

Moral Hazard, Informed Trading, and Stock Prices

Pierre Collin-Dufresne

EPFL, SFI and NBER

Vyacheslav Fos

University of Illinois at Urbana-Champaign

October 31, 2013

Abstract

We analyze a model of informed trading where an activist shareholder accumulates shares in an anonymous market and then expends costly effort to increase the firm value. We find that equilibrium prices are affected by the position accumulated by the activist, because the level of effort undertaken is increasing in the size of his acquired position. In equilibrium, price impact has two components: one due to asymmetric information (as in the seminal Kyle (1985) model) and one due to moral hazard (a new source of illiquidity). Price impact is higher the more severe the moral hazard problem, which corresponds to a more productive activist. We thus obtain a trade-off: with more noise trading (less ‘price efficiency’) the activist can build up a larger stake, which leads to more effort expenditure and higher firm value (more ‘economic efficiency’).

Keywords: Informed trading, asymmetric information, liquidity, moral hazard, shareholder activism, price efficiency, continuous time.

Email addresses: pierre.collin-dufresne@epfl.ch (Pierre Collin-Dufresne),
vfos@illinois.edu (Vyacheslav Fos)

1. Introduction

Activist shareholders play an active role in modern corporate governance. Among the most active and visible activists are outside shareholders who identify a firm with the potential for value creation, purchase a significant number of shares in the open market and then publicly announce their intention to influence management.¹ The empirical literature suggests that these activists are often successful in increasing the value of targeted companies (e.g., Brav, Jiang, Partnoy, and Thomas, 2008). The key ingredient to the success of outside activists is their ability to purchase shares in the open market before stock prices reflect their intention to become active, and therefore to increase firm value. The value created presumably depends on the activist's effort expenditure, which in turn depends on the size of the stake that was acquired during the trading period. The larger the stake the higher the incentives for the insider to provide additional effort. Thus, there is a fundamental link between market conditions that influence the block acquisition process, the activist's final position, the activist's effort expenditure, and the firm's value.

In his seminal contribution, Kyle (1985) derives the equilibrium price dynamics in a model where a large trader possesses long-lived private information about the value of a stock that will be revealed at some known date, and optimally trades into the stock to maximize his expected profits. Risk-neutral market makers try to infer from

¹Activist hedge funds constitute the core of this group. Carl Icahn, who manages several activist hedge funds, is one prime example of such outside activist shareholder. Activist shareholders also include long-term shareholders who own large stakes of a company, monitor management, and influence management in various ways. Pension funds, mutual funds, and insurance companies are among prominent representatives of this group.

aggregate order flow the information possessed by the insider. Because order flow is also driven by uninformed ‘noise traders,’ who trade solely for liquidity purposes, prices are not fully revealing. Instead, prices respond linearly to order flow. In Kyle’s model the insider has private information about a value that is independent of his actions. In other words, after the start of the trading game the insider cannot affect the liquidation value. Whether or not the insider chooses to trade in the stock, this value will be realized.

In this paper we generalize Kyle’s model (in the continuous time formulation given by Back, 1992) of informed trading to an activist who can affect the liquidation value of the firm by expending effort at some cost. We solve for the optimal effort level and trading strategy of the activist, as well as for the equilibrium price and corresponding market liquidity. The key feature of the model is that the activist’s optimal effort level is increasing in the size of the stake he has accumulated. The endogenous liquidation value of the firm is thus a function of the position accumulated by the insider, which itself depends on the noise trading activity (market liquidity). The model thus delivers an intricate relation between price and market liquidity.

Unlike in the original Kyle/Back model where the position of the insider is irrelevant for the equilibrium price in our setting the market maker’s estimate of (and uncertainty about) the position of the activist affects the equilibrium price. Therefore, price dynamics are more complex than in the standard Kyle/Back model.

First, in equilibrium, price impact has two components: an asymmetric information component due to the fact that the activist has better information about the exogenous component of the liquidation value (his ‘stock picking’ ability similar to Kyle’s model) and a moral hazard component due to the fact that the position accumulated by the

activist will affect the endogenous component of the liquidation value through his effort level (his ‘share-holder activism’ ability). While total price impact (measured by the response of price to order flow) is constant in our model, each component fluctuates over time.

In the earlier stage of the game the dominant part of the price impact is typically due to the asymmetric information component. In contrast, closer to the end of the trading period the dominant part of the price impact is due to the moral hazard component. The intuition is that the uncertainty about the position of the activist, which governs the moral hazard component of price impact, tends to grow over time in the early part of the trading game, whereas the uncertainty about the exogenous component of the liquidation value always decreases over time.

Second, in equilibrium price impact is higher the more severe the moral hazard problem, which corresponds to cases where the activist is more productive (i.e., can more influence the terminal value), or where the uncertainty about his position is larger. This is a source of price impact that has not, to the best of our knowledge, been studied in previous models.²

Third, the uncertainty about the source of value creation develops in a very special way. If there is no uncertainty about the exogenous component of firm value and all adverse selection comes from moral hazard, then the market will learn perfectly the value created by the activist. If however, there is uncertainty about the exogenous component

²For example, in Kyle-Back’s original model the prior distribution of the initial position of the insider is irrelevant to the equilibrium given a prior for the exogenous component of the liquidation value. In contrast, in our paper information about the activist’s position is directly relevant for market liquidity. This also has implications for disclosure regulations which we discuss further below.

of firm value, then there is some ‘signal jamming’: the market has difficulty separating its estimate of the exogenous and endogenous components of firm value. This seems important given the debate about the social ‘usefulness’ of activists. Are they simply better stock-pickers or do they really create value for minority shareholders by expending effort? The model suggests that this is very difficult for the market to sort out, if there is uncertainty about the activists holdings.

Fourth, the model shows how ownership disclosure regulation can affect economic efficiency. The model suggests that if we can force the activist to disclose his holdings, i.e. reduce uncertainty about his position, then this makes it easier for the market to sort out his ‘stock-picking’ from his ‘managerial’ abilities. From that perspective, if it is desirable to remunerate activists only for their efforts, ownership disclosure requirements may be useful. Of course, this may also have unintended consequences for economic efficiency, as it leads them to accumulate fewer shares and thus expend less effort on average. In addition, the model informs the debate about the optimal duration of the pre-disclosure period (e.g., Bebchuk, Brav, Jackson, and Jiang, 2013). The model shows that shortening the period during which the activist can trade anonymously may have negative consequences for economic efficiency, as it leads the activist to accumulate fewer shares and therefore expend less effort.

Finally, we obtain a trade-off between economic efficiency and price efficiency. With more noise trading the activist can build up a larger stake, and then expend more effort to increase firm value, thus leading to higher economic efficiency.

The paper concludes with a description of some empirical regularities, which are consistent with the proposed model. Using hand-collected data on trades by activist

shareholders (see Collin-Dufresne and Fos, 2013a, for detailed description of the dataset), we show that an activist’s trading strategy depends not only on the ‘valuation gap’, but also on his stake size. This is in contrast to the Kyle (1985) model, in which the trading strategy of the informed trader depends on the ‘valuation gap’ only. We also show that the activist’s stake size is positively associated with the value creation for shareholders of targeted company, conditional on the total share turnover during the accumulation period. Finally, when we analyze the characteristics of targeted companies, we find evidence consistent with the proposed model. For example Schedule 13D filers are more likely to target firms with low stock illiquidity, high idiosyncratic volatility, and high proportion of shares owned by activist hedge funds. In addition, when compared to non-targeted companies, targeted firms are older, they have lower market capitalization, higher book-to-market ratio, and higher proportion of shares owned by institutional investors. This evidence is consistent with evidence on shareholder activism (e.g., Brav et al., 2008; Fos, 2012).

Related Literature

This paper is related to several strands of literature.

First, the paper contributes to the microstructure literature. To the best of our knowledge, we are the first to endogenize the terminal firm value. For example, in Kyle (1985) the terminal firm value is exogenous. Similarly, the literature has maintained the assumption of an exogenous terminal firm value (e.g., Glosten and Milgrom, 1985; Easley and O’Hara, 1987; Back, 1992).

Second, the paper contributes to the corporate governance literature which inves-

investigates the role of blockholders in monitoring the management.³ To the best of our knowledge, there is no dynamic model which incorporates anonymous trading by an informed investor who can endogenously change the firm value. Shleifer and Vishny (1986) analyze the role of a large minority shareholder in solving the Grossman and Hart (1980a) free-rider problem and show that the large shareholder's return on his own shares suffices to cover his monitoring and takeover costs. In their static model firm value is affected by effort expenditure by the large shareholder, which is increasing in the large shareholder's stake. Admati, Pfleiderer, and Zechner (1994) develop a static model in which a large shareholder has access to a costly monitoring technology affecting securities' expected payoffs. In their model the large shareholder trades off the benefits (more profits from monitoring) and costs (more losses from risk-sharing) of owning a large stake. Similarly to Shleifer and Vishny (1986), monitoring by the large shareholder is costly and is more efficient the higher his stake in the company.

The common feature of these models is that the large shareholder cannot increase his stake by trading anonymously. This is the main innovation of our paper, in which the activist can trade in anonymous markets to change his stake size. Moreover, our model is dynamic.⁴

Third, the paper contributes to the literature that studies the role of noise trading

³See Edmans (2013) for survey of blockholders and corporate governance literature.

