FSA in an ETF World

by

Russell J Lundholm
Sauder Business School
University of British Columbia

October 2018

This paper models the value of conducting financial statement analysis (FSA) in the presence of an electronically traded fund (ETF) that gives exposure to the firm's systematic value. FSA is characterized as a costly process that yields a private signal about the idiosyncratic portion of a firm's future payoffs. The value of this signal depends on how much of the resulting information is available for free to uninformed traders by observing price, which is turn depends on the noise in the economy. A popular argument is that ETFs are attracting noise traders away from the underlying firm, which makes prices more informative and private information less valuable. While this may be true, I find that introducing an ETF into the market also provides a discrete increase in the value of FSA, but only if investors use the ETF to hedge out exposure to the portion of firm value that they are uninformed about. The net result is that in equilibrium there is an increase in the fraction of investors who conduct FSA. This result is unavailable in previous theoretical papers about ETFs because they modeled investors as being risk neutral, thus eliminating their desire to hedge out uncertainty.

INTRODUCTION

In 2018 the global market for exchange traded funds (ETFs) topped $5000 billion, up from $772 billion 10 years earlier (Blackrock 2018). The breadth of funds continues to grow as well, with almost 8000 different products offering exposure to everything from the S&P 500 index to narrowly defined industries to bit coin. It is estimated that equity ETFs account for over a third of all trading volume in the US. The advent of ETFs, and the associated shift toward passive investing, has had a significant impact on financial markets. In an article titled "*ETF growth is in 'danger of devouring capitalism'*" the *Financial Times* reports "This shift out of traditional, "active" money management is one of the most profound changes to the global financial system in history and is now powerful enough to rewire how markets function" (Wigglesworth, February 4, 2018). Empirically, it has been shown that increased ETF ownership of the underlying stock can result in decreased liquidity, increased price synchronicity between stocks in the ETF, and a decrease in the degree to which stock prices reflects future earnings (Israeli, Lee and Sridharan 2017). In short, the fear is that ETFs may be making the stock market less efficient.

On the surface this concern may seem misplaced, as we generally assume that inefficiency breeds opportunity. If the market for the firm is becoming less efficient then surely a savvy investor could profit on this by collecting private information and trading in the inefficient market. On the other hand, if liquidity is falling then it might be more difficult to profit from being privately informed. The purpose of this paper is to model the forces at work when an ETF is introduced into the market, with a particular emphasis on how this event changes the value of being privately informed about the firm.

Since the 1934 publication of "Security Analysis" by Graham and Dodd, financial statement analysis (FSA) has focused on understanding and predicting firm-specific outcomes. Thus, any change in the value of being privately informed about the firm has immediate implications for FSA. In response to the rise of ETFs, Federico Kaune of UBS Asset Management was quoted in the *Financial Times* as saying "We can no longer be the fundamental investors we want to be." In this paper I show that some of the concerns raised about the impact of ETFs on the value of information are valid. However, I also show that the introduction of the ETF creates a new trading instrument that changes investors' optimal strategies, with the net result that FSA is even more valuable.

My model features two risky assets – a firm whose payoffs have common and idiosyncratic components and an ETF whose payoffs depend only on the common component. This abstraction captures the idea that the ETF averages out the idiosyncratic variation due to its constituent firms, giving the investor a pure exposure to the common component. In this setting, I characterize FSA as collecting information about the idiosyncratic component of the firm payoff, and I study how the value of this signal changes with the introduction of the ETF. The value of being privately informed is always relative to remaining uninformed and gleaning whatever information is available from observing prices. My model finds the rational expectations equilibrium for the two risk assets in order compare the value of information with and without an ETF asset in the market.

I begin by showing that if one considers a market with only the firm asset, then indeed the value of information falls as the liquidity of the firm asset falls, as the article in the *Financial Times* would suggest. But what this analysis misses is that the introduction of ETF gives investors a mechanism for hedging out the common variation in the firm's payoff. Once we allow investors the opportunity to hedge, the value of information is actually higher after the ETF is introduced. This occurs because, once informed investors hedge out their exposure to the common component of the firm's payoff, their private signal about the idiosyncratic portion of the payoff is perfectly aligned with the variation in their net position. In short, they can bet more aggressively on the signal that FSA provides them because hedging with the ETF allows them to remove the exposure to variation that they are uninformed about. Further, if the fraction of informed investors is allowed to change in response to the changing value of being informed, I show that the equilibrium fraction of informed traders actually increases after the introduction of an ETF.

RELATED LITERATURE

While there is no direct empirical work studying the impact of ETFs on the value of information, the literature on how they impact market efficiency is closely related.[1] The evidence is mixed on whether ETFs help or hinder the efficient pricing of the underlying securities. Holding aside for a moment their impact on liquidity, to the extent ETFs offer a low-cost way to invest in a common factor, they may improve the efficient pricing of this factor. Glosten, Nallareddy, and Zou (2016) find that stocks incorporate systematic information more quickly once they are in ETF portfolios. Further, they argue that the increased comovement of stocks after their inclusion in an ETF is due to the better pricing of systematic information. Consistent with ETFs bringing more common factor bets to the market, Bhojraj, Mohanram and Zhang (2018) find that information transfers about constituent firms' earnings are greater for firms included in a sector ETF. In addition Li and Zhu (2016) argue that informed traders use ETFs to circumvent short-sale constraints on the underlying stock. They document that the portion of a stock's short interest that is due to investors shorting an ETF predicts future stock returns and conclude that ETFs help improve market efficiency.

The aforementioned predictions may be reversed once we take into account how the introduction of an ETF influences liquidity in the underlying stock. Hamm (2014) finds a negative association between ETF ownership and liquidity in the underlying stock. Consistent with this, Israeli, Lee, and Sridharan (2017) show that trading costs increase and liquidity in the underlying stock falls after an increase in firm ownership by an ETF. They also find that an increase in ETF ownership causes the underlying stock price to be less reflective of future earnings, and they offer some limited evidence that the number of analysts falls. The reduction in liquidity lowers the value of being informed, and so analysts no longer cover the stock, and

[1] There is a rich literature regarding the microstructure of ETFs; how they are created, how they are priced, and why they offer extremely low transaction costs. Lettau and Madhavan (2018) detail the inner workings of ETFs and related research into it. My model relies only on the facts that an ETF is a basket of securities that collectively provide exposure to some commonly held attribute of the constituent firms, and that they can be longed or shorted at very low cost.

the efficient pricing of its future earnings is diminished.[2] However, it is also possible that the liquidity of ETFs attracts high-frequency traders and that this new source of demand migrates to the underlying securities because ETFs and their constituent securities are tied by arbitrage (Ben-David, Franzoni, and Moussawi 2017). Therefore, the liquidity-driven demand for the ETF can actually induce greater liquidity for the underlying security.[3]

To summarize the empirical literature, ETFs have attracted a lot of attention in the capital market and, depending on the nature of the attention, it could help or hinder the efficient pricing of the underlying stock. The most direct effect is that ETFs allow exposure to a common factor at low cost. They also provide a roundabout way to short a difficult-to-short underlying stock. How the ETF affects liquidity produces more indirect effects. If liquidity traders migrate away from the underlying stock to the ETF then this would lower the value of being informed about the asset's idiosyncratic payoff, and therefore lower information production. Alternatively, if the ETF itself demands a large amount of liquidity, this will create additional liquidity demand in the underlying stock, which would increase the value of private information. There is some evidence supporting all of these results. The model presented here allows the study of both firm and ETF liquidity shocks. The model is about the value of conducting FSA, which is directly influenced by the degree of market inefficiency but, as I show, it is also impacted by the trading opportunities available to the informed investors.

Turning to the theoretical literature, Gorton and Pennacchi (1993) use a model of risk neutral traders and imperfect competition (i.e. a Kyle 1985 model) to show that uninformed liquidity traders are better off trading in the ETF than in the underlying security. This occurs because the uninformed traders don't face adverse

---

[2] The Israeli et al (2017) results may appear at odds with Glosten et al (2016). However, Israeli et al (2017) show that the Glosten et al results apply to the *concurrent* pricing of information about the *common* component of the news, while their results apply to the pricing of *future* information, and is largely driven by the *idiosyncratic* component of the news.

