As a result, the model generates similar predictions for the post-announcement drift and momentum, which is in accordance with the similarities between the empirical evidence on post-announcement drift and momentum. In particular, the empirical findings of Hong et al. (2000) and Chan (2003) are consistent with the model. Momentum is stronger for small illiquid looser stocks with news, which are likely to attract less institutional trading; and for stocks with low analyst coverage, which implies that there is less help from analysts in interpreting the announcement.

In addition, the model can now shed some light on why post-earnings announcement drift and momentum strategies do not subsume each other. Chan et al. (1996) find that momentum

[36]An alternative is to consider that investors do not make mistakes in the extraction of information from prices, but systematically overestimate the volatility of the supply shock, thus extracting a less than optimal amount of information from prices. However, there is no apparent reason why this should be the case.

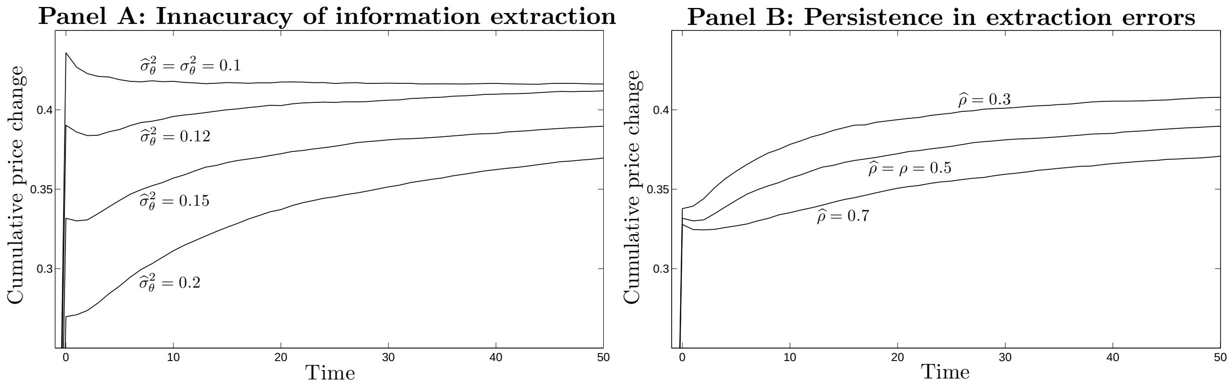

Figure 5: Inaccurate extraction of information from prices and the drift conditional on announcement date return. This figure plots the cumulative price change from dates -5 to 50 for the top quintile portfolio when proxying for investors making unbiased mistakes while extracting information from prices. Inaccurate extraction of information from prices is proxied by considering a larger supply shock volatility in the model ($\widehat{\sigma}_\theta^2$) than its real value ($\sigma_\theta^2$); the persistence of the supply shock used in the model ($\widehat{\rho}$) may be different than its real value (ρ) if the persistence of the true supply shocks differs from the persistence of the mistakes in the information extraction process. Panel A shows the effect of inaccurate extraction of information (with $\widehat{\rho} = \rho$). Panel B shows the effect of different degrees of persistence of the mistakes in the information extraction process when $\widehat{\sigma}_\theta^2 = 0.15 > \sigma_\theta^2$; $\widehat{\rho} < (>)\, \rho$ means that these mistakes are less (more) persistent than supply shocks. Everything else is as in Figure 4.

(price momentum) and post-earnings announcement drift (earnings momentum) are both profitable after controlling for the other. Brandt et al. (2008) find similar results when using the abnormal return around the announcement date instead of past returns as the basis of their price momentum strategy. Both authors hypothesize that returns contain more than just information about earnings. This would explain why price momentum strategies are more profitable and are not subsumed by earnings momentum strategies, but not why earnings momentum strategies are not subsumed by price momentum strategies. We can use the model to rationalize these findings. Consider that two signals are announced simultaneously: one is earnings related, the other non-earnings related. Econometricians can only obtain historical information about the earnings related signal, allowing them to compute the standardized unexpected earnings (SUE). Within the context of the model, knowing the SUE is like knowing the market consensus interpretation of the earnings related signal. A strategy based on high minus low SUE/market consensus stocks generates a positive return, because of the post-earnings announcement drift. It has the advantage of conditioning on a (more) direct measure of one of the signals, but completely ignores the other signal. In turn, a strategy based on high minus low earnings announcement return (EAR), generates a positive return because prices also drift when conditioning on the price reaction to the announcement. This strategy has the advantage of taking into account both signals, but it is only an indirect measure of these signals, plagued by noise trading. Therefore, both measures provide extra information over the other, reason why Brandt et al. (2008) find that both strategies are largely independent.[37] For example, a high SUE stock may have good or bad non-earnings related news. And a high EAR stock may have good earnings and non-earnings news or not so good news and a large non-informationally driven

[37] By the contrary, Sinha (2011) shows that when considering not just earnings announcements, but all news releases, a strategy based on news tonality subsumes the price momentum strategy.

demand (noise trading). But a stock with high SUE and high EAR is likely to have good earnings and non-earnings related news, especially if the goodness of both pieces of information is positively associated.

Unlike Barberis et al. (1998), Daniel et al. (1998) and Hong and Stein (1999), the model in this paper does not account for long-term reversal. Although this could be seen as a limitation, empirical evidence suggests that momentum and reversal are two unrelated phenomena, likely to be explained by different factors. For instance, Chan (2003) finds reversal for stocks with and without news, although momentum is an exclusive of stocks with news. And the 3-factor model of Fama and French (1996) successfully account for long-term reversals, but not for short-term momentum.

Compared to the models developed by Barberis et al. (1998) and Daniel et al. (1998), the advantage of this model is that it does not rely on specific behavioral biases to explain why prices are slow to incorporate information. Instead, this model builds on the intuitive notion that people process information differently when exposed to the same evidence. In a certain sense, this model is closest to the model of Hong and Stein (1999). The difference is that I do not need to exogenously impose gradual diffusion of information. Instead, this comes as an endogenous result of the model. It is important to note that Hong and Stein's assumption that momentum traders do not extract information from prices is more than a simplifying assumption to avoid a fully revealing REE in the absence of noise trading. Instead, it is a crucial assumption to generate momentum. Like in our model, there is no momentum if investors can correctly extract information from prices, regardless of the presence of noise trading.

6 Conclusion

In this paper I developed a model that incorporates the idea of differential interpretation of information and is able to generate a post-announcement drift. Differential interpretation of information transforms public raw information into private interpreted information. If investors recognize their limited ability to interpret the announcement, they will look for other investors' opinions in prices. Noise trading prevents investors from learning the market consensus interpretation of the announcement from the observation of a single price. But if noise trading is mean-reverting, investors can gradually learn the market consensus from the observation of a series of prices. As investors become more confident about their interpretation of the public announcement, they put more weight on it, and the information contained in the announcement becomes gradually incorporated into prices, generating a post-announcement drift.

Even if differential interpretation of information is not *the* explanation for the post-announcement drift, the model's ability to explain the extant empirical evidence suggests that at least it plays an *important* role. The attractiveness of the model also relies on its parsimony and plausible assumptions.

The model also highlights that gradual information flow is a necessary but not a sufficient condition to generate momentum. As long as investors flawlessly extract information from prices, they form unbiased beliefs conditional on returns and there is no momentum. This means that in Hong and Stein (1999), momentum does not automatically follow from the exogenous assumption that private information flows gradually into the market. To obtain momentum it is also crucial that those agents who observe private information do not extract information from prices.

In reality, though, the truth is likely to lie between these two extremes: investors extract information from prices, but do not do it perfectly. This allows differential interpretation of information to generate both post-announcement drift and momentum, providing a unified explanation for both phenomena. Even though the model does not account for long-run reversals, this is not necessarily a weakness. The current empirical evidence suggests that post-announcement drift and momentum are more likely to be related with each other, than with long-run reversal.

Appendix

A Proofs

To prove the results of the paper, I use the following lemmas.

Lemma 9. ***(Gaussian filtering)*** *For every t, let the m_1-dimensional vector of state variables at date t, x_t, and the m_2-dimensional vector of observations at date t, y_t, follow the processes*

$$x_t = A_t x_{t-1} + \varepsilon_{x,t},\ y_t = B_t x_t + \varepsilon_{y,t},\ t = 1, 2, ...$$

where the m_3-dimensional vector $\varepsilon_{x,t} \sim N(0, \Sigma_{x,t})$ and the m_4-dimensional vector $\varepsilon_{y,t} \sim N(0, \Sigma_{y,t})$ are independent Gaussian vectors and A_t and B_t are matrices of the appropriate order. $x_0 \sim N(\overline{x}_0, \Sigma_0)$ and independent of $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$. Let $\mathcal{F}_t = \{y_\tau\}_{\tau=1,\ldots,t}$. Then,

$$\begin{aligned}
\mathbb{E}(x_t|\mathcal{F}_t) &= A_t\mathbb{E}(x_{t-1}|\mathcal{F}_{t-1}) + K_t[y_t - \mathbb{E}(y_t|\mathcal{F}_{t-1})] \\
&= A_t\mathbb{E}(x_{t-1}|\mathcal{F}_{t-1}) + K_t[y_t - B_tA_t\mathbb{E}(x_{t-1}|\mathcal{F}_{t-1})] \\
Var(x_t|\mathcal{F}_t) &= (I_{m_1} - K_tB_t)[A_tVar(x_{t-1}|\mathcal{F}_{t-1})A_t' + \Sigma_{x,t}] \qquad (15) \\
K_t &= [A_tVar(x_{t-1}|\mathcal{F}_{t-1})A_t' + \Sigma_{x,t}]B_t'\{B_t[A_tVar(x_{t-1}|\mathcal{F}_{t-1})A_t' + \Sigma_{x,t}]B_t' + \Sigma_{y,t}\}^{-1}.
\end{aligned}$$

Proof. These results are immediate from Theorem 13.4 and Corollary 2 in Liptser and Shiryaev (2001, pp.71-72, vol. II) □

Lemma 10. *Let $\phi \equiv a + b'\varepsilon + \varepsilon' c\varepsilon$, where $\varepsilon \sim N(0, \Sigma)$ denotes a Gaussian vector, a is a constant scalar, b is a constant vector and c a constant symmetric matrix of appropriate dimensions. If $\Sigma^{-1} - 2c$ is a symmetric positive definite matrix, then*

$$\mathbb{E}\left(e^{\phi}\right) = \frac{e^{a+\frac{1}{2}b'\left(\Sigma^{-1}-2c\right)^{-1}b}}{\sqrt{|\Sigma|\,|\Sigma^{-1}-2c|}}.$$

Proof. Writing the expectation $\mathbb{E}\left(e^{\phi}\right)$ explicitly we obtain

$$\mathbb{E}\left(e^{\phi}\right) = e^a \int e^{b'\varepsilon+\varepsilon' c\epsilon} \frac{e^{-\frac{1}{2}\varepsilon'\Sigma^{-1}\varepsilon}}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}} d\varepsilon.$$

To eliminate the integral, rearrange the terms in order to obtain the integral of the probability density function of a normal distributed vector multiplied by a constant. Proceeding in this way, we find that

$$\begin{aligned}
e^a \int e^{b'\varepsilon+\varepsilon' c\epsilon} \frac{e^{-\frac{1}{2}\varepsilon'\Sigma^{-1}\varepsilon}}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}} d\varepsilon &= \frac{e^{a+\frac{1}{2}b'\left(\Sigma^{-1}-2c\right)^{-1}b}}{\sqrt{|\Sigma|\,|\Sigma^{-1}-2c|}} \int \frac{e^{-\frac{1}{2}\left[\varepsilon-\left(\Sigma^{-1}-2c\right)^{-1}b\right]'\left(\Sigma^{-1}-2c\right)\left[\varepsilon-\left(\Sigma^{-1}-2c\right)^{-1}b\right]}}{(2\pi)^{\frac{N}{2}}\left|\left(\Sigma^{-1}-2c\right)^{-1}\right|^{\frac{1}{2}}} d\varepsilon \\
&= \frac{e^{a+\frac{1}{2}b'\left(\Sigma^{-1}-2c\right)^{-1}b}}{\sqrt{|\Sigma|\,|\Sigma^{-1}-2c|}}
\end{aligned}$$

which proves the lemma. □

Lemma 11. ***(Conditional Normal distribution)*** *Let the m_1-dimensional vector x_1 and the m_2-dimensional* *vector* x_2 *be jointly normal*

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}\right).$$

Then, the distribution of x_1 *conditional on* x_2 *is*

$$x_1 | \, x_2 \sim N\left(\bar{x}_1 + \Sigma_{12}\Sigma_{22}^{-1}\left(x_2 - \bar{x}_2\right), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'\right).$$

Proof. This well know result is adapted from Theorem 13.1 of Liptser and Shiryaev (2001, p.61, vol. II). □

A.1 Proof of Lemma 2

The proof of the lemma is a straightforward application of Lemma 9. All we need is to identify the relevant unobservable state variable and observation processes, and the prior beliefs.