⁴An exception is Noe (2002), who develops a static model in which strategic investors produce private information through their own action. Monitoring induces a fixed private cost to the investor. The microstructure set up of the model is special, as the market maker does not condition quotes on aggregate order flow (see his discussion on page 311). The main reason for making this simplifying assumption is that there is no need to update market maker's prior beliefs, which is non-trivial when the terminal value of stock is endogenous. While in our dynamic model the terminal value is endogenous, we are able to solve the model when the market maker conditions prices on the order flow.

as a solution to the Grossman and Hart (1980a) free-rider problem. Kyle and Vila (1991) develop a static model in which noise trading provides camouflage that helps an outside shareholder to purchase enough shares at favorable prices so that takeovers become profitable. The main result is that noise trading has a positive impact on profits of large shareholder but has undetermined effects on economic efficiency. Specifically, noise trading encourages fewer (more) takeovers when takeovers would otherwise always (never) occur. However, unlike in our paper, Kyle and Vila (1991) assume that the takeover premium the large shareholder must pay to takeover the firm is exogenous and that the initial stake of the large shareholder is exogenously given, which essentially avoids the moral hazard problem we study. Maug (1998) endogenizes the initial stake of the large shareholder in a static model and shows that market liquidity mitigates the free-rider problem by allowing the informed activist shareholder to purchase shares at a discount from uninformed shareholders. If the activist intervenes, firm value increases by an exogenously given amount. The main innovation of our model is that we study the dynamic relation between noise trading, share accumulation and effort expenditure by the activist.

Fourth, the paper contributes to the literature that studies price informativeness. In the seminal paper of Grossman and Stiglitz (1980), market participants trade off costs and benefits of becoming informed. In equilibrium, prices depend on the relative weight of informed traders. Firm value, however, is exogenously determined and is not affected by traders' decision to become informed. Our paper suggests a trade off between economic efficiency (i.e., effort expenditure by activist shareholder) and price efficiency. We show that noise traders contribute to economic efficiency by increasing

the activist's optimal stake and therefore effort expenditure but have a negative impact on price efficiency by keeping prices from converging to their terminal value.

Fifth, the paper is related to the literature that studies optimal disclosure (e.g., Grossman and Hart, 1980b). Our paper shows that information on activists' positions can be valuable and significantly affect market liquidity, price efficiency and economic efficiency. Indeed, in our model more disclosure about activists positions may improve market liquidity (i.e., reduce price impact) at the risk of lowering economic efficiency. There is some evidence in the empirical literature that can be interpreted in light of this prediction. For example, Agarwal, Jiang, Tang, and Yang (2013) show that institutional investors often ask for a delayed reporting of quarter-end equity holdings. The authors provide several pieces of empirical evidence that support private information as the predominant motive for this activity.

Finally, there is robust empirical evidence, consistent with our model's predictions and our own empirical findings, that activists tend to target (all else equal) more liquid firms (e.g., Brav et al., 2008; Norli et al., 2010; Fos, 2012; Collin-Dufresne and Fos, 2013a).

2. Informed Trading with Hidden Action

Our model is based on the Back (1992) continuous time version of the Kyle (1985) model.⁵ The main new feature of our model is that we assume the activist trading in the stock can choose to exert effort w paying a cost $C(w) = \frac{w^2}{2\psi}$ to produce the terminal

⁵We use continuous time because of its simplicity, but in the appendix derive the discrete time one period model for completeness.

(liquidation) value of $v + w$ where v is a constant known only to him and w is a choice variable. The parameter ψ captures the productivity of the activist.⁶ We assume that the Market has an initial prior about $v \sim N(V_0, \sigma_v^2 T)$. For simplicity of interpretation we focus on the case where $v > V_0$ and the activist typically accumulates a positive number of shares and chooses a positive effort level w (though the setting is symmetric and allows, in principle, the activist to short the stock and furnish additional effort to drive the value more negative). It is natural to think of this activist as a hedge-fund activist for example.

The second important departure from the original Kyle-Back (KB) model is that we assume the activist may start with an initial position X_0 which is known only imperfectly by the Market. Indeed we assume the Market has an initial Gaussian prior $X_0 \sim N(Q_0, \sigma_X^2 T)$ which is correlated with its estimate of v and we denote their covariance by $\Sigma_{Xv}(0) = \sigma_{Xv} T$.⁷ This allows us to interpret the initial date 0 as the date at which the activist becomes informed or as the date at which the market becomes aware of the existence of the activist.⁸ It also allows us to investigate the costs and benefits of disclosure rules for activist shareholders (which can be interpreted here as reducing σ_X^2). We note that the original KB model can easily be extended to random initial endowment in stocks by the activist, since the latter plays no role in the equilibrium and indeed,

⁶The higher ψ the more productive the activist. Note that if $\psi \rightarrow 0$ it becomes optimal to choose $w = 0$ and the problem becomes identical to the original Kyle model.

⁷The only technical requirement we impose is that the prior Covariance matrix be positive definite.

⁸Of course, this is not fully consistent with rational expectation, in that such an interpretation ignores the fact that the fully rational market maker should have been aware prior to date 0 of the possibility of such information becoming available to the activist. Modeling that is possible along the lines of ?, but we leave such an extension for future work.

the equilibrium price and market liquidity are unaffected by it. This underlines the difference between both models.

The activist is risk-neutral and maximizes his expected terminal profit:

$$\max_{\theta_t \in \mathcal{A}, w} \mathbb{E} \left[\int_0^T (v + w - P_t) \theta_t dt + (v + w - \bar{P}_0) X_0 - C(w) \mid \mathcal{F}_t^Y, v, X_0, \bar{P}_0 \right], \quad (1)$$

where we denote by \mathcal{F}_t^Y the information filtration generated by observing the entire past history of aggregate order flow $\{Y_s\}_{s \leq t}$ and where \bar{P}_0 denotes the average price the activist paid for his initial stake X_0 . We note that unlike in the original Kyle-Back model, it is important to condition on the initial position of the activist shareholder, which plays an important role in our setting. For simplicity we assume that the discount rate is zero. Timing wise, we assume that the activist chooses his effort level right ‘after’ the terminal date T , given all past information on prices and trades.⁹ In particular, his optimal choice will simply maximize at T :

$$\max_w w X_T - C(w), \quad (2)$$

where $X_T = X_0 + \int_0^T \theta_t dt$ is the activist’s initial stake plus the stake accumulated during the trading game by the activist. With our choice of cost function this leads to:

$$w^* = \psi X_T. \quad (3)$$

⁹A richer model might allow the activist to work continuously during the share accumulation phase on the project that eventually will lead to the liquidation value. We leave such a model for future work.

Plugging back into his objective function we see that the activist is maximizing:

$$(v - \overline{P_0})X_0 + \max_{\theta_t \in \mathcal{A}} \mathbb{E} \left[\int_0^T (v - P_t)\theta_t dt + \frac{\psi X_T^2}{2} \mid \mathcal{F}_t^Y, v, X_0 \right]. \quad (4)$$

Following Back (1992) we assume that the activist chooses a trading rule θ in some admissible set \mathcal{A} defined to be the set of absolutely continuous trading strategies which satisfy the technical restriction that $\mathbb{E}[\int_0^T |\theta_s|^2 ds] < \infty$.¹⁰

The market maker is risk-neutral, but does not observe the terminal value. Instead (given his prior) he only observes the aggregate order flow, which is the sum of informed and uninformed order flow:

$$dY_t = \theta_t dt + \sigma dZ_t, \quad (5)$$

where Z_t is a standard Brownian motion independent of v . We assume that the uninformed order flow volatility, σ , is constant.¹¹

To solve for an equilibrium, we proceed as follows. First, we derive the dynamics of the stock price consistent with the market maker's risk-neutral filtering rule, conditional on a conjectured trading rule for the activist. Then we solve the activist's optimal portfolio choice problem, given the assumed dynamics of the equilibrium price. Finally, we show that the conjectured rule by the market maker is indeed consistent with the

¹⁰A shown in Back, it is optimal for the activist to choose an absolutely continuous trading strategy, since, in continuous time, the market maker can immediately infer from the quadratic variation of the order flow the informed component with infinite variation. The square integrability condition is a technical requirement often used in continuous time to rule out specific arbitrage strategies such as 'doubling strategies' (see Harrison and Pliska (1979) and Dybvig and Huang (1986)).

¹¹We could easily generalize idiosyncratic volatility to be a function of posterior co-variance $\sigma(\Sigma_t, t)$ as in Baruch (2002) (though at the expense of closed-form solutions). The extension to arbitrary stochastic processes as in Collin-Dufresne and Fos (2013b) is non-trivial and left for future research.

activist's optimal choice.

Since the market maker is risk-neutral, equilibrium imposes that

$$P_t = \mathbb{E} [v + w | \mathcal{F}_t^Y]. \quad (6)$$

Since the market maker is rational he knows the optimal choice of effort by the activist will be a function of his acquired stake, i.e, that $w = \psi X_T$. The market maker will thus filter the position of the activist. We define his estimate of the activist's position as:

$$Q_t = \mathbb{E} [X_t | \mathcal{F}_t^Y] \quad (7)$$

and his estimate of the constant component as:

$$V_t = \mathbb{E} [v | \mathcal{F}_t^Y]. \quad (8)$$

It follows from the definition of conditional expectation that:

$$P_t = \mathbb{E} [v + \psi X_T | \mathcal{F}_t^Y] = V_t + \psi \mathbb{E} [Q_T | \mathcal{F}_t^Y]. \quad (9)$$

We conjecture that the trading strategy of the activist will be linear in his 'valuation gap' as well as his position gap:

$$\theta_t = \beta_t(v - V_t) + \gamma_t(X_t - Q_t), \quad (10)$$

where β_t measures the speed at which the activist decides to close the gap between his

assessment of the fundamental value v (known only to him) and the market maker's estimate V_t and where γ_t is the loading on the activist's position estimation error made by the market maker.