[3] Another branch of the ETF literature considers how the ETF aids trading in assets that are otherwise illiquid. Bhattacharya and O'Hara (2017) consider the case where there is no market for the underlying asset and so informed traders must use the ETF as an indirect means of trading on their information.

selection in the ETF market. Cong and Xu (2016) show that this result holds after the ETF weights on the underlying assets are optimized. Subrahmanyam (1991) generalizes the payoff structure to one with a common component and an idiosyncratic component (as in my model) and introduces common-factor-informed investors along with idiosyncratically-informed investors. He shows that, even in the presence of factor-informed traders, liquidity traders are still generally better off meeting their needs in the ETF market. He also gives conditions whereby the equilibrium number of idiosyncratically-informed investors falls when an ETF is introduced to the market. This result is a notable contrast to the results presented here, where the value of being idiosyncratically-informed increases after the introduction of the ETF, causing the number of informed traders to increase. The key difference in the models is that Subrahmanyam assumes risk neutral investors, with the consequence that they do not hedge.[4] As I show later, risk adverse investors use the ETF asset to hedge out exposure to the common component of the payoff, and this increases the value of their private signal.

THE MODEL

The value of financial statement analysis is characterized by the value of the incremental expected utility afforded a privately informed agent relative to an uninformed agent. Precisely, it is the value Φ that equates the expected utility of an agent who pays Φ and observes the private signal Y with the expected utility of an agent who pays nothing and only gains the information available from observing the asset's price. Φ is denominated in units of the economy's riskless asset, which serves as the numeraire. This value is computed after taking the agent's optimal strategy into account, and then integrates over all the information variables in the model. As such, it is the amount the agent is willing to pay for the signal before the signal realization is known. As shown in Alles and Lundholm (1993) and Admati Pfliederer (1987), for agents with negative exponential utility and risk tolerance parameter ρ, observing normal random variables, Φ is given as

[4] To see this, note in Subrahmanyam (1991) that the derived demands for the firm asset and the ETF asset are always in the same direction (compare equation 5 on page 22 and equation 8 on page 24).

$$\Phi = \left(\frac{\rho}{2}\right) log \left\{\frac{detV_U}{detV_I}\right\}, \tag{1}$$

where V_U and V_I are the posterior variance-covariance matrices of beliefs about all risky asset payoffs held by the uninformed traders and informed traders, respectively. For instance, if there are two risky assets then each type of trader has a 2x2 matrix specifying her posterior variance for each asset and the posterior covariance between them. The determinant operator converts the matrix into a scalar. As the difference in the informed and uninformed variances grows, the value of information increases. The expression in (1) is completely general, requiring only negative exponential utility and normal random variables. However, to use this formula, we need to solve for the V_U and V_I matrices. This requires determining the amount of information the uninformed traders glean from observing price, which in turn requires specification of the equilibrium rational expectations price. For this, we need the details of the asset payoffs and information signals.

Market with One Risky Asset

In order to illustrate the impact of introducing an ETF into the market, we first need to establish the value of information in a market with only the firm asset. Assume there is a competitive market for one risky asset with payoff F, composed of a common component F_c and an idiosyncratic component F_i ; $F = F_c + F_i$. The common component can be thought of as an industry factor. The common component F_c is distributed normal with mean η and variance g, the idiosyncratic component is distributed normal with mean 0 and variance h, and the two components are independent.

There is a private signal $Y_i = F_i + \varepsilon$ held by the λ informed traders, where ε is normal with mean 0 and variance s. All informed traders see exactly the same Y_i (the subscript i denotes 'idiosyncratic').[5] There is no information about the common

[5] There are two common options for describing the signal error's covariance across investors – either the error is independent across all agents or it is perfectly correlated, as in my model. Neither

component beyond the common knowledge of its mean and variance. In this context, financial statement analysis is characterized as a process that results in possessing the private signal Y_i. The process is completely focused on learning about firm-specific payoff-relevant information. While FSA certainly entails industry analysis, in virtually all textbooks the reason given for studying the industry is to gain a better understanding of the firm's idiosyncratic place in it. Thus, as an abstraction, the private information produced by the FSA process is only about F_i. The central variable of interest in this paper is Φ, the value of possessing the private signal Y_i. We are interested in comparing its value across markets with and without an ETF asset, and in seeing how its value changes as noise trading moves from the firm asset to the ETF asset.[6]

The market for the risky asset is composed of λ informed traders and $(1\text{-}\lambda)$ uninformed traders; all with negative exponential risk aversion and risk tolerance parameter ρ. Note that to study the value of information it is desirable to begin with a model that responds to risk aversion. Absent risk aversion, investors have no reason to engage in actions that hold the mean payoff constant but lower the variance of payoffs. In particular, investors have no reason to hedge, and hedging will be turn out to be a consequence of introducing an ETF. For this reason, I begin with the standard competitive market model with risk aversion rather than Kyle's 1985 model of risk neutral traders in a non-competitive market.

While the physical supply of the risky asset is known and fixed, I model liquidity shocks, or noise trading, as random variation around this known supply. One of the main observations from Israeli et al (2017) is that liquidity trading in the underlying security diminishes as the ETF draws away this activity. The precise source of the randomness about liquidity trading is not important to my model; it could be due to "noise traders" as defined by Black (1986), where individuals

assumption is ideal. In my model, every analyst who conducts FSA reaches exactly the same conclusion. However, the alternative of independent errors is also unrealistic, implying that the analysts collectively know the value of the payoff perfectly. As shown later, there are some practical modeling advantages to assuming all informed agents see exactly the same signal.

[6] There is little tension in a model where investors have a signal about the common factor and then an ETF is added to the market. The ETF allows the factor-informed trader the perfect mechanism to use her information, and so the value of the factor information increases.

behave as if they have information even though they actually do not. Or it could be traditional liquidity trading where the demand/supply shock is driven by preferences outside the model (for example, the immediate desire to buy a boat). The random liquidity/noise shocks make it impossible for the uninformed traders to perfectly infer the informed traders' underlying private information. Instead, price is a noisy signal about their private information. The supply of the risky asset is given by X_1 with variance w_1 and is independent of the other random variables in the model. Later we will study how shifting noise from the underlying security to an ETF security impacts the value of information.

With this, we can compute the posterior beliefs of the informed and uninformed traders.[7] The λ informed traders, upon observing the private signal Y_i have the following posterior belief about the mean E_I and variance V_I of $F = F_c + F_i$.

$$E_I = \eta + \frac{hY_i}{(h+s)} \; and \; V_I = g + \frac{hs}{h+s}. \tag{2}$$

The posterior mean increases with the signal, and at an increasing rate as the signal error variance s decreases. The posterior variance increases with the common variation g. It also increases with the idiosyncratic variation h, but at a rate that decreases as the signal error variance decreases.

The $(1\text{-}\lambda)$ uninformed traders use price to infer what they can about the informed traders' signal. Assuming that price is linear in the informed traders' signal and the realized supply of the asset, $P = a_0 + a_1Y_i - a_2X_1$, the uninformed trader can compute a linear transformation of P to deduce the information signal $\hat{P}$ as

[7] The general expressions for the conditional mean and variance of normal random variables are as follows. Let F be an n-dimensional vector of payoffs with prior mean vector μ and let Y be an m-dimensional vector of signals about those payoffs, with no particular covariance structure, and mean vector υ. The covariance matrix of $F \cup Y$ is n+m dimensional and symmetric. Partition this matrix into an nxn matrix of the top left corner, labeled Σ_{11}, an mxm matrix in the bottom right corner labeled Σ_{22}, an nxm matrix of the top right corner labeled Σ_{12} and an mxn matrix in the bottom left corner, labeled Σ'_{12}. The nx1 posterior mean vector is then given by $E = \mu + \Sigma_{12}\Sigma_{22}^{-1}(Y - \upsilon)$ and the nxn posterior covariance matrix is given by $V = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma'_{12}$ (Welch 2014).

$$\hat{P} = \frac{(P - a_0)}{a_1} = Y_i - \frac{a_2}{a_1} X_1 = Y_i - Z * X_1. \tag{3}$$

The precise value of $Z = a_2/a_1$ is found later as an equilibrium. With this, the uninformed traders, upon observing price, have the following posterior belief about the mean E_U and variance V_U of $F = F_c + F_i$.

$$E_U = \eta + \frac{h\hat{P}}{(h + Q_{1a})} \; and \; V_U = g + \frac{hQ_{1a}}{h + Q_{1a}} \;, where \; Q_{1a} = s + Z^2 w_1. \tag{4}$$

The subscript '1a' on Q_{1a} stands for the one-asset economy, to distinguish between this model and the two-asset economy that comes later. We will refer to Q_{1a} as the 'noise in price.' Combining these beliefs with the negative exponential utility function gives the informed traders and uninformed traders demands, respectively, as

$$D_I = \rho V_I^{-1}[E_I - P] \; and \; D_U = \rho V_U^{-1}[E_U - P]. \tag{5}$$

Market clearing equates these demands with the supply:

$$\lambda D_I + (1 - \lambda) D_U = X_1. \tag{6}$$

The equilibrium price is found by plugging the informed and uninformed expectations into (5), and then equating the ratio of coefficients on Y_i and X_1 with $Z = a_2/a_1$. The result is a cubic equation in Z that has a unique solution, as given in the first lemma, and derived in the appendix.