Starting with the beliefs conditional on total information ($\mathcal{F}_t^i$), the vector of date t unobservable state variables and observations are

$$\underset{(n+2\times 1)}{x_t} = \begin{bmatrix} v \\ \boldsymbol{s} \\ \theta_t \end{bmatrix}, \quad \underset{(n+1\times 1)}{y_t} = \begin{bmatrix} \xi_t \\ \tilde{\boldsymbol{s}}_t^i \end{bmatrix}, \, 1 \le t \le T-1.$$

The only time dependent state variable is θ_t, whose process is given by equation (1). We can then define the process for the vector of state variables as

$$\underset{(n+2\times n+2)}{A_t} = \begin{bmatrix} I_{n+1} & \mathbf{0}_{(n+1\times 1)} \\ \mathbf{0}_{(1\times n+1)} & \rho \end{bmatrix}, \quad \underset{(n+2\times n+2)}{\Sigma_{x,t}} = \begin{bmatrix} \mathbf{0}_{(n+1\times n+1)} & \mathbf{0}_{(n+1\times 1)} \\ \mathbf{0}_{(1\times n+1)} & \sigma_\theta^2 \end{bmatrix} \; t = 1, 2, ..., T-1.$$

Using equation (3), the prior expectation and variance of the vector of state variables are

$$\mathbb{E}\left(x_0 | \, \mathcal{F}_0\right) = \mathbf{0}_{(n+2\times 1)}, \, Var\left(x_0 | \, \mathcal{F}_0\right) = \begin{bmatrix} \sigma_v^2 & \sigma_v^2 \mathbf{1}_{(1\times n)} & 0 \\ \sigma_v^2 \mathbf{1}_{(n\times 1)} & \sigma_v^2 \mathbf{1}_{(n\times n)} + \Sigma_s & \mathbf{0}_{(n\times 1)} \\ 0 & \mathbf{0}_{(1\times n)} & 0 \end{bmatrix}.$$

The process for the vector of observations is easily determined from equation (4) and the definition of ξ_t, equation (6),

$$\underset{(n+1\times n+2)}{B_t} = \begin{bmatrix} 0 & \boldsymbol{p}_t & p_{\theta,t} \\ \mathbf{0}_{(n\times 1)} & I_n & \mathbf{0}_{(n\times 1)} \end{bmatrix}, \quad \underset{(n+1\times n+1)}{\Sigma_{y,t}} = \begin{bmatrix} 0 & \mathbf{0}_{(1\times n)} \\ \mathbf{0}_{(n\times 1)} & \Sigma_{\tilde{s},t} \end{bmatrix} \; t = 1, 2, ..., T-1$$

In the case of beliefs conditional on common information ($\mathcal{F}_t^c$), the only difference is that the vector of observations reduces to one element, $y_t = \xi_t$. Therefore x_t, A_t, $\Sigma_{x,t}$, $\mathbb{E}\left(x_0 | \, \mathcal{F}_0\right)$ and $Var\left(x_0 | \, \mathcal{F}_0\right)$ are defined as above, whereas y_t and B_t correspond to the first row and $\Sigma_{y,t}$ corresponds to the first entry of the vector/matrices defined above.

Applying Lemma 9 to the state variable and observation vectors as defined above, we obtain the filtering equations (7) and (8) as well as matrices K_t and K_t^c.

A.2 Proof of Lemma 3

I will start by deriving the processes for $\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t+1}^{i}\right)$, $\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t+1}^{c}\right)$ and $\mathbb{E}\left(\theta_t|\,\mathcal{F}_{t+1}^{i}\right)$ for $t=1,2,...,T-2$, from which the process for $\boldsymbol{\psi}_{t+1}$ can be easily determined. Let K_t and K_t^c be partitioned as follows

$$\underset{(n+2\times n+1)}{K_t} = \begin{bmatrix} \underset{(1\times1)}{k_{v,\xi,t}} & \underset{(1\times n)}{k_{v,\boldsymbol{s},t}} \\ \underset{(n\times1)}{k_{\boldsymbol{s},\xi,t}} & \underset{(n\times n)}{k_{\boldsymbol{s},\boldsymbol{s},t}} \\ \underset{(1\times1)}{k_{\theta,\xi,t}} & \underset{(1\times n)}{k_{\theta,\boldsymbol{s},t}} \end{bmatrix}, \quad \underset{(n+2\times1)}{K_t^c} = \begin{bmatrix} \underset{(1\times1)}{k_{v,\xi,t}^c} \\ \underset{(n\times1)}{k_{\boldsymbol{s},\xi,t}^c} \\ \underset{(1\times1)}{k_{\theta,\xi,t}^c} \end{bmatrix},$$

where, for instance, $k_{\boldsymbol{s},\xi,t}$ is an n-dimensional column vector corresponding to entries of K_t which gives the impact of the surprise in ξ_t on the beliefs about the vector of signals $\boldsymbol{s}$.

Starting from equation (7), substituting for ξ_{t+1} and θ_{t+1} using equations (6) and (1), we obtain

$$\begin{aligned} \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t+1}^{i}\right) &= \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t}^{i}\right) + k_{\boldsymbol{s},\xi,t+1}\left[\xi_{t+1} - \mathbb{E}\left(\xi_{t+1}|\,\mathcal{F}_t^i\right)\right] + k_{\boldsymbol{s},\boldsymbol{s},t+1}\left[\boldsymbol{s}_{t+1}^i - \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^i\right)\right] \\ &= \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t}^{i}\right) + k_{\boldsymbol{s},\xi,t+1}\left[\boldsymbol{p}_{t+1}\boldsymbol{s} + p_{\theta,t+1}\theta_{t+1} - \mathbb{E}\left(\boldsymbol{p}_{t+1}\boldsymbol{s} + p_{\theta,t+1}\theta_{t+1}\middle|\,\mathcal{F}_t^i\right)\right] \\ &\quad + k_{\boldsymbol{s},\boldsymbol{s},t+1}\left[\boldsymbol{s} + \boldsymbol{\varepsilon}_{t+1}^i - \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^i\right)\right] \\ &= \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t}^{i}\right) + \left(k_{\boldsymbol{s},\xi,t+1}\boldsymbol{p}_{t+1} + k_{\boldsymbol{s},\boldsymbol{s},t+1}\right)\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^i\right)\right] + k_{\boldsymbol{s},\boldsymbol{s},t+1}\boldsymbol{\varepsilon}_{t+1}^i \\ &\quad + k_{\boldsymbol{s},\xi,t+1}p_{\theta,t+1}\varepsilon_{\theta,t+1} + \rho k_{\boldsymbol{s},\xi,t+1}p_{\theta,t+1}\left[\theta_t - \mathbb{E}\left(\theta_t|\,\mathcal{F}_t^i\right)\right]. \end{aligned} \tag{16}$$

From the definition of ξ_t, equation (6), we obtain

$$\xi_t = \boldsymbol{p}_t\boldsymbol{s} + p_{\theta,t}\theta_t \Leftrightarrow \theta_t = \frac{\xi_t - \boldsymbol{p}_t\boldsymbol{s}}{p_{\theta,t}}. \tag{17}$$

Since ξ_t is adapted to $\mathcal{F}_t^i$, we have that $\xi_t - \mathbb{E}\left(\xi_t|\,\mathcal{F}_t^i\right) = 0$. Substituting equation (17) into (16) and rearranging, we finally obtain the process for $\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t+1}^{i}\right)$,

$$\begin{aligned} \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t+1}^{i}\right) &= \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t}^{i}\right) + \left[k_{\boldsymbol{s},\boldsymbol{s},t+1} + k_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\right]\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^i\right)\right] \\ &\quad + k_{\boldsymbol{s},\boldsymbol{s},t+1}\boldsymbol{\varepsilon}_{t+1}^i + k_{\boldsymbol{s},\xi,t+1}p_{\theta,t+1}\varepsilon_{\theta,t+1}. \end{aligned}$$

The process for $\mathbb{E}\left(\theta_{t+1}|\,\mathcal{F}_{t+1}^{i}\right)$ is derived exactly in the same way and is similar to the process for $\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t+1}^{i}\right)$. The only difference is that instead of $\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t}^{i}\right)$, $k_{\boldsymbol{s},\xi,t+1}$ and $k_{\boldsymbol{s},\boldsymbol{s},t+1}$, we have $\rho\mathbb{E}\left(\theta_t|\,\mathcal{F}_t^i\right)$, $k_{\theta,\xi,t+1}$ and $k_{\theta,\boldsymbol{s},t}$, respectively.

Starting from equation (8) and following similar steps as in (16) we have that

$$\begin{aligned} \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t+1}^{c}\right) &= \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t}^{c}\right) + k_{\boldsymbol{s},\xi,t+1}^c\boldsymbol{p}_{t+1}\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right)\right] + k_{\boldsymbol{s},\xi,t+1}^c p_{\theta,t+1}\varepsilon_{\theta,t+1} \\ &\quad + \rho k_{\boldsymbol{s},\xi,t+1}^c p_{\theta,t+1}\left[\theta_t - \mathbb{E}\left(\theta_t|\,\mathcal{F}_t^c\right)\right]. \end{aligned} \tag{18}$$

Substituting equation (17) into (18), and noting that ξ_t is adapted to $\mathcal{F}_t^c$, we obtain

$$\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t+1}^{c}\right) = \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_{t}^{c}\right) + k_{\boldsymbol{s},\xi,t+1}^c\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right)\right] + k_{\boldsymbol{s},\xi,t+1}^c p_{\theta,t+1}\varepsilon_{\theta,t+1}$$

$$
\begin{aligned}
&= \mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^c\right) + k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\left[\mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^i\right) - \mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^c\right)\right] \\
&\quad + k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^i\right)\right] + k^c_{\boldsymbol{s},\xi,t+1}p_{\theta,t+1}\varepsilon_{\theta,t+1}.
\end{aligned} \tag{19}
$$

Bringing together the results above it is straightforward to define

$$
\underset{(2n+2\times 2n+2)}{F_{t+1}} = \begin{bmatrix} 1 & \mathbf{0}_{(1\times n)} & \mathbf{0}_{(1\times n)} & 0 \\ \mathbf{0}_{(n\times 1)} & I_n & \mathbf{0}_{(n\times n)} & \mathbf{0}_{(n\times 1)} \\ \mathbf{0}_{(n\times 1)} & k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right) & I_n - k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right) & \mathbf{0}_{(n\times 1)} \\ 0 & \mathbf{0}_{(1\times n)} & \mathbf{0}_{(1\times n)} & \rho \end{bmatrix} \tag{20}
$$

and

$$
\underset{(2n+2\times 2n+2)}{G_{t+1}} = \begin{bmatrix} 0 & \mathbf{0}_{(1\times n)} & \mathbf{0}_{(1\times n)} & 0 \\ \mathbf{0}_{(n\times 1)} & k_{\boldsymbol{s},\boldsymbol{s},t+1} + k_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right) & k_{\boldsymbol{s},\boldsymbol{s},t+1} & k_{\boldsymbol{s},\xi,t+1}p_{\theta,t+1} \\ \mathbf{0}_{(n\times 1)} & k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right) & \mathbf{0}_{(n\times n)} & k^c_{\boldsymbol{s},\xi,t+1}p_{\theta,t+1} \\ 0 & k_{\theta,\boldsymbol{s},t+1} + k_{\theta,\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right) & k_{\theta,\boldsymbol{s},t+1} & k_{\theta,\xi,t+1}p_{\theta,t+1} \end{bmatrix}, \tag{21}
$$

$t = 1, 2, ..., T-2$.

The process for ΔP_{t+1} is obtained in a similar way as the one for $\boldsymbol{\psi}_{t+1}$. However, because of the risky asset's liquidation at time T, the process for ΔP_{t+1} at $t = T-1$ is distinct from the process at previous dates. I will first derive the process for ΔP_{t+1} at $t = 1, 2, ..., T-2$. Using the price function (5), substituting for θ_{t+1} using (1) and for $\mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_{t+1}^c\right)$ using (19), we have

$$
\begin{aligned}
\Delta P_{t+1} &= \hat{\boldsymbol{p}}_{t+1}\mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_{t+1}^c\right) + \boldsymbol{p}_{t+1}\boldsymbol{s} + p_{\theta,t+1}\theta_{t+1} - \hat{\boldsymbol{p}}_t\mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^c\right) - \boldsymbol{p}_t\boldsymbol{s} - p_{\theta,t}\theta_t \qquad (22) \\
&= \left[\boldsymbol{p}_{t+1} - \boldsymbol{p}_t + \hat{\boldsymbol{p}}_{t+1}k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\right]\mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^i\right) \\
&\quad + \left[\hat{\boldsymbol{p}}_{t+1} - \hat{\boldsymbol{p}}_t - \hat{\boldsymbol{p}}_{t+1}k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\right]\mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^c\right) \\
&\quad + \left[\boldsymbol{p}_{t+1} - \boldsymbol{p}_t + \hat{\boldsymbol{p}}_{t+1}k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\right]\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^i\right)\right] \\
&\quad + p_{\theta,t+1}\left(1 + \hat{\boldsymbol{p}}_{t+1}k^c_{\boldsymbol{s},\xi,t+1}\right)\varepsilon_{\theta,t+1} + \left(p_{\theta,t+1}\rho - p_{\theta,t}\right)\theta_t. \qquad (23)
\end{aligned}
$$

To get rid of the term θ_t , we use the fact that ξ_t is adapted to $\mathcal{F}_t^i$ and its definition to obtain

$$
\xi_t = \mathbb{E}\left(\xi_t\middle|\,\mathcal{F}_t^i\right) \Leftrightarrow \theta_t = \mathbb{E}\left(\theta_t\middle|\,\mathcal{F}_t^i\right) - \frac{\boldsymbol{p}_t}{p_{\theta,t}}\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}\middle|\,\mathcal{F}_t^i\right)\right]. \tag{24}
$$