The novel feature relative to most of the literature is that the activist does not only trade because of his valuation gap. Instead, his trading decision is also motivated by the accumulated position. Indeed the price has two components to its value: one component that is independent of the action of the activist, and another that depends on his choice of effort, which itself depends on the position he has accumulated. Thus the price depends in equilibrium on the market's estimate of the activist's position in the stock. His accumulated position has dynamics:

$$dX_t = \theta_t dt. \tag{11}$$

Now, given his conjecture about the trading dynamics of the activist, the market maker's optimal filtered price dynamics follows from standard results in the filtering literature (e.g., Liptser and Shiryaev, 2001, Chapter 12). The novel feature, is that the market maker will estimate both the fundamental value v and the position of the activist X_t from the observed aggregate order flow. And, in equilibrium, the price dynamics will be multi-variate Markov.

Let's denote the posterior covariance matrix of the filtered state $S_t = [v; X_t]'$ by Σ_t ((2, 2) matrix):

$$\Sigma(t) = Var[S_t | \mathcal{F}_t^Y]. \tag{12}$$

Note that:

$$M_t = \mathbb{E}[S_t | \mathcal{F}_t^Y] = [V_t; Q_t]$$

and for simplicity we introduce the notation:

$$\Sigma_v(t) = \mathbb{E} [(v - V_t)^2 | \mathcal{F}_t^Y] \equiv \Sigma_{11}(t) \quad (13)$$

$$\Sigma_X(t) = \mathbb{E} [(X_t - Q_t)^2 | \mathcal{F}_t^Y] \equiv \Sigma_{22}(t) \quad (14)$$

$$\Sigma_{Xv}(t) = \mathbb{E} [(v - V_t)X_t | \mathcal{F}_t^Y] \equiv \Sigma_{12}(t), \quad (15)$$

with initial conditions $\Sigma_v(0) = \sigma_v^2 T$, $\Sigma_X(0) = \sigma_x^2 T$, and $\Sigma_{Xv}(0) = \sigma_{Xv} T$. $\Sigma_v(0)$ measures the prior uncertainty about the exogenous component of private information available to the activist, $\Sigma_X(0)$ measures the prior uncertainty about the activist's position, and $\Sigma_{Xv}(0)$ measures their covariance.

A direct application of known results on conditionally Gaussian filtering gives the following Lemma.

Lemma 1. *If the activist adopts a trading strategy of the form given in (10), then the stock price and the filtered position of the activist given by equations (9) and (7) satisfy:*

$$dV_t = \lambda_t dY_t \quad (16)$$

$$dQ_t = \Lambda_t dY_t, \quad (17)$$

where λ_t and Λ_t satisfy:

$$\lambda_t = \frac{\beta_t \Sigma_v(t) + \gamma_t \Sigma_{Xv}(t)}{\sigma^2} \quad (18)$$

$$\Lambda_t = \frac{\beta_t \Sigma_{Xv}(t) + \gamma_t \Sigma_X(t)}{\sigma^2}. \quad (19)$$

Further, the dynamics of the posterior covariance matrix is given by:

$$\begin{aligned} d\Sigma_v(t) &= -\lambda_t^2 \sigma^2 dt & (20) \\ d\Sigma_X(t) &= \Lambda_t (2 - \Lambda_t) \sigma^2 dt \\ d\Sigma_{Xv}(t) &= \lambda_t (1 - \Lambda_t) \sigma^2 dt \end{aligned}$$

Proof 1. *This follows directly from an application of theorems 12.6, 12.7 in Liptser and Shiryaev (2001). We provide a simple ‘heuristic’ motivation of the result using the Gaussian projection theorem in the appendix.*

We now try to solve the activist’s partial equilibrium problem taking as given price dynamics:

$$dV_t = \lambda_t (\theta_t dt + \sigma dZ_t) \quad (21)$$

$$dQ_t = \Lambda_t (\theta_t dt + \sigma dZ_t) \quad (22)$$

$$dX_t = \theta_t dt. \quad (23)$$

Note that if the activist indeed follows the conjectured trading rule, then both V_t and Q_t are martingales in the market maker’s filtration. It follows then from the definition of the equilibrium price in equation (9), that the price can be rewritten simply as:

$$P_t = V_t + \psi Q_t \quad (24)$$

and that its dynamics are:

$$dP_t = (\lambda_t + \psi\Lambda_t)dY_t. \quad (25)$$

Following the intuition from Kyle-Back's original model, we may conjecture that in equilibrium the total price impact $\lambda_t + \psi\Lambda_t$ will be constant. Indeed, since the activist is risk-neutral, he would otherwise seek to concentrate all his trading in periods with the lowest total price impact. We shall thus seek an equilibrium with this property. Further, we conjecture that in equilibrium the posterior variance of the price should converge to zero, since otherwise the risk-neutral activist could change his trading strategy to take advantage of positive expected return trades. We first prove that there exists a trading strategy that leads to such an outcome.

Lemma 2. *Suppose that the activist chooses his trading strategy as conjectured in (10), with*

$$\gamma_t = \psi\beta_t \quad \text{and} \quad \beta_t = \frac{\Delta\sigma^2}{\Omega_t}, \quad (26)$$

where

$$\Omega_t = \Sigma_v(t) + 2\psi\Sigma_{Xv}(t) + \psi^2\Sigma_X(t). \quad (27)$$

Then the total price impact due to Bayesian updating is constant:

$$\lambda_t + \psi\Lambda_t = \Delta \quad (28)$$

and

$$\Omega_t = \Omega_0 + (2\psi - \Delta)\Delta\sigma^2t. \quad (29)$$

Further, there exists a positive $\hat{\Delta}$ such that $\Omega_T = 0$ given by:

$$\hat{\Delta} = \psi + \sqrt{\psi^2 + \alpha^2} \quad (30)$$

where α is the ‘signal to noise ratio’ defined as

$$\alpha = \frac{\omega}{\sigma}, \quad (31)$$

where we define the annualized quantity of initial private information $\omega^2 = \frac{\Omega_0}{T} = \sigma_v^2 + 2\psi\sigma_{Xv} + \psi^2\sigma_X^2$. For this choice of $\hat{\Delta}$ the expression for Ω_t simplifies:

$$\Omega_t = \omega^2(T - t). \quad (32)$$

Proof 2. Suppose the activist adopts the conjectured trading strategy. Then using equations (18) and (19) we immediately obtain:

$$\lambda_t + \psi\Lambda_t = \Delta \quad \forall t.$$

Further, using the dynamics of the covariance matrix in (20) we find that when the activist follows such a strategy:

$$d\Omega_t = (2\psi - \Delta)\Delta\sigma^2 dt.$$

It follows that $\Omega_t = \Omega_0 + (2\psi - \Delta)\Delta\sigma^2 t$, and thus the equation $\Omega_T = 0$ admits two roots for Δ one of which is positive and given by $\hat{\Delta}$ in the Lemma.

By definition $\Omega_t = \text{Var}[v + \psi X_t | \mathcal{F}_t^Y]$. For the conjectured equilibrium Ω_t given in

(32) decays linearly, which is reminiscent of the dynamics of the posterior variance of the estimated liquidation value in the original Kyle (1985) model. Note however that the dynamics of Ω_t is affected by the prior uncertainty about the position of the activist, which is irrelevant in the original Kyle model (as can be verified by taking the limit $\psi \rightarrow 0$).

Indeed, the total price impact $\hat{\Delta}$ obtained in this model is always greater than that obtained in a model without moral hazard (i.e., when $\psi = 0$) in which case it becomes identical to price impact obtained in the KB model $\hat{\Delta}|_{\psi=0} = \frac{\sigma v}{\sigma}$. In fact, price-impact in our model depends on only two quantities: the productivity of the activist ψ and the signal to noise ratio $\frac{w}{\sigma}$. It is increasing in both. Interestingly the signal to noise ratio relevant for the moral hazard model depends on the prior uncertainty about the activist's position (σ_X, σ_{Xv}) which is irrelevant in the KB model.

Further, note that Ω_t is not equal to the posterior variance of the liquidation value (the latter is $Var[v + \psi X_T | \mathcal{F}_t^Y]$). Instead, Ω_t can be interpreted as the variance of the liquidation value of the stock if it were liquidated at the present time t . Indeed, Ω_t converges at maturity to the variance of the terminal stock price liquidation value. In fact, we have that $\Omega_T = E[(v + \psi X_T - P_T)^2 | \mathcal{F}_T^Y]$. And since the previous result shows that (for the conjectured equilibrium) $\Omega_T = 0$, this suggests that we should obtain a convergence result for the equilibrium price also in the filtration of the activist. Indeed, we can prove the following.

Lemma 3. *Suppose the trading strategy followed by the activist is as described in Lemma 2. Then, in the filtration \mathcal{F}_t of the activist, the equilibrium price process is*

two-factor Markov in state variables P_t and X_t and is given by:

$$dP_t = \frac{(\hat{\Delta}\sigma)^2}{\omega^2(T-t)}(v + \psi X_t - P_t)dt + \hat{\Delta}\sigma dZ_t \quad (33)$$

$$dX_t = \frac{(\hat{\Delta}\sigma)^2}{\omega^2(T-t)}(v + \psi X_t - P_t)dt. \quad (34)$$

In the filtration of the activist, the price process P_t converges in L^2 to $v + \psi X_T$ at maturity. Note that in the filtration of the market maker, the stock price P_t is a Brownian martingale.