<u>Lemma 1</u>: Define $G = gh + gs + hs$. The equilibrium Z and the resulting noise in the price signal Q_{1a} are given as

$$Z = \frac{G}{\lambda\rho h} \; and \; Q_{1a} = s + \left(\frac{G}{\lambda\rho h}\right)^2 w_1. \tag{7}$$

The noise in price is increasing in all the variance parameters g, h, and s, and in the exogenous liquidity noise w_1. This is a modest generalization of the rational expectations equilibrium when there is no common component to the payoff (i.e. when $F = F_i$), as given in Grossman and Stiglitz (1980). To see this, set $g = 0$ and note that $Z = s/\lambda\rho$, as in their model.[8]

By applying (1), the value of information in this case is

$$\Phi(1\text{ asset}) = \left(\frac{\rho}{2}\right) log \left\{\frac{g + \frac{hQ_{1a}}{h + Q_{1a}}}{g + \frac{hs}{h + s}}\right\}, where \; Q_{1a} = s + Z^2 w_1 \; and \; Z = \frac{G}{\lambda\rho h}. \tag{8}$$

This leads to our first result regarding the value of information.

Lemma 2: The value of information Φ in the one-asset economy is increasing in the liquidity noise w_1 and decreasing in the fraction of informed traders λ.

To see the first part, note that the equilibrium Z given in (7) does not depend on the liquidity noise w_1. Therefore the value of information in (8) is increasing in the price noise Q_{1a}, which is increasing in w_1. The second part follows because Q_{1a} is decreasing in λ.

Thus, if ETFs pull liquidity trading away from the firm, w_1 falls, and so does the resulting value of information. Less noise trading means that the uninformed traders gain more information from observing price, which lowers the difference between the informed and uninformed posterior variances, and therefore lowers

[8] Admati (1985) provides a solution to the multi-asset noisy rational expectations equilibrium with a reasonably general information structure, but her model does not accommodate the case where information is only about one component of the asset's final payoff.

the value of being informed. Later, when we also require equilibrium in the market for information, we show that in the one-asset economy the equilibrium response is that the number of informed traders falls. These predictions are in line with the results in Israeli et al (2017).

The one-asset model has a payoff structure with common and idiosyncratic components, but no means for investors to trade on the common component. In the next section we add the ETF asset to the market and show that many of the one-asset results no longer hold.

Market with Two Risky Assets

Suppose that the firm asset, and information about it, remain the same; $F_1 = F_c + F_i$ and $Y_i = F_i + \varepsilon$ as before. Now add an ETF asset whose payoff is $F_2 = F_c$. The idea is that the ETF is composed of a sufficiently large number of component assets that, to a rough approximation, the idiosyncratic components of each firm in the ETF average out and contribute trivial residual variance to the payoff. It is easiest to think of the ETF as an industry ETF, although any ETF that isolates a common component of underlying firm payoffs will fit the model. We are interesting in studying how the value of information in this two-asset economy compares to the value of information in the one-asset economy.

Investors now need to form beliefs about two assets. The λ informed traders, upon observing the private signal Y_i, have the following posterior 2x1 mean vector and 2x2 covariance matrix for $[F_1, F_2]$:

$$E_I = \begin{bmatrix} \eta + \dfrac{hY_i}{h+s} \\ \eta \end{bmatrix} \ and\ V_I = \begin{bmatrix} g + \dfrac{hs}{h+s} & g \\ g & g \end{bmatrix}, with \det V_I = \frac{ghs}{h+s}. \tag{9}$$

As before, the $(1\text{-}\lambda)$ uninformed traders use prices to infer what they can about the informed traders' signal. Because there is no information to be had about F_c, we begin with the conjecture that the price of the second asset (i.e. the ETF) is uninformative. However, the price of the first asset (i.e. the firm) should still provide some information about Y_i. We conjecture that the price of the first asset is linear in the informed traders' signal and the realized supply of both assets:

$$P_1 = a_0 + a_1 Y_i - a_2 X_1 - a_3 X_2 \; and \tag{10}$$

$$\hat{P}_1 = Y_i - \left(\frac{a_2}{a_1}\right) X_1 - \left(\frac{a_3}{a_1}\right) X_2 = Y_i - Z_1 X_1 - Z_2 X_2. \tag{11}$$

While it might be tempting to assume Z_2 is zero, we show later that the equilibrium price for the firm asset also puts weight on the supply of the ETF asset. This makes sense because both risky assets offer exposure to the common component of the payoff, and so the noisy variation in both assets should be priced. For the two-asset economy, denote the variance of the noise in the price signal as

$$Q_{2a} = Var(\varepsilon - Z_1 X_1 - Z_2 X_2) = \; s + Z_1^2 w_1 + Z_2^2 w_2, \tag{12}$$

where the subscript '2a' denotes the two-asset economy. Assuming for the moment that Z_1 and Z_2 are known values (to be determined in equilibrium), the uninformed traders' posterior 2x1 mean vector and 2x2 covariance matrix are:

$$E_U = \begin{bmatrix} \eta + \dfrac{h\hat{P}_1}{(h + Q_{2a})} \\ \eta \end{bmatrix} \; and \; V_U = \begin{bmatrix} g + \dfrac{hQ_{2a}}{h + Q_{2a}} & g \\ g & g \end{bmatrix}, with \; detV_U = \frac{ghQ_{2a}}{h + Q_{2a}}. \tag{13}$$

To generalize the demand functions in (5), let E_I and E_U be the 2x1 vectors as given above, V_I and V_U be the 2x2 covariance matrices given above, and P be a 2x1 vector of P_1 and P_2. Plugging in these values into (5) gives the 2x1 demand functions for the informed and uninformed traders:

$$D_I = \rho \begin{bmatrix} \frac{Y_i}{s} + (P_2 - P_1)\left(\frac{h+s}{hs}\right) \\ \frac{\eta - P_2}{g} - \left[\frac{Y_i}{s} + (P_2 - P_1)\left(\frac{h+s}{hs}\right)\right] \end{bmatrix} \tag{14}$$

and

$$D_U = \rho \begin{bmatrix} \frac{\hat{P}_1}{Q_{2a}} + (P_2 - P_1)\left(\frac{h+Q_{2a}}{hQ_{2a}}\right) \\ \frac{\eta - P_2}{g} - \left[\frac{\hat{P}_1}{Q_{2a}} + (P_2 - P_1)\left(\frac{h+Q_{2a}}{hQ_{2a}}\right)\right] \end{bmatrix}. \tag{15}$$

These demand functions reveal an important change that happens to the economy when the EFT asset is introduced. Both informed and uninformed traders hedge their demands. To see this, note that the demand for the ETF asset (the bottom entry in each demand vector) perfectly removes the demand for the firm asset (the top entry in each demand vector). To make this completely clear, note that absent the firm asset, the demand for the ETF asset from informed and uninformed traders would be computed from (5) as

$$\rho\left(\frac{\eta - P_2}{g}\right). \tag{16}$$

Thus, the actual demand for the ETF asset is the amount in (16) less the demand for the firm asset given in the top entry of (14) and (15). So, if the trader purchases 50 units of the firm asset, he subtracts exactly 50 units from his demand for the ETF asset. The net result is a perfect hedge against the common component F_c in the firm's payoff.