Substituting equation (24) into (23) we can then define the process for ΔP_{t+1} at $t = 1, 2, ..., T-2$ as

$$
\underset{(1\times 2n+2)}{C_{t+1}} = \begin{bmatrix} 0 \\ \left[\boldsymbol{p}_{t+1} - \boldsymbol{p}_t + \hat{\boldsymbol{p}}_{t+1}k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\right]' \\ \left[\hat{\boldsymbol{p}}_{t+1} - \hat{\boldsymbol{p}}_t - \hat{\boldsymbol{p}}_{t+1}k^c_{\boldsymbol{s},\xi,t+1}\left(\boldsymbol{p}_{t+1} - \rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right)\right]' \\ p_{\theta,t+1}\rho - p_{\theta,t} \end{bmatrix}' \tag{25}
$$

and

$$\underset{(1\times 2n+2)}{D_{t+1}} = \left[\begin{array}{ccccc} 0 & \left(1+\hat{\boldsymbol{p}}_{t+1}k^c_{\boldsymbol{s},\xi,t+1}\right)\left(\boldsymbol{p}_{t+1}-\rho\frac{p_{\theta,t+1}}{p_{\theta,t}}\boldsymbol{p}_t\right) & \mathbf{0}_{(1\times n)} & \left(1+\hat{\boldsymbol{p}}_{t+1}k^c_{\boldsymbol{s},\xi,t+1}\right)p_{\theta,t+1} \end{array} \right]. \tag{26}$$

At $t = T-1$, the excess return becomes $\Delta P_T = v - P_{T-1}$. Using the price function (5) and equation (24), we obtain

$$\begin{aligned} \Delta P_T &= v - \hat{\boldsymbol{p}}_{T-1}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}^c_{T-1}\right) - \boldsymbol{p}_{T-1}\boldsymbol{s} - p_{\theta,T-1}\theta_{T-1} \\ &= \mathbb{E}\left(v|\,\mathcal{F}^i_{T-1}\right) - \boldsymbol{p}_{T-1}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}^i_{T-1}\right) - \hat{\boldsymbol{p}}_{T-1}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}^c_{T-1}\right) - p_{\theta,T-1}\mathbb{E}\left(\theta_{T-1}|\,\mathcal{F}^i_{T-1}\right) \\ &\quad + v - \mathbb{E}\left(v|\,\mathcal{F}^i_{T-1}\right). \end{aligned} \tag{27}$$

To write ΔP_T as a function of $\boldsymbol{\psi}_{T-1}$ and $\boldsymbol{\varepsilon}_{\Delta,T}$, we need to get rid of the first $\mathbb{E}\left(v|\,\mathcal{F}^i_{T-1}\right)$ term. To that end, assume for now that the vector of signals $\boldsymbol{s}$ is observable. Equation (3) defines $\boldsymbol{s}$ as a vector of Gaussian signals for v. Therefore, we can employ the linear Gaussian filter of Lemma 9 to find $\mathbb{E}\left(v|\,\boldsymbol{s}\right)$. For that, we make the substitutions

$$x = v,\ y = \boldsymbol{s},\ A = 1,\ \Sigma_x = 0,\ B = \mathbf{1}_{(n\times 1)},\ \Sigma_y = \Sigma_s.$$

Let $\hat{K}$ be determined according to Lemma 9 for the unobservable and observations processes defined above. Then,

$$\mathbb{E}\left(v|\,\boldsymbol{s}\right) = \mathbb{E}\left(v|\,\mathcal{F}_0\right) + \hat{K}\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_0\right)\right] = \hat{K}\boldsymbol{s}, \tag{28}$$

since $\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_0\right) = \mathbb{E}\left(v|\,\mathcal{F}_0\right) = 0$. Taking expectations conditional on $\mathcal{F}^i_{T-1}$ and using the law of iterated expectations, we obtain

$$\mathbb{E}\left(v|\,\mathcal{F}^i_{T-1}\right) = \hat{K}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}^i_{T-1}\right). \tag{29}$$

Substituting equation (29) into (27), we obtain

$$\underset{(1\times 2n+2)}{C_T} = \left[\begin{array}{cccc} 0 & \hat{K} - \boldsymbol{p}_{T-1} & -\hat{\boldsymbol{p}}_{T-1} & -p_{\theta,T-1} \end{array} \right],\ \underset{(1\times 2n+2)}{D_T} = \left[\begin{array}{cccc} 1 & \mathbf{0}_{(1\times n)} & \mathbf{0}_{(1\times n)} & 0 \end{array} \right]. \tag{30}$$

Finally, $\Sigma_{\Delta,t+1}$ is given by

$$\underset{(2n+2\times 2n+2)}{\Sigma_{\Delta,t+1}} = \left[\begin{array}{cccc} Var\left(v|\,\mathcal{F}^i_t\right) & Cov\left(v,\boldsymbol{s}|\,\mathcal{F}^i_t\right) & \mathbf{0}_{(1\times n)} & 0 \\ Cov\left(\boldsymbol{s},v|\,\mathcal{F}^i_t\right) & Var\left(\boldsymbol{s}|\,\mathcal{F}^i_t\right) & \mathbf{0}_{(n\times n)} & \mathbf{0}_{(n\times 1)} \\ \mathbf{0}_{(n\times 1)} & \mathbf{0}_{(n\times n)} & \Sigma_{\tilde{s},t+1} & \mathbf{0}_{(n\times 1)} \\ 0 & \mathbf{0}_{(1\times n)} & \mathbf{0}_{(1\times n)} & \sigma^2_\theta \end{array} \right] \tag{31}$$

where $Var\left(\boldsymbol{s}|\,\mathcal{F}^i_t\right)$ and $Var\left(v|\,\mathcal{F}^i_t\right)$ are determined from (15).

A.3 Proof of Lemma 4

Based on the optimized expected utility from the static problem, I make the guess that the value function has the form

$$J\left(W^i_t,\boldsymbol{\psi}_t;t\right) = -e^{-\alpha W^i_t - \frac{1}{2}\boldsymbol{\psi}'_t H_t \boldsymbol{\psi}_t}. \tag{32}$$

Using equations (9) and (10), the expected value of the value function becomes

$$\begin{aligned}
& \mathbb{E}\left[J\left(W_t^i + X_t^i \Delta P_{t+1}, \boldsymbol{\psi}_{t+1}; t+1\right) \middle| \mathcal{F}_t^i \right] \\
= \quad & \mathbb{E}\left(-e^{-\alpha W_t^i - \alpha X_t^i \Delta P_{t+1} - \frac{1}{2}\boldsymbol{\psi}_{t+1}' H_{t+1} \boldsymbol{\psi}_{t+1}} \middle| \mathcal{F}_t^i \right) \\
= \quad & \mathbb{E}\left(-e^{-\alpha W_t^i - \alpha X_t^i \left(C_{t+1}\boldsymbol{\psi}_t + D_{t+1}\boldsymbol{\varepsilon}_{\Delta,t+1}\right) - \frac{1}{2}\left(F_{t+1}\boldsymbol{\psi}_t + G_{t+1}\boldsymbol{\varepsilon}_{\Delta,t+1}\right)' H_{t+1}\left(F_{t+1}\boldsymbol{\psi}_t + G_{t+1}\boldsymbol{\varepsilon}_{\Delta,t+1}\right)} \middle| \mathcal{F}_t^i \right) \\
= \quad & \mathbb{E}\left(e^{-\alpha X_t^i C_{t+1}\boldsymbol{\psi}_t - \alpha X_t^i D_{t+1}\boldsymbol{\varepsilon}_{\Delta,t+1} - \boldsymbol{\psi}_t' F_{t+1}' H_{t+1} G_{t+1}\boldsymbol{\varepsilon}_{\Delta,t+1} - \frac{1}{2}\boldsymbol{\varepsilon}_{\Delta,t+1}' G_{t+1}' H_{t+1} G_{t+1}\boldsymbol{\varepsilon}_{\Delta,t+1}} \middle| \mathcal{F}_t^i \right) \\
& \times \left(-e^{-\alpha W_t^i - \frac{1}{2}\boldsymbol{\psi}_t' F_{t+1}' H_{t+1} F_{t+1}\boldsymbol{\psi}_t} \right).
\end{aligned}$$

To compute the expectation above I will use the Lemma 10 with the following substitutions: $\varepsilon = \boldsymbol{\varepsilon}_{\Delta,t+1}$, $\Sigma = \Sigma_{\Delta,t+1}$, $a = -\alpha X_t^i C_{t+1}\boldsymbol{\psi}_t$, $b = -\alpha D_{t+1}' X_t^i - G_{t+1}' H_{t+1}' F_{t+1}\boldsymbol{\psi}_t$ and $c = -\frac{1}{2} G_{t+1}' H_{t+1} G_{t+1}$.

Before applying Lemma 10, we need to verify that c is symmetric and $\Sigma^{-1} - 2c$ is positive definite. As we will determine below, H_{t+1} is symmetric, which implies that c is symmetric as well. Since Σ is by definition positive definite, a sufficient condition for $\Sigma^{-1} - 2c$ to be positive definite is that c is semi-definite negative. Given the form of the value function (32), the submatrix of H_{t+1} obtained by removing the first row and column has to be positive semi-definite, because the subvector of state variables $\boldsymbol{\psi}_t$ obtained by removing the first row can assume any value in $\mathbb{R}^{n+1}$. Otherwise, the expected utility from the optimal trading strategy would be smaller than the expected utility of not trading the risky asset, which is a contradiction. Since $G_{t+1}' \Sigma_{\Delta,t+1} G_{t+1}$ is the covariance matrix of $\boldsymbol{\psi}_{t+1} | \mathcal{F}_t^i$, which is positive semi-definite (positive definite if we exclude the first row and column associated to the constant), it then follows that $G_{t+1}' H_{t+1} G_{t+1}$ is also positive semi-definite, which implies that $\Sigma^{-1} - 2c$ is positive definite.

Applying Lemma 10 we then obtain

$$\begin{aligned}
& \mathbb{E}\left[J\left(W_t^i + X_t^i \Delta P_{t+1}, \boldsymbol{\psi}_{t+1}; t+1\right) \middle| \mathcal{F}_t^i \right] \\
= \quad & -\left(\left|\Sigma_{\Delta,t+1}\right| \left|\Xi_{t+1}^{-1}\right| \right)^{-\frac{1}{2}} e^{-\alpha W_t^i - \frac{1}{2}\boldsymbol{\psi}_t' F_{t+1}' H_{t+1} F_{t+1}\boldsymbol{\psi}_t} \\
& \times e^{-\alpha X_t^i C_{t+1}\boldsymbol{\psi}_t + \frac{1}{2}\left(\alpha D_{t+1}' X_t^i + G_{t+1}' H_{t+1}' F_{t+1}\boldsymbol{\psi}_t\right)' \Xi_{t+1}\left(\alpha D_{t+1}' X_t^i + G_{t+1}' H_{t+1}' F_{t+1}\boldsymbol{\psi}_t\right)}
\end{aligned}$$

with

$$\underset{(2n+2\times 2n+2)}{\Xi_{t+1}} \equiv \left(\Sigma_{\Delta,t+1}^{-1} + G_{t+1}' H_{t+1} G_{t+1} \right)^{-1}. \tag{33}$$

The first order condition for the optimality of X_t^i is

$$-\alpha C_{t+1}\boldsymbol{\psi}_t + \alpha^2 D_{t+1}\Xi_{t+1}D_{t+1}' X_t^i + \alpha D_{t+1}\Xi_{t+1}G_{t+1}' H_{t+1}' F_{t+1}\boldsymbol{\psi}_t = 0 \Leftrightarrow X_t^i = \frac{1}{\alpha} Q_t \boldsymbol{\psi}_t$$

with

$$\underset{(1\times 2n+2)}{Q_t} \equiv \left(D_{t+1}\Xi_{t+1}D_{t+1}' \right)^{-1} \left(C_{t+1} - D_{t+1}\Xi_{t+1}G_{t+1}' H_{t+1}' F_{t+1} \right). \tag{34}$$

Substituting the optimal X_t^i into the Bellman equation, we have

$$\begin{aligned}
-e^{-\alpha W_t^i - \frac{1}{2}\boldsymbol{\psi}_t' H_t \boldsymbol{\psi}_t} \quad = \quad & -\left(\left|\Sigma_{\Delta,t+1}\right| \left|\Xi_{t+1}^{-1}\right| \right)^{-\frac{1}{2}} e^{-\alpha W_t^i - \boldsymbol{\psi}_t' Q_t' C_{t+1}\boldsymbol{\psi}_t - \frac{1}{2}\boldsymbol{\psi}_t' F_{t+1}' H_{t+1} F_{t+1}\boldsymbol{\psi}_t} \\
& \times e^{\frac{1}{2}\boldsymbol{\psi}_t' \left(D_{t+1}' Q_t + G_{t+1}' H_{t+1}' F_{t+1}\right)' \Xi_{t+1}\left(D_{t+1}' Q_t + G_{t+1}' H_{t+1}' F_{t+1}\right)\boldsymbol{\psi}_t}.
\end{aligned}$$

Since the vector of state variables $\boldsymbol{\psi}_t$ includes a constant, we can write

$$\left(|\Sigma_{\Delta,t+1}|\left|\Xi_{t+1}^{-1}\right|\right)^{-\frac{1}{2}} = e^{-\frac{1}{2}\boldsymbol{\psi}_t' M_{t+1}\boldsymbol{\psi}},$$

$$\underset{(2n+2\times 2n+2)}{M_{t+1}} \equiv \begin{bmatrix} \ln\left|\frac{\Sigma_{\Delta,t+1}}{\Xi_{t+1}}\right| & \mathbf{0}_{(1\times 2n+1)} \\ \mathbf{0}_{(2n+1,1)} & \mathbf{0}_{(2n+1,2n+1)} \end{bmatrix}, \tag{35}$$

which proves that the initial guess for the value function is correct and allows us to determine H_t. However, H_t is not uniquely defined, since $\boldsymbol{\psi}_t' Q_t' C_{t+1}\boldsymbol{\psi}_t = \boldsymbol{\psi}_t' C_{t+1}' Q_t \boldsymbol{\psi}_t$ while $Q_t' C_{t+1} \neq C_{t+1}' Q_t$. I will consider the symmetric version of H_t, which is obtained by substituting $\boldsymbol{\psi}_t' Q_t' C_{t+1}\boldsymbol{\psi}_t$ for $\frac{1}{2}\boldsymbol{\psi}_t' Q_t' C_{t+1}\boldsymbol{\psi}_t + \frac{1}{2}\boldsymbol{\psi}_t' C_{t+1}' Q_t \boldsymbol{\psi}_t$. Using (34), the symmetric version of H_t is, after simplification, defined as

$$\begin{aligned} \underset{(2n+2\times 2n+2)}{H_t} \quad = \quad & Q_t' D_{t+1}\Xi_{t+1} D_{t+1}' Q_t - F_{t+1}' H_{t+1} G_{t+1}\Xi_{t+1} G_{t+1}' H_{t+1}' F_{t+1} \\ & + F_{t+1}' H_{t+1} F_{t+1} + M_{t+1}. \end{aligned} \tag{36}$$

H_t can be found by backward induction, starting at $H_T = \mathbf{0}_{(2n+2,2n+2)}$, which is given by the boundary condition on the value function.