Proof 3. By definition $P_t = V_t + \psi Q_t$. Thus:

$$dP_t = \hat{\Delta}dY_t \quad (35)$$

$$= \hat{\Delta}\beta_t((v - V_t) + \psi(X_t - Q_t))dt + \hat{\Delta}\sigma dZ_t \quad (36)$$

$$= \frac{(\hat{\Delta}\sigma)^2}{\omega^2(T-t)}(v + \psi X_t - P_t)dt + \hat{\Delta}\sigma dZ_t. \quad (37)$$

Of course, in the filtration of the market maker price is a martingale, since $E[dY_t | \mathcal{F}_t^Y] = 0$, and $d[Y, Y]_t = \sigma^2 dt$ and $dP_t = \hat{\Delta}dY_t$.

Now, define $h_t = P_t - v - \psi X_t$. We want to show that h_t converges to zero. Its dynamics are:

$$dh_t = dP_t - \psi dX_t \quad (38)$$

$$= -\frac{1 + \kappa}{T - t}h_t dt + \hat{\Delta}\sigma dZ_t, \quad (39)$$

where:

$$\kappa = \zeta^2 + \zeta\sqrt{1 + \zeta^2} > 0 \quad (40)$$

$$\zeta = \frac{\psi\sigma}{\omega}. \quad (41)$$

We can thus compute

$$h_t = h_0 e^{-A_t} + e^{-A_t} M_t, \quad (42)$$

where $A_t = \int_0^t \frac{(1+\kappa)}{T-u} du = \log\left(\frac{T}{T-t}\right)^{1+\kappa}$ and $M_t = \int_0^t e^{-A_s} \hat{\Delta}\sigma dZ_s$ is an \mathcal{F}_t -adapted Brownian martingale. We can calculate the quadratic variation of M_t from

$$\langle M \rangle_t = \int_0^t e^{-2A_s} (\hat{\Delta}\sigma)^2 ds \quad (43)$$

$$= e^{2A_t} \Omega_t - \Omega_0 \quad (44)$$

Thus for any $t < T$ we see that M_t is a square integrable martingale and that $\lim_{t \rightarrow T} \langle M \rangle_t = \infty$. It follows that $\mathbb{E}[h_t] = h_0 e^{-A_t} \forall t < T$ and that since $\kappa > 0$ we obtain $\lim_{t \rightarrow T} \mathbb{E}[h_t] = 0$. Further, for any $t < T$ we have $\mathbb{E}[h_t^2] = h_0^2 e^{-2A_t} + e^{-2A_t} \langle M \rangle_t = (h_0^2 - \Omega_0) e^{-2A_t} + \Omega_t$. Since $\Omega_T = 0$ and $\kappa > 0$ we have $\lim_{t \rightarrow T} \mathbb{E}[h_t^2] = 0$. This establishes L^2 convergence of h_t to 0 at T .

Remark 1. The previous lemma implies that $(T-t)^\kappa (P_t - v - \psi X_t)$ converges almost surely to zero at T with κ defined in equation (40). Indeed, note that there exists an \mathcal{F}_t adapted brownian motion B_t such that $M_t = B_{\langle M \rangle_t}$. Further, by the Strong law of large numbers for Brownian motions (Karatzas and Shreve (1988) p. 104) we have $\lim_{u \rightarrow \infty} \frac{B_u}{u} = 0$ a.s.. Combining we see that $\lim_{t \rightarrow T} e^{-2A_t} M_t = \lim_{t \rightarrow T} \frac{B_{\langle M \rangle_t}}{\langle M \rangle_t} e^{-2A_t} \langle M \rangle_t = 0$ a.s.. It follows that $\left(\frac{T-t}{T}\right)^{1+\kappa} h_t = h_0 e^{-2A_t} + e^{-2A_t} M_t$ converges almost surely to 0 at time T . We note that if $\psi = 0$ then $\kappa = 0$. So this proves almost sure convergence of the price to the terminal value v if there is no Moral Hazard. We conjecture (but have not yet been able to prove) that $P_t - v - \psi X_t$ converges a.s. to zero also when $\psi > 0$.

We now turn to the optimal policy of the activist. We want to show that given the conjectured equilibrium price impact process, the strategy conjectured by the market maker is indeed a best response for the activist. Note that the expected profits of the activist given in equation (4) can be rewritten as follows:¹²

$$(v - \bar{P}_0)X_0 + \psi \frac{X_0^2}{2} + \max_{\theta_t \in \mathcal{A}} \mathbb{E} \left[\int_0^T (v - P_t + \psi X_t) \theta_t dt \mid \mathcal{F}_t^Y, v, X_0 \right]$$

Since the first two terms do not affect his choice¹³ we define the following value function that captures the optimization problem of the insider:

$$J(t) = \max_{\theta_s \in \mathcal{A}} \mathbb{E} \left[\int_t^T (v - P_s + \psi X_s) \theta_s ds \mid \mathcal{F}_t^Y, v, X_t \right]. \quad (45)$$

We show the following.

Proposition 1. *Suppose that prices have dynamics*

$$dP_t = \Delta dY_t \quad (46)$$

for some constant Δ . Suppose further that there exists an admissible trading strategy θ^* such that $\mathbb{E}[(P_T - v - \psi X_T)^2] = 0$, then θ^* is an optimal trading strategy and the optimal

¹²This follows from the fact that $\frac{X_T^2}{2} = \frac{X_0^2}{2} + \int_0^T X_t dX_t = \frac{X_0^2}{2} + \int_0^T X_t \theta_t dt$.

¹³Note that the sum of these two terms is equal to the expected profit of the insider were he to not accumulate any additional share after time 0 and then provide the optimal effort at maturity (given that $X_T = X_0$).

value function is given by:

$$J(P, X, t) = \frac{(P - v - \psi X_t)^2 + \Delta^2 \sigma^2 (T - t)}{2(\Delta - \psi)}. \quad (47)$$

Proof 4. Consider the function

$$J(P, X, t) = \frac{(P - v - \psi X_t)^2 + \Delta^2 \sigma^2 (T - t)}{2(\Delta - \psi)}. \quad (48)$$

Applying Itô's lemma we find:

$$\frac{(P_T - v - \psi X_T)^2}{2(\Delta - \psi)} - J(P_0, X_0, 0) + \int_0^T (v + \psi X_t - P_t) \theta_t dt = \int_0^T \frac{(P - v - \psi X)}{(\Delta - \psi)} \Delta \sigma dZ_t.$$

First, note that the right hand side is a square integrable martingale for any admissible trading strategy. Indeed, note that any admissible trading strategy satisfies $\mathbb{E}[\int_0^T X_t^2 dt] < \infty$. (To see this simply note that by the Cauchy-Schwartz inequality $X_t^2 = (\int_0^t \theta_s ds)^2 \leq \int_0^t \theta_s^2 ds$.) Then note that $P_t = \Delta(X_t - X_0) + \Delta \sigma(Z_t - Z_0)$ and thus $\mathbb{E}[\int_0^T (P_t - v - \psi X_t)^2 dt] < \infty$ for any admissible trading strategy.

Thus, taking expectations it follows that for any admissible trading strategy θ_t we have:

$$J(P_0, X_0, 0) = \mathbb{E} \left[\frac{(P_T - v - \psi X_T)^2}{2(\Delta - \psi)} + \int_0^T (v + \psi X_t - P_t) \theta_t dt \right]. \quad (49)$$

Since $\mathbb{E}[(P_T - v - \psi X_T)^2] \geq 0$, it follows that for any admissible θ_t we have

$$J(P_0, X_0, 0) \geq \mathbb{E} \left[\int_0^T (v + \psi X_t - P_t) \theta_t dt \right]. \quad (50)$$

And if there exists an admissible trading strategy θ_t^* such that $\mathbb{E}[(P_T - v - \psi X_T)^2] = 0$ then the (weak) inequality holds with equality. This establishes the optimality of such a

trading strategy θ_t^* and of the value function.

Combining the verification theorem with the previous lemmas we have the main result of this paper.

Proposition 2. *There exists an equilibrium characterized by deterministic functions λ_t and Λ_t such that total price impact is constant:*

$$\lambda_t + \psi\Lambda_t = \hat{\Delta} := \psi + \sqrt{\psi^2 + \frac{\omega^2}{\sigma^2}} \quad \forall t, \quad (51)$$

where $\omega^2 = \sigma_v^2 + 2\psi\sigma_{vX} + \psi^2\sigma_X^2$. The optimal trading strategy for the activist is:

$$\theta_t^* = \frac{\hat{\Delta}\sigma^2}{\Omega_t}(v + \psi X_t - P_t), \quad (52)$$

where $\Omega_t = \omega^2(T - t)$. The value function of the activist is:

$$J(P, X, t) = \frac{(P - v - \psi X_t)^2 + \hat{\Delta}^2\sigma^2(T - t)}{2(\hat{\Delta} - \psi)}. \quad (53)$$

The equilibrium is revealing in that P_t converges to $v + \psi X_t$ at time T .

Proof 5. *Follows immediately from previous results. The admissibility of θ^* is easily checked.*

We note that while the dynamics of Ω_t are simple, the separate dynamics of $\Sigma_v(t)$, $\Sigma_X(t)$, and $\Sigma_{Xv}(t)$ are less obvious. Similarly, while the total price impact deriving from both asymmetric information and moral hazard is constant: $\lambda_t + \psi\Lambda_t = \hat{\Delta}$, the individual components λ_t and Λ_t are not. We can however characterize these analytically in closed-form, which provides further insight into the equilibrium.