This result is not a simple artefact of having two risky assets in the economy. If the firm payoff was changed to be only the idiosyncratic component F_i (rather than $F_c + F_i$) and the signal was $Y_i = F_i + \varepsilon$, then the demand for each asset would depend on only the parameters related to that asset. For example, computing the conditional means and variances for the informed traders in this alternative model, then applying (5), would result in

$$D_I with\ independent\ payoffs = \rho \begin{bmatrix} \frac{Y_i}{s} - \left(\frac{h+s}{hs}\right) P_1 \\ \frac{\eta - P_2}{g} \end{bmatrix}. \tag{17}$$

Note that, with independent payoffs, the demand for each asset depends only on that asset's parameters. As we will see later, by allowing traders the opportunity to hedge out exposure to the common component of the firm payoff, the existence of the ETF asset increases the value of the signal about the idiosyncratic component of the payoffs. However, the result is not immediate because, by changing the demand for the firm asset, the ETF also changes the informativeness of that asset's price. Thus, we need to develop the equilibrium prices in the two-asset market.

Returning to our model, the market clearing conditions for each asset require that

$$\lambda D_I + (1-\lambda) D_U = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}. \tag{18}$$

Solving for P_1 and P_2 gives

$$P_1 = \eta - \frac{g}{\rho}(X_1 + X_2) + \left[\frac{\frac{\lambda Y_i}{s} + \frac{(1-\lambda)\widehat{P_1}}{Q_{2a}} - \frac{X_1}{\rho}}{\lambda\left(\frac{h+s}{hs}\right) + (1-\lambda)\left(\frac{h+Q_{2a}}{hQ_{2a}}\right)} \right]\ and \tag{19}$$

$$P_2 = \eta - \frac{g}{\rho}(X_1 + X_2). \tag{20}$$

As expected, the price P_2 of the ETF is straightforward. It is the commonly-held expectation of the payoff η less the risk premium, where the risk premium is affected equally by the shocks X_1 and X_2. This makes sense because both assets come with exposure to the common factor F_c and so the supply of each contributes to the risk. P_1, the price of the firm, is complicated by the rational expectations conjecture. Note that the RHS of (19) is linear in Y_i, X_1, and X_2 (recalling that $\hat{P}_1 = Y_i - Z_1X_1 - Z_2X_2$). As shown in the appendix, equating the coefficients on Y_i, X_1, and X_2 on the RHS of (19) with the assumed linear coefficients a_1, a_2 and a_3 in (10) gives the equilibrium condition for Z_1 and Z_2 (recalling that Q_{2a} is also a function of these values).

Lemma 3 (with proof in Appendix)
Define $G = gh + gs + hs$ and $G' = \lambda gh + gs + hs$. In the two-asset economy there exist unique equilibrium values of Z_1 and Z_2 such that

$$Z_2 = Z_1 - \frac{s}{\lambda\rho} \; and \; Z_1 \in \left(\frac{G'}{\lambda\rho h}, \frac{G}{\lambda\rho h}\right). \tag{21}$$

Note that both Z_1 and Z_2 are strictly positive (as claimed earlier). An analytic solution for Z_1 and Z_2 is not available and simulations show that the solution depends on virtually every parameter in the model. Nonetheless, the range given in (21) is sufficiently precise to allow inferences about the equilibrium.

With the existence and an approximate range of Z_1 and Z_2 established, we are in a position to compute the value of the Y_i signal in the two-asset economy.

In the two–asset economy,

$$detV_{\mathrm{I}} = \frac{ghs}{h+s} \; and \; detV_U = \frac{ghQ_{2a}}{h+Q_{2a}} \tag{22}$$

so that

$$\Phi(2 \text{ asset}) = \left(\frac{\rho}{2}\right) log \left\{ \frac{\frac{ghQ_{2a}}{h+Q_{2a}}}{\frac{ghs}{h+s}} \right\}, where \; Q_{2a} = s + Z_1^2 w_1 + Z_2^2 w_2. \tag{23}$$

The form of the value of information function in the two-asset economy has an important difference from its form in the one-asset economy. In (23) we see that the g in the numerator and denominator cancel and so, effectively, the variance of the common factor only enters the computation through the noise in price Q_{2a}. In contrast, for the one-asset economy, g enters additively in both the numerator and denominator, as seen in (8), and then again in the price noise Q_{1a}. The additive component shifts the numerator and denominator equal amounts, making any difference between them smaller in the argument of the log function. In the one-asset economy this makes sense – the information is only about the idiosyncratic portion of the payoffs and so if the common component becomes more variable then the signal is relatively less valuable. Why doesn't this same phenomenon occur in the two-asset economy? As we saw earlier, in the two-asset economy investors hedge out their exposure to the common component and so the only way a change in the common factor variance can matter is if it influences the noise in price.

As in the one-asset economy, liquidity noise still impacts the value of information although the proof, given in the appendix, is considerably more complicated.

Lemma 4

The equilibrium noise in price Q_{2a} and the value of information Φ_{2a} in the two-asset economy are increasing in the firm's liquidity shock parameter w_1 and in ETF's liquidity shock parameter w_2. Further, both are more sensitive to w_1 than to w_2.

Not surprisingly, more price noise from either source makes being privately informed more valuable. The additional observation that the value of information is more sensitive to firm noise than to ETF noise means that if a unit of liquidity trading shifts from the firm asset to the ETF asset, the net effect will be to lower the price noise and, consequently, the value of information. Thus, the phenomenon noted in the one-asset economy, that the value of being privately informed falls if the price noise falls, carries over to the two-asset economy.[9] However, what this comparative static misses is that there is a structural shift in the value of information when traders are allowed to use the ETF asset as a hedge. This leads to the main result of the paper.

Theorem One. For fixed values of noise variance w_1 and w_2, the value of information is strictly greater in the two-asset economy than in the one-asset economy. That is, $\Phi(2 \text{ asset}) > \Phi(1 \text{ asset})$.

To begin, write the difference in the two expressions as

$$\Phi(2 \text{ asset}) - \Phi(1 \text{ asset}) = \left(\frac{\rho}{2}\right) log \left[\frac{\frac{ghQ_{2a}}{h+Q_{2a}}}{\frac{ghs}{h+s}}\right] - \left(\frac{\rho}{2}\right) log \left[\frac{g+\frac{hQ_{1a}}{h+Q_{1a}}}{g+\frac{hs}{h+s}}\right] \; where \tag{24}$$

$$Q_{1a} = s + Z^2 w_1 \; and \; Q_{2a} = s + Z_1^2 w_1 + Z_2^2 w_2.$$

This value is positive if and only if

$$\left[\frac{ghQ_{2a}}{h+Q_{2a}}\right] + \left[\frac{hQ_{2a}}{h+Q_{2a}}\right]\left[\frac{hs}{h+s}\right] > \left[\frac{ghs}{h+s}\right] + \left[\frac{hQ_{1a}}{h+Q_{1a}}\right]\left[\frac{hs}{h+s}\right]. \tag{25}$$

Recall from the one-asset economy that $Q_{1a} = s + Z^2 w_1$ and $Z = G/(\lambda\rho h)$ and so Q_{1a} is a known function of parameters. In contrast, Q_{2a} is an unknown function that depends on the equilibrium values of Z_1 and Z_2. Solving (25) for Q_{2a} and substituting in for Q_{1a} gives the condition necessary for the theorem to be true:

[9] As an alternative prediction, recall that Ben-David et al (2017) conjecture that high-frequency trading increases the variance of liquidity shocks to the ETF, w_2. Lemma 4 shows that this would increase the value of information in the two-asset economy.

$$Q_{2a} > s\left[\frac{G + (g + h)Z^2 w_1}{G + gZ^2 w_1}\right]. \qquad (26)$$

Without an analytic solution for Z_1 and Q_{2a}, the proof of Theorem One, given in the Appendix, has to rely on an indirect method. The basic idea is to consider both sides of (26) as a function of w_1 and then show that the LHS dominates the RHS at all levels of w_1. The result is illustrated in Figure 1. As the figure shows, when there is no price noise, the uninformed traders know everything the informed traders know, so the value of information in both economies is zero. However, as the price noise increases, the value of information in the two-asset economy increases faster. Eventually there is so much noise that the uninformed traders learn nothing from price; at this point the value of information in the two-asset economy asymptotes at a strictly higher level than the value of information in the one-asset economy.