A.4 Proof of Lemma 5

Let the $(2n+2)$-dimensional row vector Q_t be partitioned as follows,

$$Q_t \equiv \begin{bmatrix} \underset{(1\times 1)}{q_{1,t}} & \underset{(1\times n)}{q_{\boldsymbol{s},t}^i} & \underset{(1\times n)}{q_{\boldsymbol{s},t}^c} & \underset{(1\times 1)}{q_{\theta,t}^i} \end{bmatrix}.$$

From the definitions of Q_t, C_t, G_t and F_t, one can easily check that $q_{1,t} = 0$, $t = 1, 2, ..., T-1$. Using equation (12) and the definition of $\boldsymbol{\psi}_t$ from Lemma 3, the market clearing condition can be written as

$$\begin{aligned} & \frac{1}{\alpha}\int_i \left[q_{\boldsymbol{s},t}^i \mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^i\right) + q_{\boldsymbol{s},t}^c \mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^c\right) + q_{\theta,t}^i \mathbb{E}\left(\theta_t\middle|\mathcal{F}_t^i\right)\right] = \theta_t \\ \Leftrightarrow \quad & q_{\boldsymbol{s},t}^i \int_i \mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^i\right) + q_{\boldsymbol{s},t}^c \mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^c\right) + q_{\theta,t}^i \int_i \mathbb{E}\left(\theta_t\middle|\mathcal{F}_t^i\right) - \alpha\theta_t = 0 \end{aligned} \tag{37}$$

where the equivalence follows from $\mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^c\right)$ being the same for every investor and from the unit mass of investors. Rearranging equation (24), which uses the fact that ξ_t is adapted to $\mathcal{F}_t^i$, we have

$$\mathbb{E}\left(\theta_t\middle|\mathcal{F}_t^i\right) = \theta_t + \frac{\boldsymbol{p}_t}{p_{\theta,t}}\left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^i\right)\right]. \tag{38}$$

Substituting equation (38) into (37), we get

$$\left(q_{\boldsymbol{s},t}^i - \frac{q_{\theta,t}^i}{p_{\theta,t}}\boldsymbol{p}_t\right)\int_i \mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^i\right) + q_{\boldsymbol{s},t}^c \mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^c\right) + \frac{q_{\theta,t}^i}{p_{\theta,t}}\boldsymbol{p}_t\boldsymbol{s} + \left(q_{\theta,t}^i - \alpha\right)\theta_t = 0 \tag{39}$$

Noticing that $\mathcal{F}_t^i = \left\{\mathcal{F}_t^c, \tilde{\boldsymbol{s}}_\tau^i : \tau = 1, ..., t\right\}$, where $\tilde{\boldsymbol{s}}_t^i$ is a Gaussian vector, we can use Lemma 11 on conditional normal distributions to compute $\int_i \mathbb{E}\left(\boldsymbol{s}\middle|\mathcal{F}_t^i\right)$. Let $\tilde{\boldsymbol{S}}_t^i$ denote the stack of $\tilde{\boldsymbol{s}}_t^i$ from date 1 to t, i.e.

$\underset{(tn\times 1)}{\tilde{\boldsymbol{S}}_t^i} \equiv \begin{bmatrix} \tilde{\boldsymbol{s}}_1'^i & \cdots & \tilde{\boldsymbol{s}}_t'^i \end{bmatrix}'$, and $\boldsymbol{\Sigma}_{\tilde{s},t}$ be given by

$$\underset{(tn\times tn)}{\boldsymbol{\Sigma}_{\tilde{s},t}} = \begin{bmatrix} \Sigma_{\tilde{s},1} & \mathbf{0}_{(n\times n)} & \cdots & \mathbf{0}_{(n\times n)} \\ \mathbf{0}_{(n\times n)} & \Sigma_{\tilde{s},2} & \cdots & \mathbf{0}_{(n\times n)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{(n\times n)} & \mathbf{0}_{(n\times n)} & \cdots & \Sigma_{\tilde{s},t} \end{bmatrix}.$$

Then we have that

$$\begin{bmatrix} \boldsymbol{s} \\ \tilde{\boldsymbol{S}}_t^i \end{bmatrix} \Bigg| \mathcal{F}_t^c \sim N\left(\begin{bmatrix} \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) \\ \mathbf{1}_{(t\times 1)} \otimes \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) \end{bmatrix}, \begin{bmatrix} Var\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) & \mathbf{1}_{(1\times t)} \otimes Var\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) \\ \mathbf{1}_{(t\times 1)} \otimes Var\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) & \mathbf{1}_{(t\times t)} \otimes Var\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) + \boldsymbol{\Sigma}_{s,t}^i \end{bmatrix}\right),$$

where $\otimes$ denotes the Kronecker product of matrices. Applying Lemma 11, we obtain

$$\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^i\right) = \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) + \Gamma_t\left[\tilde{\boldsymbol{S}}_t^i - \mathbf{1}_{(t\times 1)} \otimes \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right)\right]$$

where

$$\underset{(n\times tn)}{\Gamma_t} \equiv \left[\mathbf{1}_{(1\times t)} \otimes Var\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right)\right]\left[\mathbf{1}_{(t\times t)} \otimes Var\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) + \boldsymbol{\Sigma}_{\tilde{s},t}\right]^{-1}. \tag{40}$$

And integrating over i, we obtain

$$\begin{aligned} \int_i \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^i\right) &= \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) + \Gamma_t\left\{\mathbf{1}_{(t\times 1)} \otimes \left[\boldsymbol{s} - \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right)\right]\right\} \\ &= \hat{\Gamma}_t \boldsymbol{s} + \left(I_n - \hat{\Gamma}_t\right)\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right). \end{aligned} \tag{41}$$

where

$$\underset{(n\times n)}{\hat{\Gamma}_t} \equiv \Gamma_t\left(\mathbf{1}_{(t\times 1)} \otimes I_n\right). \tag{42}$$

Substituting equation (41) into (39) the market clearing condition becomes

$$\left[\left(q_{\boldsymbol{s},t}^i - \frac{q_{\theta,t}^i}{p_{\theta,t}}\boldsymbol{p}_t\right)\hat{\Gamma}_t + \frac{q_{\theta,t}^i}{p_{\theta,t}}\boldsymbol{p}_t\right]\boldsymbol{s} + \left[q_{\boldsymbol{s},t}^c + \left(q_{\boldsymbol{s},t}^i - \frac{q_{\theta,t}^i}{p_{\theta,t}}\boldsymbol{p}_t\right)\left(I_n - \hat{\Gamma}_t\right)\right]\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right) + \left(q_{\theta,t}^i - \alpha\right)\theta_t = 0.$$

For the market condition to hold for all possible realizations of $\boldsymbol{s}$, $\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^c\right)$ and θ_t it must be the case that their coefficients are all equal to zero. This condition is satisfied when

$$q_{\theta,t}^i = \alpha \tag{43}$$

$$q_{\boldsymbol{s},t}^i = -q_{\boldsymbol{s},t}^c = \frac{\alpha}{p_{\theta,t}}\boldsymbol{p}_t\left(I_n - \hat{\Gamma}_t^{-1}\right) \tag{44}$$

which concludes the proof of the first part of the lemma.

To prove the second part of the lemma, I start by determining the date $T-1$ price coefficients. Plugging expressions (30), (31) and $H_T = \mathbf{0}_{(2n+2,2n+2)}$ into (34), we obtain

$$q_{\boldsymbol{s},T-1}^i = \frac{\hat{K} - \boldsymbol{p}_{T-1}}{Var\left(v|\,\mathcal{F}_{T-1}^i\right)},\ q_{\boldsymbol{s},T-1}^c = -\frac{\hat{\boldsymbol{p}}_{T-1}}{Var\left(v|\,\mathcal{F}_{T-1}^i\right)},\ q_{\theta,T-1}^i = -\frac{p_{\theta,T-1}}{Var\left(v|\,\mathcal{F}_{T-1}^i\right)},$$

noting that $Q_{T-1} = Var\left(v|\,\mathcal{F}^i_{T-1}\right)^{-1} C_T$. From the market clearing conditions (43) and (44), we then obtain that

$$\begin{aligned}
\boldsymbol{p}_{T-1} &= \hat{K}\hat{\Gamma}_{T-1} && (45)\\
\hat{\boldsymbol{p}}_{T-1} &= \hat{K}\left(I_n - \hat{\Gamma}_{T-1}\right) &&\\
p_{\theta,T-1} &= -\alpha Var\left(v|\,\mathcal{F}^i_{T-1}\right) && (46)
\end{aligned}$$

and so $\boldsymbol{p}_{T-1} + \hat{\boldsymbol{p}}_{T-1} = \hat{K}$. This implies that the $(j+1)$-th column of C_T and its $(j+n+1)$-th column are symmetric, for $j = 1, ..., n$. Knowing that $Q_{T-1} = Var\left(v|\,\mathcal{F}^i_{T-1}\right)^{-1} C_T$ and recalling that $H_T = \mathbf{0}_{(2n+2,2n+2)}$, we can then determine H_{T-1} from (36) as

$$H_{T-1} = \frac{1}{Var\left(v|\,\mathcal{F}^i_{T-1}\right)} C_T' C_T,$$

and it follows that the $(j+1)$-th column and row of H_{T-1} and its $(j+n+1)$-th column and row are symmetric, $j = 1, ..., n$. We can then write

$$H_{T-1} = J'\tilde{H}_{T-1}J,\ J' = \begin{bmatrix} 1 & \mathbf{0}_{(1,n)} & 0 \\ \mathbf{0}_{(n,1)} & I_n & \mathbf{0}_{(n,1)} \\ \mathbf{0}_{(n,1)} & -I_n & \mathbf{0}_{(n,1)} \\ 0 & \mathbf{0}_{(1,n)} & 1 \end{bmatrix} \tag{47}$$

with $\tilde{H}_{T-1}$ being composed by the first $n+1$ and the last columns and rows of H_{T-1}. Now, substituting the expression above into expression (34) for Q_{T-2}, and inverting for C_{T-1}, we obtain

$$\begin{aligned}
C_{T-1} &= D_{T-1}\left(\Sigma^{-1}_{\Delta,T-1} + G'_{T-1}H_{T-1}G_{T-1}\right)^{-1} D'_{T-1}Q_{T-2} \\
&\quad + D_{T-1}\left(\Sigma^{-1}_{\Delta,T-1} + G'_{T-1}H_{T-1}G_{T-1}\right)^{-1} G'_{T-1}J'\tilde{H}_{T-1}JF_{T-1}.
\end{aligned}$$

The market clearing condition implies that $(j+1)$-th column of Q_{T-2} and its $(j+n+1)$-th column are symmetric, $j = 1, ..., n$. We can verify that the same holds for the second term due to JF_{T-1}, since

$$JF_{T-1} = \\
\begin{bmatrix} 1 & \mathbf{0}_{(1,n)} & \mathbf{0}_{(1,n)} & 0 \\ \mathbf{0}_{(n,1)} & I_n - k^c_{\boldsymbol{s},\xi,T-1}\left(\boldsymbol{p}_{T-1} - \rho\frac{p_{\theta,T-1}}{p_{\theta,T-2}}\boldsymbol{p}_{T-2}\right) & -I_n + k^c_{\boldsymbol{s},\xi,T-1}\left(\boldsymbol{p}_{T-1} - \rho\frac{p_{\theta,T-1}}{p_{\theta,T-2}}\boldsymbol{p}_{T-2}\right) & \mathbf{0}_{(n,1)} \\ 0 & \mathbf{0}_{(1,n)} & \mathbf{0}_{(1,n)} & \rho \end{bmatrix},$$

and so the C_{T-1} also shares this property. Therefore, from expression (25) for C_{T-1} we obtain

$$\boldsymbol{p}_{T-1} - \boldsymbol{p}_{T-2} + \hat{\boldsymbol{p}}_{T-1} - \hat{\boldsymbol{p}}_{T-2} = 0 \Leftrightarrow \boldsymbol{p}_{T-2} + \hat{\boldsymbol{p}}_{T-2} = \hat{K}.$$

Next, substituting $H_{T-1} = J'\tilde{H}_{T-1}J$ into the expression for H_{T-2} we obtain

$$\begin{aligned}
H_{T-2} &= Q'_{T-2}D_{T-1}\Xi_{T-1}D'_{T-1}Q_{T-2} - F'_{T-1}J'\tilde{H}_{T-1}JG_{T-1}\Xi_{T-1}G'_{T-1}J'\tilde{H}_{T-1}JF_{T-1} \\
&\quad + F'_{T-1}J'\tilde{H}_{T-1}JF_{T-1} + M_{T-1}
\end{aligned}$$

and we can see that all terms share the property that their $(j+1)$-th column and row is the symmetric of their $(j+n+1)$-th column and row, $j=1,...,n$. Therefore, we can write $H_{T-2}=J'\tilde{H}_{T-2}J$. Using the same steps as above for every date until $t=1$ we can then show that $\boldsymbol{p_t}+\boldsymbol{\hat{p}_t}=\hat{K}$, $t=1,...,T-1$, which proves the second part of the lemma.