Lemma 4. *In equilibrium, $\lambda_t = \lambda_0 \left(\frac{T-t}{T}\right)^\kappa$ with $\lambda_0 = \frac{\hat{\Delta}}{\omega^2}(\sigma_v^2 + \psi\sigma_{Xv})$ and κ defined in equation (40). Thus if $\lambda_0 = \sigma_v^2 + \psi\sigma_{Xv} > 0$ (resp. < 0) then λ_t is positive (resp. negative) and strictly decreasing (resp. increasing) and $\lim_{t \rightarrow T} \lambda_t = 0$.*

It follows that if $\lambda_0 > 0$ (resp. < 0) then $\Lambda_t = \frac{\hat{\Delta} - \lambda_t}{\psi}$ is strictly increasing (resp. decreasing) and that $\lim_{t \rightarrow T} \Lambda_t = \hat{\Delta}$.

Further, we can solve the system of ODE for all the posterior covariance matrix in closed-form.

$$\Sigma_v(t) = \sigma_v^2 T + \frac{\lambda_0^2 \sigma^2 \left((T-t) \left(\frac{T-t}{T}\right)^{2\kappa} - T \right)}{2\kappa + 1} \quad (54)$$

$$\Sigma_{Xv}(t) = \sigma_{Xv} T - \frac{\Sigma_v(t) - \sigma_v^2 T}{\psi} + \frac{\sqrt{\frac{1}{\zeta^2} + 1} \lambda_0 \sigma^2 (T-t) \left(\frac{T-t}{T}\right)^\kappa}{\kappa + 1} - \frac{\sqrt{\frac{1}{\zeta^2} + 1} \lambda_0 \sigma^2 T}{\kappa + 1} \quad (55)$$

$$\Sigma_X(t) = (\Omega(t) - 2\psi\Sigma_{Xv}(t) - \psi^2\Sigma_v(t)) / \psi^2 \quad (56)$$

The results show that if either $\sigma_X = 0$ or $\sigma_v = 0$, then $\lim_{t \rightarrow T} \Sigma_v(t) = \lim_{t \rightarrow T} \Sigma_X(t) = 0$. In other words, if there is prior uncertainty about only one of the two sources of uncertainty (position or exogenous component of asset payoff) then the equilibrium is fully revealing about both the position and the terminal value of the asset (in that both $V_T = v$ and $Q_T = X_T$ at maturity). Instead, if there is prior uncertainty about both sources (both σ_X and σ_v are greater than zero), then the equilibrium reveals the total payoff $v + \psi X_T$ but not the individual components (both $\Sigma_X(T) > 0$ and $\Sigma_v(T) > 0$). In particular, the market cannot infer perfectly the effort expended by the activist in that case.

Proof 6. *From its definition $\lambda_t = \frac{\hat{\Delta}}{\Omega_t}(\Sigma_v(t) + \psi\Sigma_{Xv}(t))$. Differentiating and using the dynamics of the covariance matrix we obtain $\dot{\lambda}_t = \frac{\lambda_t \sigma^2}{\Omega_t} \left(\frac{\omega^2}{\sigma^2} - \hat{\Delta}(\hat{\Delta} - \psi) \right) = -\frac{\kappa}{\lambda_t}$, where κ has been defined previously. This ODE is easily solved for λ_t given its initial condition.*

The results on Λ_t follow from the fact that $\Lambda_t = \hat{\Delta} - \psi\lambda_t$.

Given the solutions for λ_t, Λ_t the covariance matrix ODE can be solved explicitly given their initial conditions.

We can then check the terminal value $\Sigma_v(T)$ and $\Sigma_X(T)$ and observe that $\Sigma_v(T) = 0 \Leftrightarrow \Sigma_X(T) = 0 \Leftrightarrow \{\sigma_v = 0 \text{ or } \sigma_X = 0\}$. This follows immediately from the closed-form solution and using the fact that at any time t we have $\Sigma_{Xv}(t) = 0 \Leftrightarrow \{\Sigma_v(t) = 0 \text{ or } \Sigma_X(t) = 0\}$.

We next provide some pictures of the dynamics of these variables for specific parameter choices. We fix $T = \sigma = 1$, and $\psi = 1$ and consider three cases for the initial two sources of adverse selection:

1. Initial position is known: $\sigma_X^2 = 0, \sigma_v^2 = 1$.
2. Exogenous component is known: $\sigma_X^2 = 1, \sigma_v^2 = 0$.
3. Both are unknown: $\sigma_X^2 = 0.5, \sigma_v^2 = 0.5$.

In all cases, we set $\sigma_{Xv} = 0$ so that the total signal to noise ratio $\frac{\omega}{\sigma}$ remains unchanged in all three cases.

Figure 1 shows how the private information available to the activist is revealed through his trading activity over the trading period. $\Sigma_v(t)$ measures the uncertainty about the exogenous component of private information available to the activist. $\Sigma_X(t)$ measures the uncertainty about the activist's position, i.e., the moral hazard component of adverse selection in our model. $\Sigma_{Xv}(t)$ measures their dependence.

The figure shows that, even though total uncertainty $\Omega(t)$ behaves exactly the same across all three scenarios and whence prices behave similarly, the type of information that gets into prices is very different depending on the initial conditions.

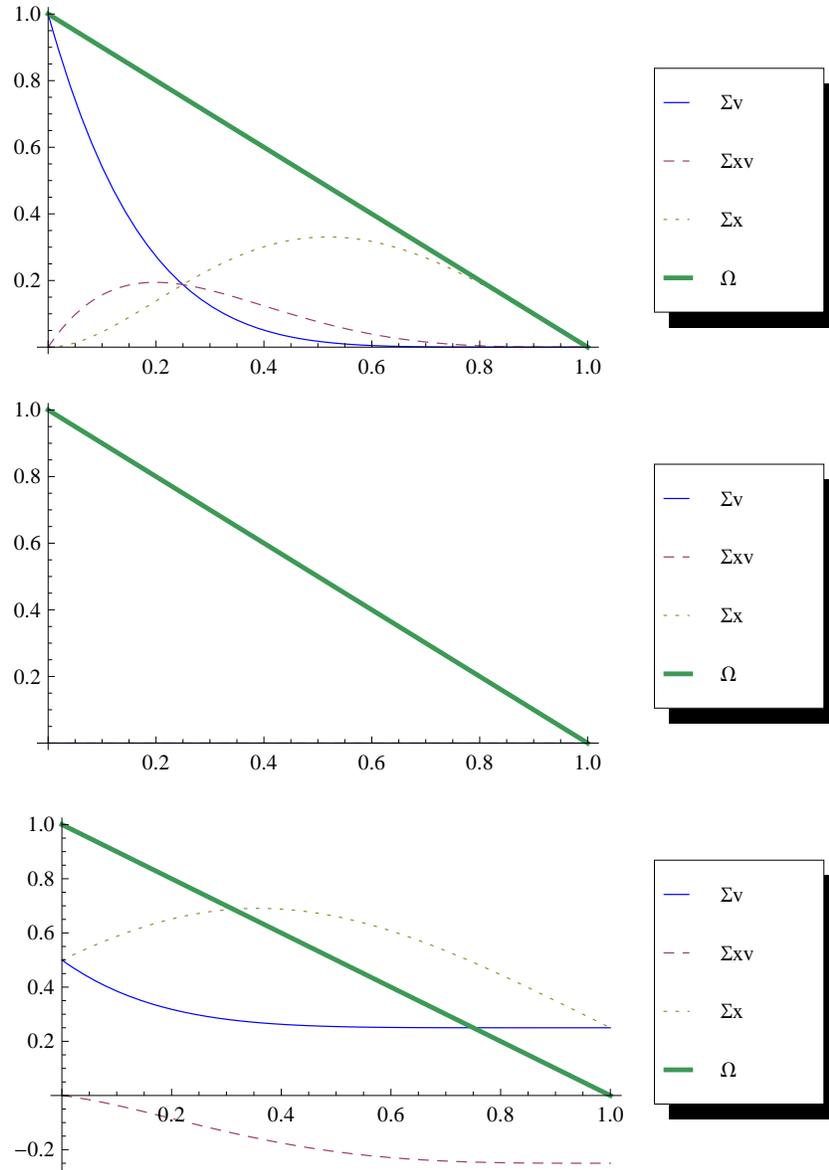


Figure 1: **Flow of private information into Prices.** This figure presents the flow of private information into prices. $\Sigma_v(t)$ summarizes the residual uncertainty about the exogenous terminal value. $\Sigma_X(t)$ the residual uncertainty about the activist's position. $\Sigma_{Xv}(t)$ is the covariance between both. The upper panel corresponds to the case with $\Sigma_X(0) = 0$, the middle panel to $\Sigma_v(0) = 0$ and the lower panel to $\Sigma_X(0) = \Sigma_v(0) = 0.5$. In all cases we set $\Sigma_{Xv}(0) = 0$.

If there is no uncertainty about the exogenous component of firm value ($\Sigma_v(t) = 0 \quad \forall t$) and all adverse selection comes from moral hazard, then Σ_X decays over time and the market will learn perfectly the ability of the activist to generate firm value (middle panel). If however, there is uncertainty about the exogenous component of firm value (that not affected by the activist), then the moral hazard component of adverse selection tends to first grow over time as the activist is expected to accumulate more shares and there is some ‘signal jamming’: the market has difficulty separating its estimate of exogenous and endogenous component of firm value. In the end when maturity approaches the market will learn the ability of the activist to add to firm value only if there is no initial uncertainty about his position (i.e., if $\Sigma_X(0) = 0$ as in the upper panel). If the market is at the outset uncertain about both position and exogenous firm value (i.e., if $\Sigma_X(0)\Sigma_v(0) > 0$ as in the lower panel), then the market will never be able to separate the endogenous from the exogenous firm value. For the calibration presented the results are pretty dramatic. Starting from $\Sigma_X(0) = \Sigma_v(0) = 0.5$ we end up at $\Sigma_X(T) = \Sigma_v(T) \approx 0.3$, so there is considerable uncertainty about both components remaining.