Theorem One captures a feature of ETFs that has not been studied previously. The prior literature has focused on how the addition of an EFT draws liquidity away from the underlying security, and benefits an investor who is informed about the common factor. Both effects work against an investor who is informed about the idiosyncratic component of the asset's payoff – the very thing that financial statement analysis is best at. While these marginal forces are still at work in my model, Theorem One shows that there is a discrete increase in the value of information when the ETF asset is added to the market. By allowing investors to hedge out their exposure to the common variation, the value of their idiosyncratic information is greater. This result is not available in the prior literature because those models assumed risk neutral investors, and risk neutral investors do not hedge.

Equilibrium in the Market for Information

The results derived so far are for a fixed fraction λ of informed agents. Given sufficient friction in the labour market for financial analysts' skills, this is not an unreasonable assumption over horizons of a few years. However, if the value of information changes because of the introduction of ETFs, then it is reasonable to assume that, eventually, the fraction of agents who collect private information will also change.

Assume that an informed trader pays a fixed cost C to acquire the signal $Y_i = F_i + \varepsilon$, and $Var(\varepsilon) = s$, as before. Insofar as the change in noise/liquidity trading brought on by the EFT has no impact on the cost of becoming privately informed, a fixed cost structure is appropriate.

Definition: The competitive equilibrium in the market for private information is defined by the fraction λ^* that makes $\Phi(\lambda^*) = C$.

As discussed in the introduction, Israeli et al (2017) offer some mixed evidence that the number of analysts following a firm falls when there is an increase in ETF ownership of a firm's shares. Treating λ^* as a function of w_1 and totally differentiating $\Phi(\lambda^*) = C$, the prediction can be written as

$$\frac{\partial \Phi}{\partial w_1} = \frac{\partial \Phi}{\partial \lambda}\frac{\partial \lambda^*}{\partial w_1} + \frac{\partial \Phi}{\partial w_1}|fixed\ \lambda = \frac{\partial C}{\partial w_1} = 0. \tag{27}$$

In the one-asset economy, Lemma 2 shows that $\frac{\partial \Phi}{\partial \lambda} < 0\ and\ \frac{\partial \Phi}{\partial w_1}|fixed\ \lambda > 0$ and so, by (27), it must be that $\frac{\partial \lambda^*}{\partial w_1} > 0$.

In other words, if the noise in the firm's asset decreases (because liquidity trading is moving to the ETF asset), the equilibrium fraction of informed traders in the one-asset economy also decreases, with the consequence that the value of information remains constant. This is a standard result that can be found in Grossman and Stiglitz (1980).

This force remains in the two-asset economy, but there is a competing force created by the ETF asset. We can characterize the equilibrium in both markets as:

$$\Phi(2\text{ asset}, \lambda_{2a}^{*}) = \Phi(1\text{ asset}, \lambda_{1a}^{*}) = \mathrm{C}, \text{where} \tag{28}$$

λ_{2a}^{*} denotes the equilibrium fraction of informed traders in the two-asset market and λ_{1a}^{*} denotes the equilibrium fraction of informed traders in the one-asset market. Our second theorem compares the two fractions of informed traders.

Theorem Two. In equilibrium, the fraction of informed traders in the two-asset economy is greater than the fraction of informed traders in the one-asset economy. That is, $\lambda_{2a}^{*} > \lambda_{1a}^{*}$.

The theorem follows because Φ(1 asset) is decreasing in λ, as given in lemma 2, and $\Phi(2\text{ asset}) > \Phi(1\text{ asset})$ at all levels of λ, as shown in Theorem One. Thus, while the exact shape of Φ(2 asset) is unknown, it must lie above and to the right of the Φ(1 asset) function. Thus, at the equilibrium where both Φ(1 asset) and Φ(2 asset) equal C, the fraction of informed traders in the two-asset economy is greater than in the fraction of informed traders in the one-asset economy. That is, $\lambda_{2a}^{*} > \lambda_{1a}^{*}$. This result is illustrated in Figure 2. The figure shows that, regardless of the new equilibrium of informed traders that would occur in the single asset economy as a response to the decline in liquidity trading, once the EFT asset is taken into account, more traders choose to become informed.

SUMMARY AND EMPIRICAL IMPLICATIONS

The most practical implication of the model is that, to reap the highest value of information, the financial analyst needs to pair a recommendation based on his idiosyncratic information about the firm with a recommendation to hedge out the common component of value using the appropriate ETF. For example, through the use of financial statement analysis the analyst could conclude that Eldorado Gold Corporation is unusually efficient at mining gold. The analyst should then combine her recommendation that investors go long in Eldorado with a recommendation that they short the iShares Global Gold Index in order to remove exposure to mining industry factors that she is uninformed about (for example, the future price of gold).

The model also highlights the difference between studying the introduction of an ETF into the market and studying changes in the level of ETF ownership. Thus, the model is consistent with the marginal effects found in Israeli et al (2017); assuming that an increase in ownership of an existing ETF causes a reduction in liquidity/noise trading in the underlying firms, then the value of information will drop and the fraction of informed traders will fall. At the same time, the prediction of Theorems One and Two are that, in a pre/post research design, the introduction of an ETF should increase the value of information and the fraction of informed traders.

The model also highlights the importance of identifying the full impact of the ETF on liquidity trading. Theorem One says that the value of information increases discretely, holding the noise parameters constant. But once we allow the noise parameters to change as well, the predictions are more complicated. Once the ETF asset exists, Lemma 4 describes the marginal effects of changing the noise parameters. If liquidity/noise shifts from the firm asset to the ETF asset, but the total remains constant, then the net marginal effect is to lower the value of information. Alternatively, if the introduction of an ETF draws in a new source of liquidity/noise trading in the ETF, the lemma says that the value of information will increase.

Finally, Theorem Two highlights the importance of the researcher's assumption about the equilibrium conditions of all markets. If, in addition to equilibrium in the asset markets, the researcher believes that the market for analysts is also in equilibrium, then the empirical focus should be on finding the counter-balancing empirical forces that maintain the equilibrium. For instance, as the fraction of informed traders increases following the introduction of an ETF, there should be a compensating increase in the noise in the price signal, such that the value of information remains equal to its cost. Alternatively, one could argue that frictions in the labour market for analysts support predictions based on a fixed fraction of informed traders.

Figure 1

The value of information is higher in the two-asset market than the one-asset market for all levels of liquidity noise
$\Phi(2\text{ asset}) > \Phi(1\text{ asset})$ for all w_1

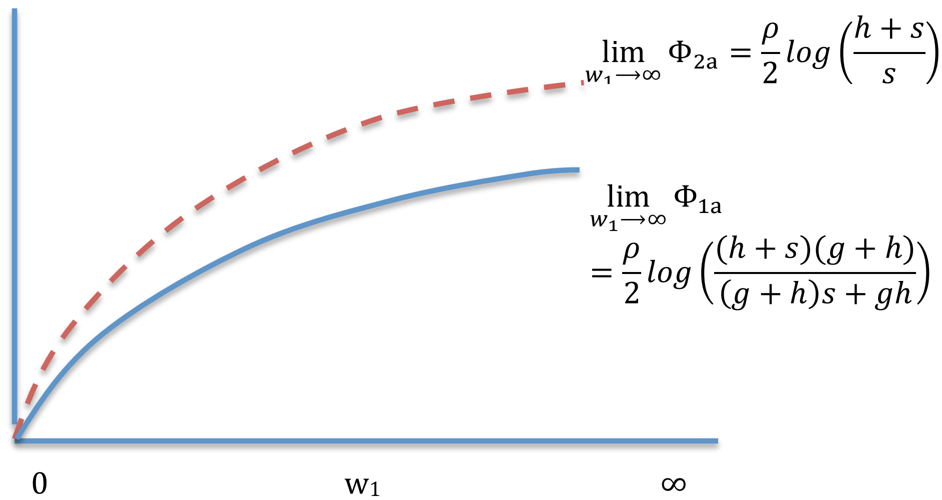
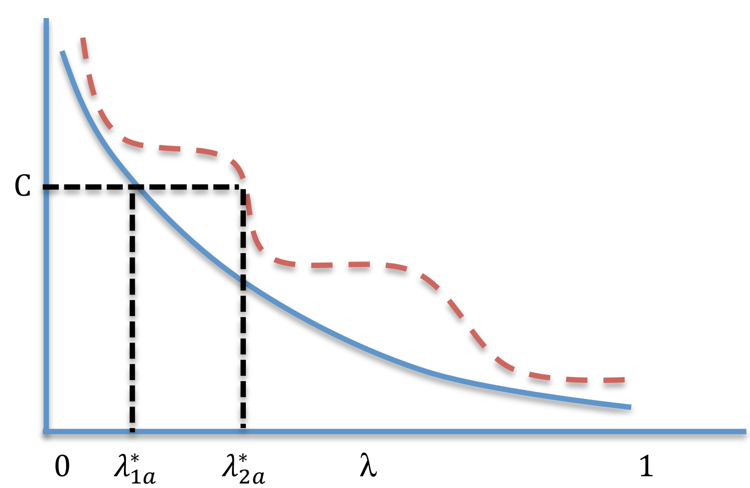