A.5 Proof of Theorems 7 and 8

Parts (i) to (iv) of Theorem 7 are proved in Appendix C.4 and part (v) is immediate. Parts (i) and (ii) of Theorem 8 are proved in Appendix C.5 and the proof of parts (iii) to (v) is similar to those of Theorem 7 and so are omitted.

B Numerical Algorithm to Solve the Equilibrium

Let $\underline{\boldsymbol{p}}$ denote the $(T-1)\times(2n+1)$ vector of parameters of the price function. I solve for the $\underline{\boldsymbol{p}}$ that solves the fixed point problem associated to the $T-1$ market clearing conditions. For that, I recast the set of market clearing conditions as a function $\underline{\boldsymbol{p}}=f\left(\underline{\boldsymbol{p}}\right)$.

We can rewrite C_{t+1} as

$$C_{t+1}\equiv\hat{C}_{t+1}-\left[\begin{array}{cccc}0 & \boldsymbol{p}_t & \boldsymbol{\hat{p}}_t & p_{\theta,t}\end{array}\right].$$

As a result, the market clearing condition at date t can be rewritten as

$$\begin{aligned}\left[\begin{array}{cccc}0 & \boldsymbol{p}_t & \boldsymbol{\hat{p}}_t & p_{\theta,t}\end{array}\right] &= \hat{C}_{t+1}-D_{t+1}\Xi_{t+1}G'_{t+1}H'_{t+1}F_{t+1}\\ &\quad -D_{t+1}\Xi_{t+1}D'_{t+1}\left[\begin{array}{ccc}\alpha & \frac{\alpha}{p_{\theta,t}}\boldsymbol{p}_t\left(I_n-\hat{\Gamma}_t^{-1}\right) & -\frac{\alpha}{p_{\theta,t}}\boldsymbol{p}_t\left(I_n-\hat{\Gamma}_t^{-1}\right)\end{array}\right].\end{aligned}$$

The $T-1$ market clearing conditions then give us the $\underline{\boldsymbol{p}}=f\left(\underline{\boldsymbol{p}}\right)$ fixed point problem for $\underline{\boldsymbol{p}}$. The algorithm I use to solve the fixed point problem is the following:

1. Make and initial guess $\underline{\boldsymbol{p}}^0$, and choose the damping factor $\delta>0$;
2. Obtain the j-th estimate of $\underline{\boldsymbol{p}}$ as $\underline{\boldsymbol{p}}^j=\underline{\boldsymbol{p}}^{j-1}(1-\delta)+\delta f\left(\underline{\boldsymbol{p}}^{j-1}\right)$ and compute the norm $\left\|\underline{\boldsymbol{p}}^{j-1}-f\left(\underline{\boldsymbol{p}}^{j-1}\right)\right\|$;
3. Whenever $\left\|\underline{\boldsymbol{p}}^{j-1}-f\left(\underline{\boldsymbol{p}}^{j-1}\right)\right\|\geq\left\|\underline{\boldsymbol{p}}^{j-2}-f\left(\underline{\boldsymbol{p}}^{j-2}\right)\right\|$ decrease the damping factor and recompute the i-th estimate with the new damping factor, until $\left\|\underline{\boldsymbol{p}}^{j-1}-f\left(\underline{\boldsymbol{p}}^{j-1}\right)\right\|<\left\|\underline{\boldsymbol{p}}^{j-2}-f\left(\underline{\boldsymbol{p}}^{j-2}\right)\right\|$. If the damping factor becomes smaller than a threshold value, abort the algorithm with no solution found;
4. Repeat steps 2 and 3 until $\left\|\underline{\boldsymbol{p}}^{j-1}-f\left(\underline{\boldsymbol{p}}^{j-1}\right)\right\|$ is smaller than some threshold value, and take $\underline{\boldsymbol{p}}^j$ as the solution.

C 3-period example

Here I will solve in closed form a 3-period version of the model without residual uncertainty. This version of the model can be solved by direct application of Theorem 1 and Lemmas 2 to 5. However, it is useful to introduce some modifications to the general model. The objective of these modifications is to simplify the procedure for the specific setting under analysis.

The assumption of no residual uncertainty implies that $s=v$. This originates redundancies that can be eliminated. Specifically, I drop the conditional beliefs on v (Lemma 2) and the term $v-\mathbb{E}\left(v|\,\mathcal{F}_t^i\right)$ in $\varepsilon_{\Delta,t+1}$ (Lemma 3). Moreover, since $s=v$, it is straightforward to see that $\hat{K}=1$. Using equation (14), we then have that $\hat{p}_t = 1 - p_t$. To simplify the notation, I use this expression to substitute for $\hat{p}_t$ everywhere. I also redefine $\xi_t \equiv s + \frac{p_{\theta,t}}{p_t}\theta_t$. The price function (5) then becomes

$$P_t = (1 - p_t)\,\mathbb{E}\left(s|\,\mathcal{F}_t^c\right) + p_t\xi_t.$$

Finally, using (14) on expressions (25) and (30), it is immediate that ΔP_{t+1} depends on $\mathbb{E}\left(s|\,\mathcal{F}_t^i\right)$ and $\mathbb{E}\left(s|\,\mathcal{F}_t^c\right)$ only through $\mathbb{E}\left(s|\,\mathcal{F}_t^i\right) - \mathbb{E}\left(s|\,\mathcal{F}_t^c\right)$. This allows me to redefine

$$\psi_t \equiv \begin{bmatrix} 1 & \mathbb{E}\left(s|\,\mathcal{F}_t^i\right) - \mathbb{E}\left(s|\,\mathcal{F}_t^c\right) & \mathbb{E}\left(\theta_t|\,\mathcal{F}_t^i\right) \end{bmatrix}',$$

decreasing the dimensionality of C_{t+1}, F_{t+1}, G_{t+1}, Q_{t+1} and H_{t+1}.

As stated in Corollary 6, in the general case the price function coefficients for all dates are determined simultaneously by solving a system of nonlinear equations. If the conditional beliefs are given, one can solve the price function coefficients backwards from $t = T - 1$. However, conditional beliefs are solved forward from $t = 0$ and require information about those coefficients. Specifically time t conditional beliefs require the knowledge of the ratio $\left\{\frac{p_{\theta,\tau}}{p_\tau}\right\}_{\tau\leq t}$. Without residual uncertainty, however, we can determine the value of $\frac{p_{\theta,t}}{p_t}\forall t$ beforehand. Adapting from Corollary 2 of He and Wang (1995) to our notation, $T = 3$ and defining $\sigma^2_{\tilde{s},2} \equiv \frac{\beta}{1-\beta}\sigma^2_{\tilde{s},1}$, $0 \leq \beta \leq 1$, we have

$$\frac{p_{\theta,1}}{p_1} = -\alpha\sigma^2_{\tilde{s},1},\ \frac{p_{\theta,2}}{p_2} = -\alpha\beta\sigma^2_{\tilde{s},1}. \tag{48}$$

Thus, we can start by solving for the beliefs forward from $t = 0$, and then solve for the price function coefficients backwards from $t = 2$ separately.

C.1 Conditional Beliefs

Beliefs on s and θ_t conditional on $\mathcal{F}_t^i$ are obtained by applying Lemma 9 to

$$x_t = \begin{bmatrix} s \\ \theta_t \end{bmatrix},\ y_t = \begin{bmatrix} \xi_t \\ \tilde{s}_t^i \end{bmatrix},\ A_t = \begin{bmatrix} 1 & 0 \\ 0 & \rho \end{bmatrix},\ \Sigma_{x,t} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2_\theta \end{bmatrix},\ B_t = \begin{bmatrix} 1 & \frac{p_{\theta,t}}{p_t} \\ 1 & 0 \end{bmatrix},\ \Sigma_{y,t} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2_{\tilde{s},t} \end{bmatrix},$$

with prior beliefs given by $\mathbb{E}\left(x_0|\,\mathcal{F}_0\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $Var\left(x_0|\,\mathcal{F}_0\right) = \begin{bmatrix} \sigma^2_v & 0 \\ 0 & 0 \end{bmatrix}$. Applying Lemma 9 and using (48), it is straightforward to obtain

$$\mathbb{E}\left(s|\,\mathcal{F}_1^i\right) = \frac{\frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\xi_1 + \frac{1}{\sigma^2_{\tilde{s},1}}\tilde{s}_1^i}{\frac{1}{\sigma^2_v} + \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{1}{\sigma^2_{\tilde{s},1}}},\ \mathbb{E}\left(s|\,\mathcal{F}_2^i\right) = \frac{\frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\xi_1 + \frac{1}{\sigma^2_{\tilde{s},1}}\tilde{s}_1^i + \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\left(\xi_2 - \beta\rho\xi_1\right) + \frac{1}{\frac{\beta}{1-\beta}\sigma^2_{\tilde{s},1}}\tilde{s}_2^i}{\frac{1}{\sigma^2_v} + \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{1}{\sigma^2_{\tilde{s},1}} + \frac{(1-\beta\rho)^2}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{1}{\frac{\beta}{1-\beta}\sigma^2_{\tilde{s},1}}}, \tag{49}$$

$$\mathbb{E}\left(\theta_1|\,\mathcal{F}_1^i\right) = \frac{\sigma^2_v\tilde{s}_1^i - \left(\sigma^2_{\tilde{s},1} + \sigma^2_v\right)\xi_1}{\alpha\sigma^4_{\tilde{s},1}\sigma^2_v}V_1^i, \tag{50}$$

$$V_1^i \equiv Var\left(s|\,\mathcal{F}_1^i\right) = \frac{1}{\frac{1}{\sigma^2_v} + \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{1}{\sigma^2_{\tilde{s},1}}},\ V_2^i \equiv Var\left(s|\,\mathcal{F}_2^i\right) = \frac{1}{\frac{1}{\sigma^2_v} + \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{1}{\sigma^2_{\tilde{s},1}} + \frac{(1-\beta\rho)^2}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{1}{\frac{\beta}{1-\beta}\sigma^2_{\tilde{s},1}}}, \tag{51}$$

$$K_1 = \begin{bmatrix} \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} & \frac{1}{\sigma^2_{\tilde{s},1}} \\ -\frac{1}{\alpha\sigma^2_{\tilde{s},1}}\left(\frac{1}{V^i_1} - \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\right) & \frac{1}{\alpha\sigma^4_{\tilde{s},1}} \end{bmatrix} V^i_1,\ K_2 = \begin{bmatrix} \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} & \frac{1}{\frac{\beta}{1-\beta}\sigma^2_{\tilde{s},1}} \\ -\frac{1}{\alpha\beta\sigma^2_{\tilde{s},1}}\left(\frac{1}{V^i_2} - \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\right) & \frac{1}{\alpha\frac{\beta^2}{1-\beta}\sigma^4_{\tilde{s},1}} \end{bmatrix} V^i_2.$$

Beliefs conditional on $\mathcal{F}^c_t$ are obtained in the same way, by removing the second row of y_t and B_t and the second row and column of $\Sigma_{y,t}$, yielding

$$\mathbb{E}\left(s|\,\mathcal{F}^c_1\right) = \frac{\frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\xi_1}{\frac{1}{\sigma^2_v} + \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}},\ \mathbb{E}\left(s|\,\mathcal{F}^c_2\right) = \frac{\frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\xi_1 + \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\left(\xi_2 - \beta\rho\xi_1\right)}{\frac{1}{\sigma^2_v} + \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{(1-\beta\rho)^2}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}}, \tag{52}$$

$$V^c_1 \equiv Var\left(s|\,\mathcal{F}^c_1\right) = \frac{1}{\frac{1}{\sigma^2_v} + \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}},\ V^c_2 \equiv Var\left(s|\,\mathcal{F}^c_2\right) = \frac{1}{\frac{1}{\sigma^2_v} + \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{(1-\beta\rho)^2}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}},$$

$$K^c_1 = \begin{bmatrix} \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} \\ -\frac{1}{\alpha\sigma^2_{\tilde{s},1}}\left(\frac{1}{V^c_1} - \frac{1}{\alpha^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\right) \end{bmatrix} V^c_1,\ K^c_2 = \begin{bmatrix} \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} \\ -\frac{1}{\alpha\beta\sigma^2_{\tilde{s},1}}\left(\frac{1}{V^c_2} - \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta}\right) \end{bmatrix} V^c_2.$$

The following relations between V^i_1, V^i_2, V^c_1 and V^c_2 will be used to simplify expressions later on

$$\frac{1}{V^i_2} = \frac{1}{V^i_1} + \frac{(1-\beta\rho)^2}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} + \frac{1}{\frac{\beta}{1-\beta}\sigma^2_{\tilde{s},1}},\ \frac{1}{V^c_2} = \frac{1}{V^c_1} + \frac{(1-\beta\rho)^2}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta},\ \frac{1}{V^i_2} = \frac{1}{V^c_2} + \frac{1}{\beta\sigma^2_{\tilde{s},1}},\ \frac{1}{V^i_1} = \frac{1}{V^c_1} + \frac{1}{\sigma^2_{\tilde{s},1}} \tag{53}$$

C.2 Solving for p_2 and $p_{\theta,2}$

From equations (40) and (42), and using $\frac{1}{V^i_2} = \frac{1}{V^c_2} + \frac{1}{\beta\sigma^2_{\tilde{s},1}}$, we obtain $\hat{\Gamma}_2 = \frac{V^i_2}{\beta\sigma^2_{\tilde{s},1}}$. Thus, from equations (45) and (46), we can determine that

$$p_2 = \frac{V^i_2}{\beta\sigma^2_{\tilde{s},1}},\ p_{\theta,2} = -\alpha V^i_2.$$

Note that $\frac{p_{\theta,2}}{p_2} = -\alpha\beta\sigma^2_{\tilde{s},1}$ as it was initially assumed.