This seems important given the debate about the social ‘usefulness’ of activists. Are these better stock-pickers or do they really create value for minority shareholders by expending effort? The model presented suggests that this is very difficult for the market to sort out, if at the outset there is some uncertainty about the activists holdings. Given our interpretation of date 0 in the model as the first time the market becomes aware of the existence of a potential activist present in the market place, this scenario seems the more likely. Instead, the model suggests that if we can force the activist to disclose

information on his holdings, i.e. reduce $\Sigma_X(0)$, then this makes it easier for the market to sort out his ‘stock-picking’ from his ‘managerial’ abilities. From that perspective, if it is desirable to remunerate activists only for their efforts, disclosure requirements may be useful. Of course, as we discuss further below this may also have unintended consequences for economic efficiency, as it leads the activist to accumulate less shares and thus expend less effort on average.

Figure 2 presents the equilibrium behavior of both components of price impact Λ_t and λ_t (see Lemma 4). Recall that total price impact $\lambda_t + \psi\Lambda_t = \hat{\Delta}$ (see equation (51)) is constant in this model. However, the figure shows that the ‘position impact’ Λ_t increases monotonically closer to maturity. This is consistent with information about the activist’s position (and therefore effort expenditure) becomes more important as maturity approaches. Instead, λ_t decreases monotonically and converges to zero at maturity.

Next, we present the optimal trading strategy of the activist. 52 presents the optimal trading strategy of the activist as a function of his valuation gap. $\beta(t)$ in Figure 3 plots the speed at which the activist reacts to the gap between his assessment of the fundamental value and the market maker’s estimate of the terminal firm value. Note that as in the KB model the activist becomes more ‘aggressive’ as maturity approaches so as to not leave any money on the table. Since he is risk-neutral, any difference between price and expected payoff must lead to aggressive trading on the part of the activist. In equilibrium, however, his actual trading is the product of this rate and of the valuation gap, which disappears as maturity approaches.

Interestingly, when we focus on the actual trading rate θ_t its behavior is quite different

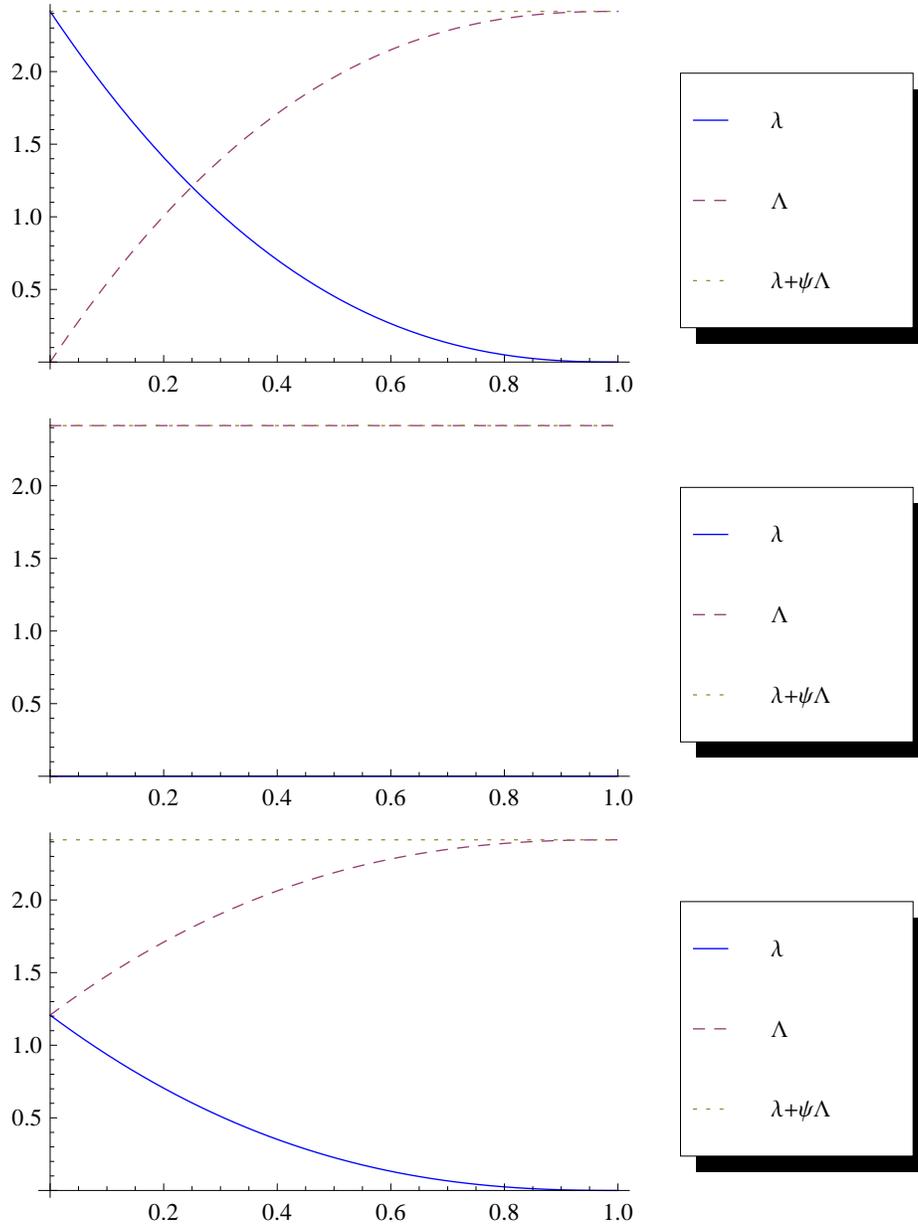


Figure 2: **Components of price impact.** This figure plots the equilibrium paths of the position-impact Λ_t and price-impact λ_t factors. The upper panel corresponds to the case with $\Sigma_X(0) = 0$, the middle panel to $\Sigma_v(0) = 0$ and the lower panel to $\Sigma_X(0) = \Sigma_v(0) = 0.5$. In all cases we set $\Sigma_{Xv}(0) = 0$.

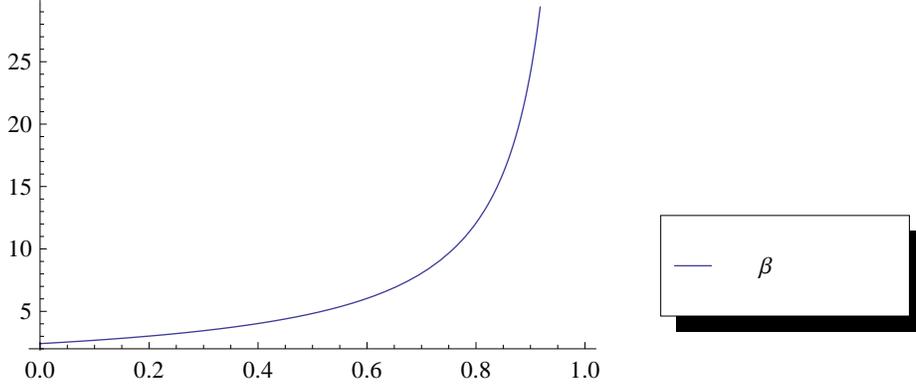


Figure 3: **Optimal trading strategy.** This figure plots the optimal trading strategy of the activist (see equation (10)). $\beta(t)$ plots the speed at which the activist decides to close the gap between his assessment of the fundamental value and the market maker's estimate of the terminal firm value. The figure is drawn for $\Omega(0) = 1$.

than in the traditional Kyle model. Indeed, we can compute explicitly the unconditional expected trading rate of the activist in his filtration $E[\theta_t|\mathcal{F}_0, v, X_0]$. We find:

$$E[\theta_t|\mathcal{F}_0, v, X_0] = (v + \psi X_0 - P_0) \left(\frac{T-t}{T}\right)^\kappa \hat{\Delta} / (T\alpha^2) \quad (57)$$

Clearly the unconditional trading rate is expected to decrease over time. Instead, when there is no moral hazard (when $\psi = 0$), where the model reverts to the traditional Kyle model, the unconditional expected trading rate is constant equal to the initial gap normalized by price impact and time to maturity: $E[\theta_t|\mathcal{F}_0, v, X_0]_{\psi=0} = \frac{v-P_0}{\sigma v T}$. So the presence of moral hazard changes the trading strategy of the activist unconditionally leading him to be more aggressive early on, even though his total price impact ($\hat{\Delta}$) is constant over time. Figure 4 below plots both expected trading rates.

Moral hazard and expected profits of the activist

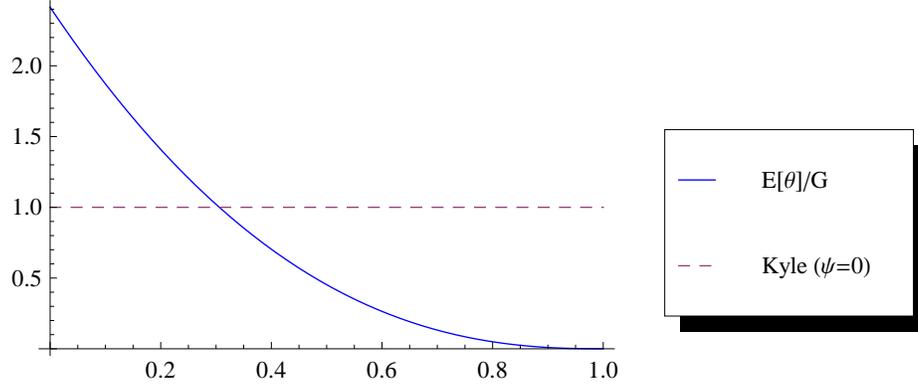


Figure 4: **Optimal trading strategy.** This figure shows the unconditional expected trading rate of the activist shareholder normalized by the initial valuation gap $E[\theta_t|\mathcal{F}_0, v, X_0]/G_0$ with $G_0 = (v + \psi X_0 - P_0)$ as a function of time and compares that to the expected trading rate in the absence of moral hazard, i.e., when $\psi = 0$.