Figure 2

The equilibrium fraction of informed traders is higher in the two-asset market than the one asset market
$\Phi(2\text{ asset}, \lambda^*_{2a}) = \Phi(1\text{ asset}, \lambda^*_{1a}) = \text{C and } \lambda^*_{2a} > \lambda^*_{1a}$



APPENDIX

Derivation of equilibrium price and value of information in one-asset market:

Lemma 1

Substitute in the mean and variance values from (2) and (4) into the demand functions in (5), then substitute these demands into the market clearing condition in (6), then solve for *P* to get

$$P = \left\{ \lambda \left[\eta + \frac{hY_i}{h+s} \right] \left[g + \frac{hQ_{1a}}{h+Q_{1a}} \right] + (1-\lambda) \left[\eta + \frac{h(Y_i - ZX_1)}{h+Q_{1a}} \right] \left[g + \frac{hs}{h+s} \right] \right.$$
$$\left. - \frac{\left(g + \frac{hs}{h+s} \right)\left(g + \frac{hQ_{1a}}{h+Q_{1a}} \right) X_1}{\rho} \right\} \tag{A1}$$
$$\div \left\{ \lambda \left(g + \frac{hs}{h+s} \right) + (1-\lambda) \left(g + \frac{hQ_{1a}}{h+Q_{1a}} \right) \right\}.$$

Note that (A1) is linear in Y_i and X_1. The equilibrium price is determined by equating the coefficients in (A1) with the coefficients in our conjectured price, $P = a_0 + a_1 Y_i - a_2 X_1$. More precisely, recalling that $Z = a_2/a_1$, we need to equate the ratio of coefficients on X_1 to Y_i in (A1) to Z. Defining $G = gh + gs + hs$, this gives

$$Z = \frac{(1-\lambda)hZG + G(gh + gQ_{1a} + hQ_{1a})/\rho}{\lambda h(gh + gQ_{1a} + hQ_{1a}) + (1-\lambda)hG}. \tag{A2}$$

Substitute in $Q_{1a} = s + Z^2 w_1$ from (4) and simplify to get

$$Z^3 - Z^2\left[\frac{G}{\lambda\rho h}\right] + Z\left[\frac{G}{(g+h)w_1}\right] - \frac{G^2}{\lambda\rho h(g+h)w_1} = 0. \tag{A3}$$

The solution to this cubic defines the equilibrium Z, and thus the equilibrium coefficients in the price equation. Factoring the cubic yields the unique solution

$$Z = \left(\frac{a_2}{a_1}\right) = \frac{G}{\lambda\rho h}. \tag{A4}$$

<u>Derivation of equilibrium prices and value of information in two-asset market</u>

Picking up from equations (19) and (20) in the main body of the paper, market clearing gives us

$$P_2 = \eta - \frac{g}{\rho}(X_1 + X_2)\ and$$

$$P_1 = \eta - \frac{g}{\rho}(X_1 + X_2) + \left[\frac{\frac{\lambda Y_i}{s} + \frac{(1-\lambda)\widehat{P_1}}{Q_{2a}} - \frac{X_1}{\rho}}{\lambda\left(\frac{h+s}{hs}\right) + (1-\lambda)\left(\frac{h+Q_{2a}}{hQ_{2a}}\right)}\right] \tag{A5}$$

Recall that $\hat{P}_1 = Y_i - \left(\frac{a_2}{a_1}\right)X_1 - \left(\frac{a_3}{a_1}\right)X_2 = Y_i - Z_1X_1 - Z_2X_2\ and$

$$Q_{2a} = Var(\epsilon - Z_1X_1 - Z_2X_2) = (s + Z_1^2w_1 + Z_2^2w_2). \tag{A6}$$

The RHS of P_1 is linear in Y_i, X_1 and X_2. Denoting the denominator D as

$$D = \lambda\left(\frac{h+s}{hs}\right) + (1-\lambda)\left(\frac{h+Q_{2a}}{hQ_{2a}}\right), \tag{A7}$$

the coefficients in P_1 are as follows:

$$on\ Y_i: \frac{\frac{\lambda}{s} + \frac{(1-\lambda)}{Q_{2a}}}{D}, \tag{A8}$$

$$on - X_1: \frac{g}{\rho} + \left[\frac{\frac{1}{\rho} + \frac{(1-\lambda)Z_1}{Q_{2a}}}{D}\right], and \tag{A9}$$

$$on - X_2: \frac{g}{\rho} + \left[\frac{\frac{(1-\lambda)Z_2}{Q_{2a}}}{D}\right]. \tag{A10}$$

Equating the ratio of coefficients in $P_1 = a_0 + a_1 Y_i - a_2 X_1 - a_3 X_2$ given above with their assumed values Z_1 and Z_2 gives

$$\frac{a_2}{a_1} = Z_1 = \frac{\frac{gD}{\rho} + \frac{1}{\rho} + \frac{(1-\lambda)Z_1}{Q_{2a}}}{\frac{\lambda}{s} + \frac{(1-\lambda)}{Q_{2a}}}\ and \tag{A11}$$

$$\frac{a_3}{a_1} = Z_2 = \frac{\frac{gD}{\rho} + \frac{(1-\lambda)Z_2}{Q_{2a}}}{\frac{\lambda}{s} + \frac{(1-\lambda)}{Q_{2a}}}. \tag{A12}$$

The equilibrium price is determined by the solution to the system given in (A11) and (A12). To reduce this to one equation, define *G = gh + gs + hs* and *G' = λgh + gs + hs*. (A11) and (A12) can then be written as

$$Q_{2a}(Z_1 \lambda \rho h - G') - (1-\lambda) ghs = 0\ and \tag{A13}$$

$$Q_{2a}(Z_2\lambda\rho h - G' + hs) - (1-\lambda)ghs = 0. \tag{A14}$$

Equating the two equations gives $Z_2 = Z_1 - s/\lambda\rho$. By substituting $Z_2 = Z_1 - s/\lambda\rho$ into (A13), we can reduce the system to a cubic in Z_1. In particular, the equilibrium is determined by the root to the following equation:

$$\lambda\rho h(w_1 + w_2)Z_1^3 - [G'(w_1 + w_2) + 2hsw_2]Z_1^2 \tag{A15}$$

$$+\left\{\lambda\rho h\left[s + \left(\frac{s}{\lambda\rho}\right)^2 w_2\right] + G'\left(\frac{2s}{\lambda\rho}\right)w_2\right\}Z_1 - G'\left[s + \left(\frac{s}{\lambda\rho}\right)^2 w_2\right] - (1-\lambda)ghs = 0.$$

An analytic solution is not readily available, and simulations show that the root depends on every parameter in the model. However, the solution lies in a known range. Substituting $G'/\lambda\rho h$ for Z_1 in (A15) gives a strictly negative result and substituting $G/\lambda\rho h$ for Z_1 gives a strictly positive result. Thus, any solution to the cubic resides in the range ($G'/\lambda\rho h$, $G/\lambda\rho h$). As a check, note that setting $g = 0$, so that $G' = G = hs$, gives $Z_1 = s/\lambda\rho$ as a root to (A15). This is the same as the solution to the cubic given in (A4) when $g = 0$ for the one-asset market.