C.3 Solving for p_1 and $p_{\theta,1}$

Using equations (40), (42) and $\frac{1}{V^i_1} = \frac{1}{V^c_1} + \frac{1}{\sigma^2_{\tilde{s},1}}$, we have that $\hat{\Gamma}_1 = \frac{V^i_1}{\sigma^2_{\tilde{s},1}}$. Using (48), $\frac{1}{V^i_1} = \frac{1}{V^c_1} + \frac{1}{\sigma^2_{\tilde{s},1}}$ and $\hat{\Gamma}_1$ we can then write the market clearing condition (13) as

$$Q_1 = \begin{bmatrix} 0 & \frac{1}{V^c_1} & \alpha \end{bmatrix}. \tag{54}$$

Notice that, from (25), C_2 can be written as $C_2 = \hat{C}_2 - \begin{bmatrix} 0 & p_1 & p_{\theta,1} \end{bmatrix}$, where $\hat{C}_2$ depends on date 1 price function parameters only through the ratio $\frac{p_{\theta,1}}{p_1}$, whose value is known. Using equation (34) for Q_1 and the market clearing condition, p_1 and $p_{\theta,1}$ are given by

$$\begin{bmatrix} p_1 & p_{\theta,1} \end{bmatrix} = -D_2\Xi_2 D'_2 \begin{bmatrix} \frac{1}{V^c_1} & \alpha \end{bmatrix} + \left(\hat{C}_2 - D_2\Xi_2 G'_2 H_2 F_2\right)\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{55}$$

C_2 is given by (25), with some changes to reflect the modifications introduced above: the modification of ψ_t implies that third column disappears; and the redefinition of ξ_t (which is now the old value of ξ_t divided by p_t) implies that $k^c_{\boldsymbol{s},\xi,t+1}$ is divided by p_t. Using (53) to simplify, we obtain C_2 and $\hat{C}_2$ as

$$C_2 = \hat{C}_2 - \begin{bmatrix} 0 & p_1 & p_{\theta,1} \end{bmatrix}, \hat{C}_2 = \begin{bmatrix} 0 & 1 - \frac{V_2^i}{V_1^c} & -\alpha V_2^i \rho \end{bmatrix}. \tag{56}$$

F_2 is given by (20), with the following changes: the second row becomes the old second row minus the old third row, and the third row and column disappear (due to the change in ψ_t); $k^c_{\boldsymbol{s},\xi,t+1}$ is divided by p_t (due to the redefinition of ξ_t). Using (53) to simplify, we get

$$F_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{V_2^c}{V_1^c} & 0 \\ 0 & 0 & \rho \end{bmatrix}. \tag{57}$$

G_2 is given by (21), with the following changes: the first column disappears (to reflect the change in $\varepsilon_{\Delta,t}$); the second row becomes the difference between the old second and third rows, and the third row disappears (due to the change in ψ_t); and $k^c_{\boldsymbol{s},\xi,t+1}$ and $k^c_{\theta,\xi,t+1}$ are divided by p_t (due to the redefinition of ξ_t). After simplification, using (48) and (53), we obtain

$$G_2 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{V_2^c}{V_1^c} - \frac{V_2^i}{V_1^i} & \frac{V_2^i(1-\beta)}{\beta\sigma^2_{\tilde{s},1}} & -\frac{\left(V_2^i - V_2^c\right)(1-\beta\rho)}{\alpha\beta\sigma^2_{\tilde{s},1}\sigma^2_\theta} \\ \frac{1}{\alpha}\left(\frac{V_2^i}{V_2^c V_1^i} - \frac{1}{V_1^c} - \frac{1-\rho}{\sigma^2_{\tilde{s},1}}\right) & \frac{V_2^i(1-\beta)}{\alpha\beta^2\sigma^4_{\tilde{s},1}} & 1 - \frac{V_2^i(1-\beta\rho)}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} \end{bmatrix}.$$

By definition, $H_3 = \mathbf{0}_{(3,3)}$. As a result, $\Xi_3 = \Sigma_{\Delta,3}$, $M_3 = \mathbf{0}_{(3,3)}$, $D_3\Xi_3D_3' = V_2^i$ and we obtain, from (36), $H_2 = Q_2'Q_2V_2^i$. Using the market clearing condition (13) and simplifying we get $Q_2 = \begin{bmatrix} 0 & \frac{1}{V_2^c} & \alpha \end{bmatrix}$, and so

$$H_2 = V_2^i \begin{bmatrix} 0 \\ \frac{1}{V_2^c} \\ \alpha \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{V_2^c} \\ \alpha \end{bmatrix}'. \tag{58}$$

Tedious algebra, using (53) to simplify expressions, shows that

$$G_2'H_2G_2 = V_2^i \begin{bmatrix} -\frac{1-\rho}{\sigma^2_{\tilde{s},1}} \\ \frac{1-\beta}{\beta\sigma^2_{\tilde{s},1}} \\ \alpha \end{bmatrix} \begin{bmatrix} -\frac{1-\rho}{\sigma^2_{\tilde{s},1}} \\ \frac{1-\beta}{\beta\sigma^2_{\tilde{s},1}} \\ \alpha \end{bmatrix}', \; G_2'H_2F_2 = V_2^i \begin{bmatrix} -\frac{1-\rho}{\sigma^2_{\tilde{s},1}} \\ \frac{1-\beta}{\beta\sigma^2_{\tilde{s},1}} \\ \alpha \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{V_1^c} \\ \alpha\rho \end{bmatrix}'. \tag{59}$$

From (31), removing the first row and column, we obtain

$$\Sigma_{\Delta,2} = \begin{bmatrix} V_1^i & 0 & 0 \\ 0 & \frac{\beta}{1-\beta}\sigma^2_{\tilde{s},1} & 0 \\ 0 & 0 & \sigma^2_\theta \end{bmatrix}.$$

Because $G_2'H_2G_2$ is the outer product of a vector, we can determine $\Xi_2 \equiv \left(\Sigma^{-1}_{\Delta,2} + G_2'H_2G_2\right)^{-1}$ without inverting matrices, obtaining

$$\Xi_2 = \Sigma_{\Delta,2} - \frac{V_2^i}{1 + V_2^i\left(\frac{V_1^i(1-\rho)^2}{\sigma^4_{\tilde{s},1}} + \frac{1-\beta}{\beta\sigma^2_{\tilde{s},1}} + \alpha^2\sigma^2_\theta\right)} \begin{bmatrix} -\frac{(1-\rho)V_1^i}{\sigma^2_{\tilde{s}.1}} \\ 1 \\ \alpha\sigma^2_\theta \end{bmatrix} \begin{bmatrix} -\frac{(1-\rho)V_1^i}{\sigma^2_{\tilde{s}.1}} \\ 1 \\ \alpha\sigma^2_\theta \end{bmatrix}'. \tag{60}$$

D_2 is given by (26), with the first column removed and, once again, $k^c_{\boldsymbol{s},\xi,t+1}$ is divided by p_t. After simplification, we get

$$D_2 = V_2^i \left(\frac{1}{\beta \sigma^2_{\tilde{s},1}} + \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} \right) \left[\begin{array}{ccc} 1-\beta\rho & 0 & -\alpha\beta\sigma^2_{\tilde{s},1} \end{array} \right].$$

We can now obtain

$$\begin{aligned} D_2\Xi_2 D_2' &= -\left(V_2^i\right)^2 \left(\frac{1}{\beta\sigma^2_{\tilde{s},1}} + \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} \right)^2 \times \\ &\quad \times \left[V_1^i \left(1-\beta\rho\right)^2 + \alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta - \frac{V_2^i \left(\alpha^2\beta\sigma^2_{\tilde{s},1}\sigma^2_\theta + \frac{V_1^i(1-\rho)(1-\beta\rho)}{\sigma^2_{\tilde{s},1}} \right)^2}{1 + V_2^i \left(\frac{V_1^i(1-\rho)^2}{\sigma^4_{\tilde{s},1}} + \frac{1-\beta}{\beta\sigma^2_{\tilde{s},1}} + \alpha^2\sigma^2_\theta \right)} \right] \end{aligned} \tag{61}$$

and

$$D_2\Xi_2 G_2' H_2 F_2 = -\frac{\left(V_2^i\right)^2 \left(\frac{1}{\beta\sigma^2_{\tilde{s},1}} + \frac{1-\beta\rho}{\alpha^2\beta^2\sigma^4_{\tilde{s},1}\sigma^2_\theta} \right) \left(\alpha^2\beta\sigma^2_{\tilde{s},1}\sigma^2_\theta + \frac{V_1^i(1-\rho)(1-\beta\rho)}{\sigma^2_{\tilde{s},1}} \right)}{1 + V_2^i \left(\frac{V_1^i(1-\rho)^2}{\sigma^4_{\tilde{s},1}} + \frac{1-\beta}{\beta\sigma^2_{\tilde{s},1}} + \alpha^2\sigma^2_\theta \right)} \left[\begin{array}{ccc} 0 & \frac{1}{V_1^c} & \alpha\rho \end{array} \right].$$

At this point we have everything needed to obtain p_1 and $p_{\theta,1}$ from (55). However, the expressions for the general case are too complex and difficult to work with. For that reason, I consider two special cases. In the first, information at $t=2$ is exclusively endogenous ($\beta \to 1$, $0 \leq \rho < 1$), whereas in the second it is exclusively exogenous ($0 \leq \beta < 1$, $\rho = 1$).

C.4 Model with Endogenous Information

Taking the limits of $D_2\Xi_2 G_2' H_2 F_2$, and $D_2\Xi_2 D_2'$ as $\beta \to 1$, and substituting into (55), we obtain the date 1 price coefficients as

$$p_{\theta,1} = -\alpha V_2^i \frac{\left(1 + \frac{1-\rho}{\sigma^2_{\tilde{s},1}\alpha^2\sigma^2_\theta}\right)^2 + \frac{\rho}{\alpha^2\sigma^2_\theta}\left(\frac{1}{V_1^i} - \frac{1-\rho}{\sigma^2_{\tilde{s},1}}\right)}{1 + \frac{1}{V_1^i\alpha^2\sigma^2_\theta}},\; p_1 = \frac{V_2^i}{\sigma^2_{\tilde{s},1}} \frac{\left(1 + \frac{1-\rho}{\sigma^2_{\tilde{s},1}\alpha^2\sigma^2_\theta}\right)^2 + \frac{\rho}{\alpha^2\sigma^2_\theta}\left(\frac{1}{V_1^i} - \frac{1-\rho}{\sigma^2_{\tilde{s},1}}\right)}{1 + \frac{1}{V_1^i\alpha^2\sigma^2_\theta}}.$$

$p_1 \geq 0$ and $p_{\theta,1} \leq 0$ if $\frac{1}{V_1^i} - \frac{1-\rho}{\sigma^2_{\tilde{s},1}} \geq 0 \Leftrightarrow V_1^i(1-\rho) \leq \sigma^2_{\tilde{s},1}$ which is always the case since, by definition, $V_1^i \leq \sigma^2_{\tilde{s},1}$.

C.4.1 Average, Announcement Date Reaction and Drift in Expectations

Averaging over net supply shocks and interpretation errors, we have $\xi_1 = \xi_2 = \tilde{s}_1^i = \tilde{s}_2^i = s$. In this scenario, expectations (49), (52) and (50) become

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(s|\,\mathcal{F}_1^i\right)\right] = \left(1 - \frac{V_1^i}{\sigma^2_v}\right)s,\; \mathbb{E}_{\theta,i}\left[\mathbb{E}\left(s|\,\mathcal{F}_2^i\right)\right] = \left(1 - \frac{V_2^i}{\sigma^2_v}\right)s, \tag{62}$$

$$\mathbb{E}_{\theta}\left[\mathbb{E}\left(s|\,\mathcal{F}_1^c\right)\right] = \left(1 - \frac{V_1^c}{\sigma^2_v}\right)s,\; \mathbb{E}_{\theta}\left[\mathbb{E}\left(s|\,\mathcal{F}_2^c\right)\right] = \left(1 - \frac{V_2^c}{\sigma^2_v}\right)s, \tag{63}$$

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\theta_1|\,\mathcal{F}_1^i\right)\right] = -\frac{V_1^i}{\alpha\sigma^2_{\tilde{s},1}\sigma^2_v}s. \tag{64}$$

Since $s = v$, the conditional expectations of s and v are the same.