Turning to the profits of the activist investor, we plot his value function on Figure 5 as a function of his valuation gap for different levels of his productivity parameter ψ and on Figure 6 as a function of the productivity parameter for different levels of the valuation gap. Recall that:

$$J(P, X, 0) = \frac{(P - v - \psi X_0)^2 + \hat{\Delta}^2 \sigma^2 T}{2(\hat{\Delta} - \psi)}.$$

We see that for small deviations of the price from its ‘fundamental’ value, the expected profits of the activist are increasing in his ability to create more value for shareholders. However, when deviations are very large, then it may become decreasing in his ability. A possible intuition for this surprising result is that increasing his ability ψ has two effects. On the one hand, for a given stake at maturity it raises the payoff of the activist. On the other hand, it increases his price-impact as the market maker anticipates that, which

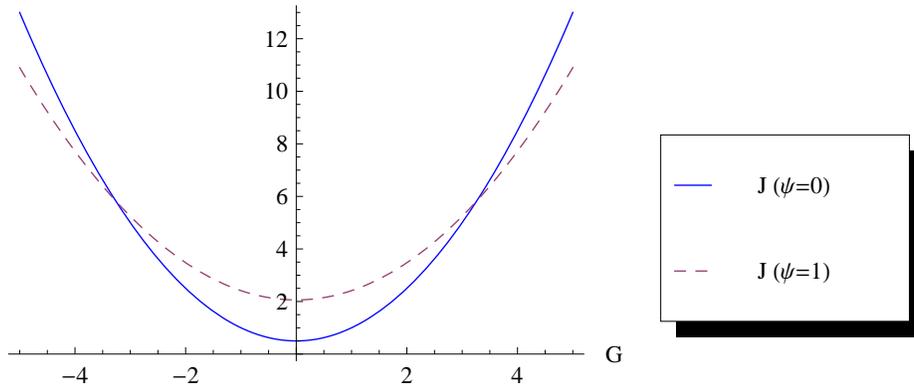


Figure 5: **Value function as function of initial valuation gap.** This figure plots the optimal value function as a function of the initial valuation gap $G = v + \psi X_0 - P_0$ for different values of the productivity level of the activist: $\psi = 0$ which corresponds to the Kyle-Back model with no moral hazard, and $\psi = 1$.

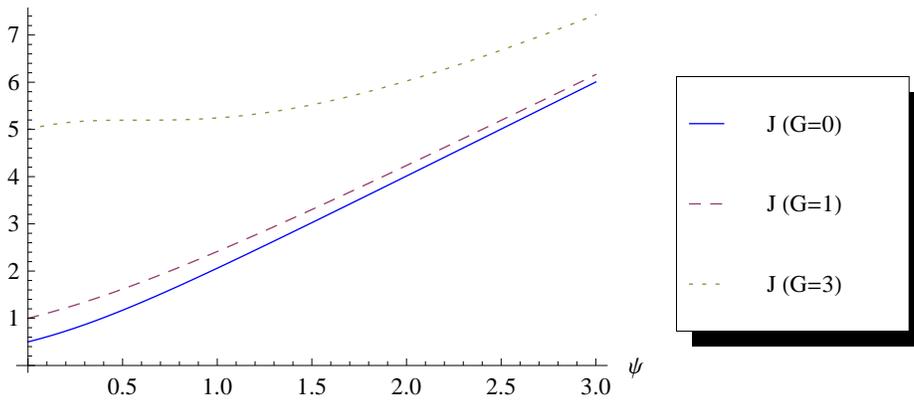


Figure 6: **Value function as a function of productivity.** This figure plots the optimal value function as a function of the productivity parameter ψ for different levels of the initial valuation gap $G = v + \psi X_0 - P_0$: $G = 0$, $G = 1$, $G = 3$.

reduces his profits. For a very large initial valuation gap, the second effect seems to dominate.

We can also calculate the unconditional expected profits of the activist by integrating his value function.

Lemma 5. *The unconditional expected profits of an activist is*

$$U(\omega, \psi, \sigma, T) = \omega\sigma T \left(\zeta + \sqrt{1 + \zeta^2} \right)$$

with $\zeta = \frac{\psi\sigma}{\omega}$ and $\omega^2 = \sigma_v^2 + 2\psi\sigma_{Xv} + \psi^2\sigma_X^2$ as before.

Proof 7. *Taking unconditional expectation of the value function of the activist we see that its unconditional profits are given by:*

$$E[J] = \frac{\omega^2 T + \hat{\Delta}^2 \sigma^2 T}{2(\hat{\Delta} - \psi)}.$$

Now recall that $\omega^2 + (2\psi - \hat{\Delta})\hat{\Delta}\sigma^2 = 0$. Substituting and rearranging we get the expression in the lemma.

We see that total unconditional profits of the activist converge to $U(\omega, \psi = 0, \sigma, T) = \sigma\sigma_v T$ identical to KB in the absence of moral hazard. Further, differentiating the expression given in the lemma we see that unconditional profits $U(\omega, \psi, \sigma, T)$ of an activist are increasing in (a) managerial ability (ψ), (b) (annualized) noise trading volatility (σ), (c) initial level of (annualized) private information (ω), (d) and length of the trading period (T).

Now suppose that an investor needs to pay a fixed cost to establish the potential under or over-valuation of a given firm ($v + \psi X_0 - P_0$), based on which he would decide

to accumulate more shares and eventually become an activist shareholder. These costs would be related to his research time, legal costs, etc. Say these costs amount to I . Then it is clear that an investor would from an ex-ante perspective only spend these initial costs if his unconditional expected profits exceed these costs, i.e., if $U(\omega, \psi, \sigma, T) > I$. This suggests that all else equal we would expect shareholder activism to target firms with greater (a) uncertainty about its fundamental value, (b) greater potential for managerial impact, (c) higher stock price liquidity, and (d) longer trading period. The model thus delivers an interesting link between ex-ante stock liquidity and corporate governance. It also has implications for the current regulatory debate on the allowable discretionary trading period for 13-D filers prior to their filing with the SEC (e.g., Bebchuk et al., 2013).

3. Discussion

The model suggests an interesting relation between economic efficiency and market (or price) efficiency.¹⁴

Economic Efficiency

With respect to economic efficiency, we assume that the first best would be achieved if the activist could acquire the entire firm and produce the optimal amount of effort

¹⁴As in many models with noise-traders, since their utility is not defined, the notion of Pareto efficiency is not well-defined. Our notion of economic efficiency refers instead to the value added by the activist through his effort expenditure. We simply posit that the first best outcome in terms of economic efficiency would be attained if the activist produced at the level he would if he owned the entire firm. Similarly, the notion of price efficiency is also not well defined. Clearly in our model prices are ‘efficient’ in the sense that they always reveal all the public information available (to the market maker). Instead, we use the term ‘price efficiency’ to indicate the degree to which prices reveal the private information available to the activist.

(given his full ownership). Denote the total number of shares issued by the firm by \bar{X} .¹⁵ Then the first best effort level would be such that the firm is worth $v + \psi\bar{X}$ at time T . Now, in general, we would expect the parameters of the model to be such that the activist will not be able to acquire the entire amount of outstanding shares before maturity (i.e., we expect that $X_T < \bar{X}$ most of the time). In that sense, first best ‘economic efficiency’ would not be achieved as long as $X_T < \bar{X}$ since the activist would not expend as much effort as he would if he owned all the firm. The model generates clear predictions as to what drive X_T the ultimate position achieved by the activist.

Lemma 6. *In equilibrium the shares accumulated by the activist X_T are given by*

$$X_T = X_0 + \frac{v + \psi X_0 - P_0}{\sqrt{\psi^2 + \alpha^2}} - \sigma \left(1 + \frac{1}{\sqrt{1 + \frac{1}{\zeta^2}}}\right) Z_T.$$

Thus X_T is normally distributed with mean $E[X_T | \mathcal{F}_0, v, X_0] = X_0 + \frac{v + \psi X_0 - P_0}{\sqrt{\psi^2 + \alpha^2}}$ and variance $V[X_T] = \sigma^2 T \left(1 + \frac{1}{\sqrt{1 + \frac{1}{\zeta^2}}}\right)^2$.

Proof 8. *This follows from $dP_t = \hat{\Delta} dY_t = \hat{\Delta}(dX_t + \sigma dZ_t)$. Integrating and using the fact that $P_T = v + \psi X_T$ at maturity we find:*

$$v + \psi X_T - P_0 = \hat{\Delta}(X_T - X_0) + \hat{\Delta}\sigma(Z_T - Z_0).$$

¹⁵Note that this upper bound is not stricto-sensu part of the model. In the Kyle-type setup we investigate, the implicit assumption is that a trader could potentially accumulate a long position larger than the maximum available number of shares. In that case, the excess would have to be considered a ‘derivative’ (e.g., forward) trade with the market maker that would be settled in cash and not in kind. This interpretation poses some issues when combined with the effort model, where it would seem ‘efficient’ to take the largest possible stake in the firm to expand the largest possible effort. If we impose that the maximum amount of effort is bounded by the available number or shares in the firm, this introduces a non-linearity in the cost function, that one would need to explicitly take into account at the outset. For the discussion here we ignore this complication.

Solving for X_T and using (51) gives the result.