Further, the solution to the cubic in(A15) is unique. To see this, note from (A13) that Q_{2a} and $(\lambda\rho h Z_1 - G')$ are both strictly positive, both are increasing in Z_1, and their product must equal the positive constant $(1-\lambda)ghs$ at a root to the equation. Suppose a second solution Z_1' exists that is greater than the root Z_1. This would increase both parts of the product but not change the constant and so could not constitute a solution to the equation; similarly, a candidate second solution Z_1' that is less than Z_1 would reduce both parts of the product and not change the constant and so also could not constitute a solution. Consequently, there is only one solution.

Other properties of the equilibrium Z_1

We need to establish some properties of the equilibrium that will be used later to prove Theorem One. First, viewing the cubic as an implicit function that defines Z_1, we can compute the first derivatives of Z_1 with respect to w_1 and w_2. Differentiate both sides of (A15) with respect to w_2 to get

$$\lambda\rho h w_1 3Z_1^2\frac{\partial Z_1}{\partial w_2}+\lambda\rho h\left(w_2 3Z_1^2\frac{\partial Z_1}{\partial w_2}+Z_1^3\right)-G'w_1 2Z_1\frac{\partial Z_1}{\partial w_2}-(G'+2sh)\left(w_2 2Z_1\frac{\partial Z_1}{\partial w_2}+Z_1^2\right)$$

$$+\lambda\rho hs\frac{\partial Z_1}{\partial w_2}+(2G'+sh)\frac{sh}{\lambda\rho h}\left(w_2\frac{\partial Z_1}{\partial w_2}+Z_1\right)-G'\left(\frac{sh}{\lambda\rho h}\right)^2=0 \tag{A16}$$

Solving for $\frac{\partial Z_1}{\partial w_2}$ gives

$$\frac{\partial Z_1}{\partial w_2}=\frac{-\left(Z_1-\frac{sh}{\lambda\rho h}\right)^2(\lambda\rho h Z_1-G')}{D_1}\ where \tag{A17}$$

$$D_1=Z_1w_1(3\lambda\rho h Z_1-2G')+\lambda\rho hs$$

$$+\frac{w_2}{\lambda\rho h}\left[2(\lambda\rho h Z_1-G')(\lambda\rho h Z_1-sh)+(\lambda\rho h Z_1-sh)^2\right].$$

The lower bound on Z_1 implies that $(\lambda\rho h Z_1-G')$ is strictly positive and so the numerator of (A17) is negative. Further, the denominator D_1 is positive. To see this note that $G' > sh$ and so the lower bound also implies that $(\lambda\rho h Z_1-sh)$ is also positive. Thus, $\frac{\partial Z_1}{\partial w_2}<0$.

Next differentiate the cubic in (A15) with respect to w_1 to get

$$\lambda\rho h w_2 3Z_1^2 \frac{\partial Z_1}{\partial w_1} + \lambda\rho h\left(w_1 3Z_1^2 \frac{\partial Z_1}{\partial w_1} + Z_1^3\right) - w_2(G' + 2sh)2Z_1 \frac{\partial Z_1}{\partial w_1} - G'\left(w_1 2Z_1 \frac{\partial Z_1}{\partial w_1} + Z_1^2\right)$$

$$+\left[\lambda\rho hs + \lambda\rho h w_2\left(\frac{sh}{\lambda\rho h}\right)^2 + G' w_2\left(\frac{2sh}{\lambda\rho h}\right)\right] \frac{\partial Z_1}{\partial w_1} = 0 \qquad \text{(A18)}$$

Solving for $\frac{\partial Z_1}{\partial w_1}$ gives

$$\frac{\partial Z_1}{\partial w_1} = \frac{-Z_1^2(\lambda\rho h Z_1 - G')}{D_1} \textit{ where } D_1 \textit{ is positive as given above.} \qquad \text{(A19)}$$

The same argument as above yields $\frac{\partial Z_1}{\partial w_1} < 0$.

Finally, we need the second derivative with respect to w_1 when w_2 is set to zero. Evaluating $\frac{\partial Z_1}{\partial w_1}$ at $w_2 = 0$ gives

$$\frac{\partial Z_1}{\partial w_1} = \frac{-Z_1^2(\lambda\rho h Z_1 - G')}{Z_1 w_1(3\lambda\rho h Z_1 - 2G') + \lambda\rho hs}. \qquad \text{(A20)}$$

Differentiating (A20) by w_1 gives

$$\frac{\partial^2 Z_1}{\partial w_1} = \left\{\begin{matrix} [w_1 Z_1(3\lambda\rho h Z_1 - 2G') + \lambda\rho hs] * \left[-3\lambda\rho h Z_1^2 \frac{\partial Z_1}{\partial w_1} + 2G' Z_1 \frac{\partial Z_1}{\partial w_1}\right] \\ - [Z_1^2(\lambda\rho h Z_1 - G')] * \left[3\lambda\rho h\left(2w_1 Z_1 \frac{\partial Z_1}{\partial w_1} + Z_1^2\right) - 2G'\left(w_1 Z_1 \frac{\partial Z_1}{\partial w_1} + Z_1\right)\right] \end{matrix}\right\} \qquad \text{(A21)}$$

$$\div [Z_1 w_1(3\lambda\rho h Z_1 - 2G') + \lambda\rho hs]^2.$$

The sign of (A21) is determined by the numerator. Substituting in (A19) for $\frac{\partial Z_1}{\partial w_1}$ gives

$$numerator = 2Z_1^3(\lambda\rho h Z_1 - G')(3\lambda\rho h Z_1 - 2G') - \frac{Z_1^3(\lambda\rho h Z_1 - G')^2(6\lambda\rho h Z_1 - 2G')}{(3\lambda\rho h Z_1 - 2G') + \lambda\rho hs}. \qquad \text{(A22)}$$

The $\lambda\rho hs$ term in the denominator of the second term makes this negative value smaller. Ignoring the $\lambda\rho hs$ term and some simplification gives

$$numerator = \frac{2Z_1^3(\lambda\rho h Z_1 - G')\,[3(2\lambda\rho h Z_1 - G')(\lambda\rho h Z_1 - G') + \lambda\rho h Z_1 G']}{(3\lambda\rho h Z_1 - 2G')} > 0, \qquad \text{(A23)}$$

so that $\frac{\partial^2 Z_1}{\partial w_1} > 0$.

The final properties of Z_1 we need are, with $w_2 = 0$, where it lies when $w_1 = 0$ and where it converges. Setting $w_2 = w_1 = 0$ in (A15) gives

$$Z_1(w_1 = 0) = \frac{G}{\lambda\rho h}. \qquad \text{(A24)}$$

Taking the limit with respect to w_1 is a bit more complicated, but the key is to remember that, even though Z_1 is unknown and depends on w_1, it is bound in ($G'/\lambda\rho h$, $G/\lambda\rho h$). Therefore none of the Z_1 terms go to infinity. With $w_2 = 0$, taking the limit of both sides of the equality in (A15) gives

$$\lim_{w_1\to\infty} Z_1(w_1) = \lim[w_1 Z_1^2(\lambda\rho h Z_1 - G')] = 0, \textit{which occurs at } \frac{G'}{\lambda\rho h}. \qquad \text{(A25)}$$

Proof of Lemma 4

The sensitivity of the value of information to liquidity noise is driven by the sensitivity of price noise Q_{2a} to liquidity noise. Differentiating Q_{2a} with respect to w_1 gives

$$\frac{\partial Q_{2a}}{\partial w_1} = 2Z_1 \frac{\partial Z_1}{\partial w_1} w_1 + Z_1^2 + 2Z_2 \frac{\partial Z_2}{\partial w_1} w_2. \tag{A26}$$