Notice that, on average, the announcement date reaction of the expectation of v (or s) conditional on $\mathcal{F}^i$,

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)-\mathbb{E}\left(v|\,\mathcal{F}_0\right)\right]=\left(1-\frac{V_1^i}{\sigma_v^2}\right)s \tag{65}$$

has the same sign of s as long as

$$\frac{V_1^i}{\sigma_v^2}<1\Leftrightarrow\sigma_v^2\left(\frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}+\frac{1}{\sigma_{\tilde{s},1}^2}\right)>0$$

Therefore, expectations react to the announcement if and only if: (i) the liquidation value is not known before the announcement ($\sigma_v^2>0$); and (ii) the announcement is informative ($\sigma_{\tilde{s},1}^2<\infty$).

In turn, the drift in conditional expectations,

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_2^i\right)-\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)\right]=\frac{V_1^i-V_2^i}{\sigma_v^2}s, \tag{66}$$

has the same sign of s if and only if $\sigma_v^2<\infty$ and

$$V_2^i<V_1^i\Leftrightarrow\frac{1}{\frac{1}{\sigma_v^2}+\frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}+\frac{1}{\sigma_{\tilde{s},1}^2}+\frac{(1-\rho)^2}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}}<\frac{1}{\frac{1}{\sigma_v^2}+\frac{1}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}+\frac{1}{\sigma_{\tilde{s},1}^2}}.$$

The latter is the case as long as: (i) the liquidation value is not known by the announcement ($\sigma_v^2>0$, $\alpha^2\sigma_\theta^2>0$ and $\sigma_{\tilde{s},1}^2>0$); (ii) the date 1 announcement is informative ($\sigma_{\tilde{s},1}^2<\infty$); (iii) there is endogenous information available at date 2 ($\rho<1$ and $\alpha^2\sigma_\theta^2<\infty$).

C.4.2 Average, Announcement Date Reaction and Drift in Prices

Given the common prior $\mathbb{E}\left(v|\,\mathcal{F}_o\right)=0$ and initial net supply $\theta_0=0$, the date 0 price equals the prior on the liquidation value, that is, $P_0=0$.

Averaging over supply shocks and substituting for $\mathbb{E}_\theta\left[\mathbb{E}\left(s|\,\mathcal{F}_2^c\right)\right]$ from (63), the average price at date 1 and 2 are given by

$$\mathbb{E}_\theta\left(P_1\right)=\left[1-(1-p_1)\frac{V_1^c}{\sigma_v^2}\right]s,\ \mathbb{E}_\theta\left(P_2\right)=\left[1-(1-p_2)\frac{V_2^c}{\sigma_v^2}\right]s. \tag{67}$$

Trivially, the price reaction to the announcement is

$$\mathbb{E}_\theta\left(\Delta P_1\right)=\mathbb{E}_\theta\left(P_1\right)=\left[1-(1-p_1)\frac{V_1^c}{\sigma_v^2}\right]s. \tag{68}$$

As we saw above, $p_1\geq0$ and, by definition, $V_1^c\leq\sigma_v^2$. Therefore, the sign of the price reaction to the announcement is the same as that of s unless $p_1=0\wedge\frac{V_1^c}{\sigma_v^2}=1$. The first equality equates to

$$p_1=0\Leftrightarrow\frac{V_2^i}{\sigma_{\tilde{s},1}^2}=0\Leftrightarrow\frac{1}{\frac{\sigma_{\tilde{s},1}^2}{\sigma_v^2}+\frac{1}{\alpha^2\sigma_{\tilde{s},1}^2\sigma_\theta^2}+1+\frac{(1-\rho)^2}{\alpha^2\sigma_{\tilde{s},1}^2\sigma_\theta^2}}=0,$$

which holds if any of the following is true: $\sigma_v^2=0$, $\alpha^2\sigma_\theta^2=0$ or $\sigma_{\tilde{s},1}^2=\infty$. The second equality becomes

$$\frac{V_1^c}{\sigma_v^2}=1\Leftrightarrow\frac{\sigma_v^2}{\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2}=0$$

which holds if any of the following is true: $\sigma_v^2 = 0$, $\sigma_{\tilde{s},1}^2 = \infty$ or $\alpha^2\sigma_\theta^2 = \infty$. Consequently, the price reaction to the announcement has the same sign of s if and only if: (i) the liquidation value is not known before the announcement ($\sigma_v^2 > 0$); and (ii) the announcement is informative ($\sigma_{\tilde{s},1}^2 < \infty$).

Substituting for p_1 and p_2 and taking the limit as $\beta \to 1$, the average price drift is

$$\mathbb{E}_\theta\left(\Delta P_2\right) = \frac{V_2^i\left(1-\rho+\alpha^2\sigma_{\tilde{s},1}^2\sigma_\theta^2\right)}{\sigma_v^2+\alpha^2\sigma_{\tilde{s},1}^2\sigma_\theta^2\left(\sigma_v^2+\sigma_{\tilde{s},1}^2+\alpha^2\sigma_{\tilde{s},1}^2\sigma_v^2\sigma_\theta^2\right)}\left(1-\rho\right)s. \tag{69}$$

It is easy to see that when there is no exogenous information at date 2, on average, prices drift on the same direction of s if and only if: there is exogenous information at date 1 and it is interpreted differentially ($0 < \sigma_{\tilde{s},1}^2 < \infty$); endogenous information is produced at date 2 ($\rho < 1$ and $\alpha^2\sigma_\theta^2 < \infty$); and not all uncertainty is resolved by date 1 ($\sigma_v^2 > 0$ and $\alpha^2\sigma_\theta^2 > 0$).

C.4.3 Average Expectation of Price Changes and Average Expected Demand

The date 1 expectation of future price changes ΔP_2 and ΔP_3 conditional on $\mathcal{F}_1^i$ are easily obtained from equations (9) and (10) as $\mathbb{E}\left(\left.\Delta P_2\right|\mathcal{F}_1^i\right) = C_2\psi_1$ and $\mathbb{E}\left(\left.\Delta P_3\right|\mathcal{F}_1^i\right) = C_3F_2\psi_1$. Averaging over supply shocks and interpretation errors, using expressions (62), (63) and (64), we get

$$\mathbb{E}_{\theta,i}\left(\psi_1\right) \equiv \begin{bmatrix} 1 \\ \mathbb{E}\left(\left.s\right|\mathcal{F}_1^i\right)-\mathbb{E}\left(\left.s\right|\mathcal{F}_1^c\right) \\ \mathbb{E}\left(\left.\theta_1\right|\mathcal{F}_1^i\right)\end{bmatrix} = \begin{bmatrix} 1 \\ \frac{V_1^c-V_1^i}{\sigma_v^2}s \\ -\frac{V_1^i}{\alpha\sigma_{\tilde{s},1}^2\sigma_v^2}s\end{bmatrix}.$$

C_3 is given by (30) without the third column, that is, $C_T = \begin{bmatrix} 0 & 1-p_2 & -p_{\theta,2}\end{bmatrix}$. Substituting for C_2 and F_2 determined above (equations (56) and (57) respectively), taking the limit as $\beta \to 1$ and simplifying, we obtain

$$\begin{aligned}
\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\left.\Delta P_2\right|\mathcal{F}_1^i\right)\right] &= -\frac{V_2^iV_1^i\left(1-\rho+\alpha^2\sigma_{\tilde{s},1}^2\sigma_\theta^2\right)\alpha^2\sigma_\theta^2}{\sigma_v^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\left(\sigma_v^2+\sigma_{\tilde{s},1}^2+\alpha^2\sigma_{\tilde{s},1}^2\sigma_v^2\sigma_\theta^2\right)}\left(1-\rho\right)s \\
\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\left.\Delta P_3\right|\mathcal{F}_1^i\right)\right] &= \frac{V_2^iV_1^i}{\sigma_{\tilde{s},1}^2\sigma_v^2}\left(1-\rho\right)s \\
\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\left.\Delta P_2+\Delta P_3\right|\mathcal{F}_1^i\right)\right] &= \frac{V_2^iV_1^i\left[\sigma_v^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\left(\sigma_{\tilde{s},1}^2+\rho\sigma_v^2\right)\right]}{\sigma_{\tilde{s},1}^2\sigma_v^2\left[\sigma_v^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\left(\sigma_v^2+\sigma_{\tilde{s},1}^2+\alpha^2\sigma_\theta^2\sigma_{\tilde{s},1}^2\sigma_v^2\right)\right]}\left(1-\rho\right)s.
\end{aligned}$$

It is easy to see that, on average, the price is expected to move in the opposite (same) direction of s from date 1 to date 2 (date 2 to date 3 and date 1 to date 3) in the same circumstances that, on average, prices drift.

The average date 1 expectation of the date 2 demand is obtained from equations (10) and (12) as

$$\mathbb{E}_{\theta,i}\left[E\left(\left.X_2^i\right|\mathcal{F}_1^i\right)\right] = \frac{1}{\alpha}Q_2F_2\mathbb{E}_{\theta,i}\left(\psi_1\right) = \frac{V_1^i\left(1-\rho\right)}{\alpha\sigma_{\tilde{s},1}^2\sigma_v^2}s$$

which has the same sign of s in the same circumstances that, on average, prices drift.

Conditioning on $\tilde{s}_1^i$ only, it is easy to verify that, in the limit case of $\beta \to 1$:

$$\mathbb{E}\left(\left.s\right|\tilde{s}_1^i\right) = \mathbb{E}\left(\left.\xi_1\right|\tilde{s}_1^i\right) = \mathbb{E}\left(\left.\xi_2\right|\tilde{s}_1^i\right) = \frac{\frac{1}{\sigma_{\tilde{s},1}^2}\tilde{s}_1^i}{\frac{1}{\sigma_v^2}+\frac{1}{\sigma_{\tilde{s},1}^2}},\ \mathbb{E}\left(\left.\theta_1\right|\tilde{s}_1^i\right) = \mathbb{E}\left(\left.\theta_2\right|\tilde{s}_1^i\right) = 0,$$

$$\mathbb{E}\left[\mathbb{E}\left(s|\,\mathcal{F}_1^c\right)|\,\tilde{s}_1^i\right]=\frac{\sigma_v^4\tilde{s}_1^i}{\left(\sigma_{\tilde{s},1}^2+\sigma_v^2\right)\left(\sigma_v^2+\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2\right)},\ \mathbb{E}\left[\mathbb{E}\left(s|\,\mathcal{F}_2^c\right)|\,\tilde{s}_1^i\right]=\frac{\left[1+\left(1-\rho^2\right)\right]\sigma_v^4\tilde{s}_1^i}{\left(\sigma_{\tilde{s},1}^2+\sigma_v^2\right)\left(\left[1+\left(1-\rho^2\right)\right]\sigma_v^2+\alpha^2\sigma_{\tilde{s},1}^4\sigma_\theta^2\right)}.$$

The expected price changes conditional on $\tilde{s}_1^i$ are then given by

$$\begin{aligned}
\mathbb{E}\left(\Delta P_2|\,\tilde{s}_1^i\right) &= (1-p_2)\,\mathbb{E}\left[\mathbb{E}\left(s|\,\mathcal{F}_2^c\right)|\,\tilde{s}_1^i\right]-(1-p_1)\,\mathbb{E}\left[\mathbb{E}\left(s|\,\mathcal{F}_1^c\right)|\,\tilde{s}_1^i\right]+(p_2-p_1)\,\mathbb{E}\left(s|\,\tilde{s}_1^i\right)\\
&= \frac{V_2^i\left(1-\rho+\alpha^2\sigma_{\tilde{s},1}^2\sigma_\theta^2\right)}{\sigma_v^2+\alpha^2\sigma_{\tilde{s},1}^2\sigma_\theta^2\left(\sigma_v^2+\sigma_{\tilde{s},1}^2+\alpha^2\sigma_{\tilde{s},1}^2\sigma_v^2\sigma_\theta^2\right)}\frac{\sigma_v^2}{\sigma_v^2+\sigma_{\tilde{s},1}^2}\left(1-\rho\right)\tilde{s}_1^i\\
\mathbb{E}\left(\Delta P_3|\,\tilde{s}_1^i\right) &= (1-p_2)\left\{\mathbb{E}\left(s|\,\tilde{s}_1^i\right)-\mathbb{E}\left[\mathbb{E}\left(s|\,\mathcal{F}_2^c\right)|\,\tilde{s}_1^i\right]\right\}\\
&= \frac{V_2^i}{\sigma_v^2}\frac{\sigma_v^2}{\sigma_v^2+\sigma_{\tilde{s},1}^2}\tilde{s}_1^i.
\end{aligned}$$

Since, $\mathbb{E}_i\left(\tilde{s}_1^i\right)=s$, on average when conditioning only on $\tilde{s}_1^i$, investors expect prices to drift in the same direction of s.

C.5 Model with Exogenous Information

Following the same steps as in the previous case but setting $\rho=1$, we obtain the date 1 price coefficients

$$p_1=\frac{V_1^i}{\sigma_{\tilde{s},1}^2},\ p_{\theta,1}=-\alpha V_1^i.$$

The drift in average expectations and prices (using equation (67)) is given by

$$\begin{aligned}
\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)-\mathbb{E}\left(v|\,\mathcal{F}_0\right)\right]=\mathbb{E}_\theta\left[\mathbb{E}\left(\Delta P_1|\,s\right)\right] &= \left(1-\frac{V_1^i}{\sigma_v^2}\right)s\\
\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(v|\,\mathcal{F}_2^i\right)-\mathbb{E}\left(v|\,\mathcal{F}_1^i\right)\right]=\mathbb{E}_\theta\left[\mathbb{E}\left(\Delta P_2|\,s\right)\right] &= \frac{V_1^i-V_2^i}{\sigma_v^2}s\\
\mathbb{E}_{\theta,i}\left[v-\mathbb{E}\left(v|\,\mathcal{F}_2^i\right)\right]=\mathbb{E}_\theta\left[\mathbb{E}\left(\Delta P_3|\,s\right)\right] &= \frac{V_2^i}{\sigma_v^2}s
\end{aligned}$$

It is straightforward to determine that the reaction to the announcement and drifts have the same sign as s in the same conditions we determined for the model with endogenous information. The only difference is that now drifts occur if there is additional exogenous information at date 2 ($\beta<1$) instead of additional endogenous information ($\rho<1$ and $\alpha^2\sigma_\theta^2<\infty$).

The average date 1 expectation of future price changes is

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\Delta P_2|\,\mathcal{F}_1^i\right)\right]=\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(\Delta P_3|\,\mathcal{F}_1^i\right)\right]=0$$

and the average date 1 expectation of date 2 demand is

$$\mathbb{E}_{\theta,i}\left[\mathbb{E}\left(X_2^i\middle|\,\mathcal{F}_1^i\right)\right]=0.$$

Conditioning only on $\tilde{s}_1^i$, the expected price change is

$$\mathbb{E}\left(\Delta P_2|\,\tilde{s}_1^i\right)=\frac{V_1^i-V_2^i}{\sigma_v^2+\sigma_{\tilde{s},1}^2}\tilde{s}_1^i,\ \mathbb{E}\left(\Delta P_3|\,\tilde{s}_1^i\right)=\frac{V_2^i}{\sigma_v^2+\sigma_{\tilde{s},1}^2}\tilde{s}_1^i.$$

D Extension to Two Types of Investors

The model can be easily extended to allow for any number of distinct types of investors, differing in their levels of risk aversion and/or the precision of their interpretation of information. The only requirement is that there exists an infinite number of investors of each type so that, by the law or large numbers, the average interpretation by each type of investor is unbiased. I will only consider the case of two types of investors, but the generalization to more types of investors is straightforward.

The two types of investors are indexed by $j = \{1, 2\}$; the fraction of type 1 investors is λ. Since different types of investors interpret information with different degrees of precision, they will form different beliefs when conditioning on their interpretation ($\mathcal{F}_t^i$), even if they interpret information exactly in the same way.[38] To obtain the optimal demand for an investor of each type, we use Lemmas 2 to 4 with the necessary adaptations for each type of investors. Specifically, K_t in Lemma 2 becomes type dependent through $\Sigma_{y,t}$, which reflects the different precision of interpretation across types of investors. In Lemma 3, ψ_t, G_t and $\Sigma_{\Delta,t}$ also become type dependent. Finally, in Lemma 4 Q_t and α become type dependent as well.

Denoting type specifics variables by superscripting them with (j), the market clearing condition becomes

$$\begin{aligned}
&\frac{\lambda}{\alpha^{(1)}}\int_i \left[q_{\boldsymbol{s},t}^{i(1)}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{i(1)}\right) + q_{\boldsymbol{s},t}^{c(1)}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{c}\right) + q_{\theta,t}^{i(1)}\mathbb{E}\left(\theta_t|\,\mathcal{F}_t^{i(1)}\right)\right] + \\
&\frac{1-\lambda}{\alpha^{(2)}}\int_i \left[q_{\boldsymbol{s},t}^{i(2)}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{i(2)}\right) + q_{\boldsymbol{s},t}^{c(2)}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{c}\right) + q_{\theta,t}^{i(2)}\mathbb{E}\left(\theta_t|\,\mathcal{F}_t^{i(2)}\right)\right] = \theta_t \\
\Leftrightarrow\quad &\left[\lambda\frac{q_{\boldsymbol{s},t}^{c(1)}}{\alpha^{(1)}} + (1-\lambda)\frac{q_{\boldsymbol{s},t}^{c(2)}}{\alpha^{(2)}}\right]\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{c}\right) + \left[\lambda\frac{q_{\theta,t}^{i(1)}}{\alpha^{(1)}} + (1-\lambda)\frac{q_{\theta,t}^{i(2)}}{\alpha^{(2)}} - 1\right]\theta_t + \\
&\left[\lambda\frac{q_{\theta,t}^{i(1)}}{\alpha^{(1)}} + (1-\lambda)\frac{q_{\theta,t}^{i(2)}}{\alpha^{(2)}}\right]\frac{\boldsymbol{p}_t}{p_{\theta,t}}\boldsymbol{s} + \frac{\lambda}{\alpha^{(1)}}\left[q_{\boldsymbol{s},t}^{i(1)} - q_{\theta,t}^{i(1)}\frac{\boldsymbol{p}_t}{p_{\theta,t}}\right]\int_i \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{i(1)}\right) + \\
&\frac{1-\lambda}{\alpha^{(2)}}\left[q_{\boldsymbol{s},t}^{i(2)} - q_{\theta,t}^{i(2)}\frac{\boldsymbol{p}_t}{p_{\theta,t}}\right]\int_i \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{i(2)}\right) = 0
\end{aligned}$$

with the equivalence following from substitution of equation (38). Substituting $\int_i \mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{i(j)}\right)$, $j = \{1, 2\}$ for equation (41) with type dependent matrices Γ_t and $\hat{\Gamma}_t$, the market clearing condition can be written as

$$\begin{aligned}
&\left\{\frac{\lambda}{\alpha^{(1)}}\left[q_{\boldsymbol{s},t}^{c(1)} + \left(q_{\boldsymbol{s},t}^{i(1)} - q_{\theta,t}^{i(1)}\frac{\boldsymbol{p}_t}{p_{\theta,t}}\right)\left(I_n - \hat{\Gamma}_t^{(1)}\right)\right] + \frac{1-\lambda}{\alpha^{(2)}}\left[q_{\boldsymbol{s},t}^{c(2)} + \left(q_{\boldsymbol{s},t}^{i(2)} - q_{\theta,t}^{i(2)}\frac{\boldsymbol{p}_t}{p_{\theta,t}}\right)\left(I_n - \hat{\Gamma}_t^{(2)}\right)\right]\right\}\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{c}\right) \\
&+\left\{\frac{\lambda}{\alpha^{(1)}}\left[q_{\theta,t}^{i(1)}\frac{\boldsymbol{p}_t}{p_{\theta,t}} + \left(q_{\boldsymbol{s},t}^{i(1)} - q_{\theta,t}^{i(1)}\frac{\boldsymbol{p}_t}{p_{\theta,t}}\right)\hat{\Gamma}_t^{(1)}\right] + \frac{1-\lambda}{\alpha^{(2)}}\left[q_{\theta,t}^{i(2)}\frac{\boldsymbol{p}_t}{p_{\theta,t}} + \left(q_{\boldsymbol{s},t}^{i(2)} - q_{\theta,t}^{i(2)}\frac{\boldsymbol{p}_t}{p_{\theta,t}}\right)\hat{\Gamma}_t^{(2)}\right]\right\}\boldsymbol{s} \\
&+\left[\lambda\frac{q_{\theta,t}^{i(1)}}{\alpha^{(1)}} + (1-\lambda)\frac{q_{\theta,t}^{i(2)}}{\alpha^{(2)}} - 1\right]\theta_t = 0.
\end{aligned}$$

All the entries of the vectors multiplying $\mathbb{E}\left(\boldsymbol{s}|\,\mathcal{F}_t^{c}\right)$ and $\boldsymbol{s}$ and the scalar multiplying θ_t have to be simultaneously equal to zero $\forall t$ for the market clearing condition to be satisfied at all dates. This gives us $(T-1)\times(2n+1)$ nonlinear equations from which the same number of price function coefficients can be determined. Note that, although the system is underidentified for vectors $Q_t^{(j)}$, $j = \{1, 2\}$, $t = \{1, 2, ..., T-1\}$, it is exactly identified for the price function coefficients. Vectors $Q_t^{(j)}$ can be determined once we obtain the price function coefficients.

[38] In contrast, investors of the same type form different expectations only if they interpret information differently.

References

Allen, F., S. Morris, and H. Shin (2006). Beauty contests and iterated expectations in asset markets. *Review of Financial Studies 19*(3), 719–752.

Ball, R. and E. Bartov (1996, August). How naive is the stock market's use of earnings information? *Journal of Accounting and Economics 21*(3), 319–337.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research 6*(2), 159–178.

Banerjee, S., R. Kaniel, and I. Kremer (2009). Price drift as an outcome of differences in higher order beliefs. *Review Of Financial Studies 22*(9), 3707–3734.

Banerjee, S. and I. Kremer (2010). Disagreement and learning: Dynamic patterns of trade. *The Journal of Finance 65*(4), 1269–1302.

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics 49*(3), 307–343.

Bartov, E., S. Radhakrishnan, and I. Krinsky (2000, January). Investor sophistication and patterns in stock returns after earnings announcements. *The Accounting Review 75*(1), 43–63.

Berkman, H. and P. Koch (2008). Noise trading and the price formation process. *Journal of Empirical Finance 15*(2), 232–250.

Bernard, V. and J. Thomas (1989). Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research 27*(1), 1–35.

Bernard, V. and J. Thomas (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics 13*(4), 305–340.

Biais, B., P. Bossaerts, and C. Spatt (2010). Equilibrium asset pricing and portfolio choice under asymmetric information. *Review of Financial Studies 23*(4), 1503–1543.

Brandt, M., R. Kishore, P. Santa-Clara, and M. Venkatachalam (2008). Earnings announcements are full of surprises. Working paper, Duke University and UCLA.

Campbell, J., S. Grossman, and J. Wang (1993). Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics 108*(4), 905–939.

Cespa, G. and X. Vives (2010). Dynamic trading and asset prices: Keynes vs. Hayek. *Review of Economic Studies*, forthcoming.

Chan, L., N. Jegadeesh, and J. Lakonishok (1996, December). Momentum strategies. *The Journal of Finance 51*(5), 1681–1713.

Chan, W. (2003). Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics 70*(2), 223–260.

Choi, W. and J.-W. Kim (2001). Underreaction, trading volume, and post-earnings announcement drift. Working paper, Harvard University and University of Alberta.

Chordia, T. and L. Shivakumar (2006, June). Earnings and price momentum. *Journal of Financial Economics 80*(3), 627–656.

Cohen, R., P. Gompers, and T. Vuolteenaho (2002). Who underreacts to cash-flow news? Evidence from trading between individuals and institutions. *Journal of Financial Economics 66*(3), 409–462.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998, December). Investor psychology and security market under- and overreactions. *The Journal of Finance 53*(6), 1839–1885.

DellaVigna, S. and J. Pollet (2009). Investor inattention and Friday earnings announcements. *The Journal of Finance 64*(2), 709–749.

Dow, J. and G. Gorton (2008). Noise traders. In S. Durlauf and L. Blume (Eds.), *The new Palgrave dictionary of economics*. New York: Palgrave Macmillan.

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

Fama, E. and K. French (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance 51*(1), 55–84.

Foster, G., C. Olsen, and T. Shevlin (1984). Earnings releases, anomalies, security returns and the behavior of security returns. *The Accounting Review 59*(4), 574–603.

Garfinkel, J. and J. Sokobin (2006). Volume, opinion divergence, and returns: A study of post-earnings announcement drift. *Journal of Accounting Research 44*(1), 85–112.

Gilbert, T., S. Kogan, L. Lochstoer, and A. Ozyildirim (2011). Investor inattention and the market impact of summary statistics. Working paper, University of Texas at Austin.

Gompers, P. and J. Lerner (1998). Venture capital distributions: Short-run and long-run reactions. *The Journal of Finance 53*(6), 2161–2183.

Grinblatt, M., R. Masulis, and S. Titman (1984). The valuation effects of stock splits and stock dividends. *Journal of Financial Economics 13*(1), 97–112.

Grossman, S. and J. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review 70*(3), 393–408.

Grundy, B. and M. Mcnichols (1989). Trade and the revelation of information through prices and direct disclosure. *Review of Financial Studies 2*(4), 495–526.

Harris, M. and A. Raviv (1993). Differences of opinion make a horse race. *Review of Financial Studies 6*(3), 473–506.

He, H. and J. Wang (1995). Differential information and dynamic behavior of stock trading volume. *Review of Financial Studies 8*(4), 919–972.

Hirshleifer, D., S. Lim, and S. Teoh (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance 64*(5), 2289–2325.

Hirshleifer, D. and S. Teoh (2003, December). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics 36*(3), 337–386.

Hong, H., T. Lim, and J. Stein (2000, February). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance 55*(1), 265–295.

Hong, H. and J. Stein (1999). A unified theory of underreaction, momentum trading and overreaction in asset markets. *The Journal of Finance 54*(6), 2143–2184.

Huberman, G. and T. Regev (2001). Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *The Journal of Finance 56*(1), 387–396.

Ikenberry, D., J. Lakonishok, and T. Vermaelen (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics 39*(2), 181–208.

Kandel, E. and N. Pearson (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy 103*(4), 831–872.

Liptser, R. and A. Shiryaev (2001). *Statistics of Random Processes I & II* (second ed.). New York: Springer.

Loughran, T. and J. Ritter (1995). The new issues puzzle. *The Journal of Finance 50*(1), 23–52.

Michaely, R., R. Thaler, and K. Womack (1995). Price reactions to dividend initiations and omissions: Overreaction or drift? *The Journal of Finance 50*(2), 573–608.

Michaely, R. and K. Womack (1999). Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies 12*(4), 653–686.

Sinha, N. (2011). Underreaction to news in the US stock market. Working paper, University of Illinois at Chicago.

Zhang, F. (2006). Information uncertainty and stock returns. *The Journal of Finance 61*(1), 105–136.