Because $Z_2 = Z_1 - \left(\frac{s}{\rho h}\right)$, the derivatives of Z_2 with respect to w_1 and w_2 are the same as the derivatives for Z_1, as given in (A17) and (A19). Substituting in $\frac{\partial Z_1}{\partial w_1}$ for $\frac{\partial Z_2}{\partial w_1}$ in (A26) and then substituting in (A19) for $\frac{\partial Z_1}{\partial w_1}$ gives

$$\frac{\partial Q_{2a}}{\partial w_1} = Z_1^2 \left[1 - \frac{2(\lambda\rho h Z_1 - G')(Z_1 w_1 + Z_2 w_2)}{D_1}\right], \tag{A27}$$

where D_1 is given in (A17) as

$$D_1 = Z_1 w_1 (3\lambda\rho h Z_1 - 2G') + \lambda\rho h s$$

$$+ \frac{w_2}{\lambda\rho h} [2(\lambda\rho h Z_1 - G')(\lambda\rho h Z_1 - sh) + (\lambda\rho h Z_1 - sh)^2].$$

The bracketed term in (A27) is positive because the fraction in the bracketed term is less than one. To see this, note that

$$Z_1 w_1 (3\lambda\rho h Z_1 - 2G') > 2(\lambda\rho h Z_1 - G') Z_1 w_1 \text{ and} \tag{A28}$$

$$2(\lambda\rho h Z_1 - G')(\lambda\rho h Z_1 - sh)\frac{w_2}{\lambda\rho h} > 2(\lambda\rho h Z_1 - G') Z_2 w_2 \tag{A29}$$

and so $\frac{\partial Q_{2a}}{\partial w_1} > 0$. A parallel argument establishes $\frac{\partial Q_{2a}}{\partial w_2} > 0$.

Comparing the two derivatives gives

$$\frac{\partial Q_{2a}}{\partial w_1} - \frac{\partial Q_{2a}}{\partial w_2} = (Z_1 - Z_2)(Z_1 + Z_2)\left[1 - \frac{2(\lambda\rho h Z_1 - G')(Z_1 w_1 + Z_2 w_2)}{D_1}\right] > 0 \qquad \text{(A30)}$$

where the bracketed term is positive, as shown above.

<u>Proof of Theorem One.</u>

Equation (26) in the text gives the following condition for the theorem to hold:

$$Q_{2a} > s\left[\frac{G + (g+h)Z^2 w_1}{G + gZ^2 w_1}\right]. \qquad \text{(A31)}$$

To bring the equilibrium condition into play, use (A13) to write

$$Q_{2a} = \frac{(1-\lambda)ghs}{(Z_1\lambda\rho h - G')} \qquad \text{(A32)}$$

.

and then solve for Z_1 to get the following condition imposed on the equilibrium Z_1 for the theorem to be true.

$$Z_1 < \frac{G'}{\lambda\rho h} + \frac{(1-\lambda)gh}{\lambda\rho h}\left[\frac{G + gZ^2 w_1}{G + (g+h)Z^2 w_1}\right]. \qquad \text{(A33)}$$

The value $G'/\lambda\rho h$ is the lower bound on Z_1 and the value $(1\text{-}\lambda)gh/\lambda\rho h$ is the difference between the upper bound and the lower bound of Z_1. So, for example, if $w_1 = 0$, the bracket term equals one, the RHS is the upper bound on Z_1 and the condition for the theorem is met. The general proof involves treating both sides of (A33) as a function of w_1 and showing that the RHS is greater than the LHS at all values of w_1.

Start with the RHS of (A33). From (A4) substitute in $Z= G/\lambda\rho h$ from the one-asset and evaluate at $w_1 = 0$ to get the RHS equals $G/\lambda\rho h$. The derivative of the RHS with respect to w_1, evaluated at $w_1=0$, is

$$\frac{\partial RHS}{\partial w_1} = \frac{-(1-\lambda)gh^2G}{(\lambda\rho h)^3}. \tag{A34}$$

Finally, the limit of the RHS as w_1 goes to infinity is

$$\frac{G'}{\lambda\rho h} + \frac{(1-\lambda)gh}{\lambda\rho h}\left[\frac{g}{g+h}\right]. \tag{A35}$$

In sum, the RHS starts at $G/\lambda\rho h$ and decreases monotonically to the limit given in (A35).

Now consider the LHS of (A33); the equilibrium Z_1. From (A17) above, $\frac{\partial Z_1}{\partial w_2} < 0$ at all values of w_1. Because we are looking for a Z_1 lower than a bound, without loss of generality we can consider the maximal value of Z_1 by setting $w_2=0$. With this, we see from (A24) that $Z_1(w_1 = 0) = G/\lambda\rho h$; the same starting point as the RHS. Next, (A19) gives $\frac{\partial Z_1}{\partial w_1}$. Evaluating this at $w_1 = w_2 = 0$, and substituting in $Z_1= G/\lambda\rho h$ gives

$$\frac{\partial Z_1}{\partial w_1} = \frac{-(1-\lambda)ghG^2}{(\lambda\rho h)^3 s}. \tag{A36}$$

Comparing the derivatives of the RHS and LHS with respect to w_1 evaluated at 0, as given in (A34) and (A36), shows that the LHS decreases from zero at a greater rate. Next, as w_1 goes to infinity, the LHS goes to $G'/\lambda\rho h$, as given in (A25), which is strictly lower than the limit of the RHS as given in (A35). Thus, the LHS starts at the same place as the RHS, then decreases faster and limits out lower. Finally, (A23) shows that the second derivative of Z_1 is positive everywhere. This rules out the possibility that the Z_1 function flattens out and crosses the RHS condition before finally limiting to a lower value.

REFERENCES

Admati, A. 1985. A noisy rational expectations equilibrium for multi-asset securities markets. Econometrica 53 (May): 629-57.

Admati, A. and P. Pfleiderer. 1987. Viable Allocations of Information in Financial Markets. Journal of Economic Theory 43:76-115.

Alles, M. and R. Lundholm. 1993. On the optimality of public signals in the presence of private information. The Accounting Review 66(1):93-112.

Bhattacharya, A., and M. O'Hara. 2017. Can ETFs increase market fragility? Effect of Information linkages in ETF markets. City University of New York Working Paper.

Ben-David, I., F. Franzoni, and R. Moussawi. Do ETFs Increase Volatility? 2017. Ohio State University working paper.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi. 2017. Exchange traded funds, *Annual Review of Financial Economics* 9:165–185.

Black, F. (1986). Noise. *Journal of Finance* 41:529–543.

BlackRock Global ETP Landscape Monthly Snapshot. September 2018.

Bhojraj, S., P. Mohanram and S. Zhang. ETFs and Information Transfer Across Firms. Cornell University 2018 working paper.

Cong, L., and D. Xu. 2016. Rise of factor investing: Asset prices, informational efficiency, and security design. University of Chicago working paper.

Gorton, G. and G. Pennacchi. 1993. Security baskets and index-linked securities. *The Journal of Business* 66(1):1-27.

Grossman, S., and J. Stiglitz. 1980. On the Impossibility of Informationally Efficient Markets. *American Economic Review* 70: 393-408.

Hamm, S. 2014. The Effect of ETFs on Stock Liquidity. Ohio State University Working Paper.

Israeli, D., C. Lee, and S. Sridharan. 2017. Is there a Dark Side to Exchange Traded Funds (ETFs)? An Information Perspective. *Review of Accounting Studies* 22 (3): 1048–1083.

Kyle, A. 1985. Continuous auctions and insier trading. *Econometrica* 53 (November): 1315-35.

Lettau, M. and A. Madhavan Exchange-Traded Funds 101 for Economists. 2018. *Journal of Economic Perspectives* 32(1): 135–154.

Li, Frank Weikai, and Qifei Zhu. 2016. Synthetic Shorting with ETFs. University of Texas at Austin Working Paper.

Marshall B., N. Nguyen, and N. Visaltanachoti. 2015. ETF Liquidity, Massey University Working Paper.

Subrahmanyam. A. 1991. A theory of trading in stock index futures. *Review of Financial Studies* 4(1):17-51.

Welch, W. 2014. Properties of the Normal and Multivariate Normal Distributions. Class Note. https://www.stat.ubc.ca/~will/cx/private/normal_properties.pdf

Wigglesworth, R. ETF growth is 'in danger of devouring capitalism.' Financial Times. February 4, 2018.