This suggests that when the firm appears under-valued initially to the activist and he would like to accumulate more shares, then the expected number of shares he accumulates is actually decreasing in the signal to noise ratio $\alpha = \omega/\sigma$. This of course reflects the increased price impact. So typically the more noise trading (as measured by σ), the larger the expected number of shares the activist will accumulate (on a positive NPV venture). As a result, the activist will expend more effort and will create more value for shareholders. It implies a positive impact of noise trading on Economic efficiency. Also, note that the variance of her position is increasing in the quantity $\zeta = \frac{\sigma\psi}{\omega}$.

Lastly, note that the expected number of shares purchased is independent of maturity T . This is true for a given *annualized* level of private information ω . If instead, we assume that total initial amount of private information is given i.e., that $\Omega(0)$ is constant, then increasing the trading period T will decrease the signal to noise ratio α (since $\omega^2 := \Omega(0)/T$) leading to an increase in the expected number of shares accumulated by the activist. Therefore, for a given amount of initial private information, the longer the period when the activist can accumulate shares anonymously, the larger the expected number of shares the activist will accumulate (on a positive NPV venture). As a result, the activist will expend more effort and will create more value for shareholders. It implies a positive impact of longer pre-disclosure period on economic efficiency.

These relation suggests there exists a trade-off between economic efficiency and price efficiency.

Price efficiency

Note that in the Kyle model, $\psi = 0$ and markets are most efficient in the absence

of noise traders, i.e., when $\sigma = 0$. In that case, the profits of the activist are driven to zero, since even an infinitesimal trade will reveal all of his information. So given that the informed is indifferent, prices might as well jump to v immediately.¹⁶ The larger the amount of noise trading volatility the larger the activist's profits and the less prices reveal his information. When there is moral hazard and the activist can affect prices, the activist will accumulate fewer shares in expectation when there is less noise trading. Thus, he will create less wealth (since his effort is proportional to the size of his acquired stake). Thus somewhat paradoxically, more 'price efficiency' (in the sense of less noise trading activity and thus more informative prices) will lead to less economic efficiency in the sense that the activist will create less shareholder value. This is a dynamic version of the Grossman and Hart (1980a) free-rider problem when there are noise traders, which offer a partial solution to that problem, as also emphasized in a one period model by Kyle and Vila (1991) and Maug (1998).

4. Empirical Regularities

We developed a model with endogenous effort expenditure by an activist. As we discussed in the introduction, it is natural to think of this activist as a hedge-fund activist for example. An activist hedge fund typically accumulates shares of a public company in the open market and then expends effort to improve the company. Data on trading strategies of such activist shareholders could provide an interesting testing ground for the model. We focus on four implications of the model. First, the model

¹⁶Of course, with an infinitesimal cost, he might not want to trade at all, leading to a Grossman-Stiglitz style paradox.

predicts that the value created by an activist is increasing in his stake size. The higher an activist's stake, the higher effort he will expend and the higher increase in the firm value will be achieved. Second, the model predicts that the unconditional trading rate is expected to decrease closer to the terminal date (see Figure 4). Third, the model shows that an activist's trading strategy should depend not only on the stock price but also on his stock ownership. Specifically, the activist's trading rate is increasing in the size of his acquired position (see proposition 2). Finally, the model predicts that shareholder activism should target relatively more liquid stocks.

We use the novel data-set built by Collin-Dufresne and Fos (2013a) on individual trades by activist share-holders as identified from so-called 13D SEC filings, to investigate these implications of our model. We first describe the data succinctly and then describe the evidence.

4.1. Sample Description

In this section we describe data used in the empirical analysis. Data are compiled from several sources. Stock returns, volume, and prices come from the Center for Research in Security Prices (CRSP). Data on trades by Schedule 13D filers come from a unique hand-collected database, described below (Detailed description of the sample is in Collin-Dufresne and Fos, 2013a).

These authors exploit a disclosure requirement to identify trades that rely on valuable private information. Rule 13d-1(a) of the 1934 Securities Exchange Act requires investors to file with the SEC within 10 days of acquiring more than 5% of any class of securities of a publicly traded company if they have an interest in influencing the management of

the company. In particular, Item 5(c) of Schedule 13D requires the filer to “... describe any transactions in the class of securities reported on that were effected during the past sixty days or since the most recent filing of Schedule 13D, whichever is less.” Thus, Schedule 13D filings reveal the date and price at which all trades by the Schedule 13D filer were executed during the 60 days that precede the filing date. For each event we extract the following information from the Schedule 13D filings: CUSIP of the underlying security, date of every transaction, transaction type (purchase or sell), transaction size, and transaction price. In addition, we extract filing date, event date (date of crossing the 5% threshold), and the beneficial ownership of the Schedule 13D filer at the filing date.

The sample of trades by Schedule 13D filers is constructed as follows. First, using an automatic search script, we identify 19,026 Schedule 13D filings from 1994 to 2010. We then apply criteria described in Collin-Dufresne and Fos (2013a) and are left with the final sample of 3,126 Schedule 13D filings from 1994 to 2010.

4.2. Value Creation

We begin by documenting shareholder activism generates (on average) statistically and economically significant positive excess returns. Figure 7 plots the average buy-and-hold return, in excess of the buy-and-hold return on the value-weighted NYSE/AMEX/NASDAQ index from CRSP, from sixty days prior to the filing date to forty days afterward. The sample includes data from the 1994 to 2010 sample period. There is a run-up of about 3% from sixty days to one day prior to the filing date. The two-day jump in excess return observed at the filing date is around 2.5%. After that

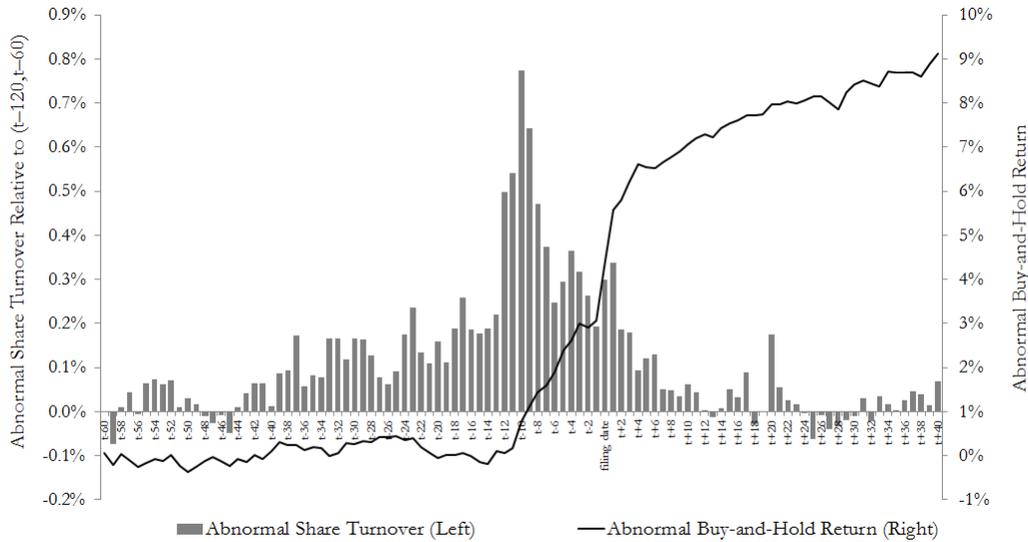


Figure 7: **Buy-and-Hold Abnormal Return around the Filing Date.** In Panel A the solid line (right axis) plots the average buy-and-hold return around the filing date in excess of the buy-and-hold return of the value-weighted market from sixty days prior the filing date to forty days afterwards. The filing date is the day on which a Schedule 13D filing is submitted to the SEC. The dark bars (left axis) plot the increase (in percentage points) in the share turnover during the same time window compared to the average turnover rate during the preceding (t-120, t-60) event window. In Panel B the solid line plots the daily abnormal return. The abnormal return is the average daily return in excess of the value-weighted market return. The dashed lines plot the lower and upper 1% confidence bounds.

the excess return remains positive and the post-filing ‘drift’ cumulates to a total of 9%. Thus, the short-term announcement event-day returns suggest that Schedule 13D filers indeed possess valuable private information during the pre-announcement period.¹⁷

¹⁷Please note that neither our model nor any other microstructure model is consistent with the jump in stock prices upon the announcement. We believe that this empirical regularity is consistent with activists being risk averse. Generalization of the current model to the case of a risk averse activist is outside the scope of this paper and is work in progress.

The dataset reveals that activist shareholders purchase a significant number of shares in targeted companies and therefore have strong incentives to increase the value of their firms. For instance, the average (median) stock ownership of a Schedule 13D filer on the filing date is 7.51% (6.11%). The average (median) filer purchases 3.8% (2.8%) of outstanding shares during the sixty-day period prior to the filing date. It corresponds to an average (median) purchase of 899,692 (298,807) shares at an average (median) cost of \$16.4 (\$2.5) million. Such a significant stake creates a strong incentive for the activist to increase the firm value.

The model predicts that an activist's stake size is positively associated with the value creation. To study this relation, we estimate the following cross-sectional regression:

$$car_i = a + bX_i + \gamma to_i + \varepsilon_i, \tag{58}$$

where car_i is the cumulative abnormal return between the event date and the filing date, X_i is the percentage of outstanding shares of company i owned by a Schedule 13D filer on the filing date, and to_i is the average daily turnover between the event date and the filing date. The daily turnover is defined as the daily volume divided by the number of outstanding shares. The event date is the day when a Schedule 13D filer's ownership in stock i crosses the 5% threshold. The filing date is the day when a Schedule 13D filing is submitted to the SEC. Table 1 reports the results.

[Insert Table 1 here]

Table 1 reveals that activist's stake size is positively associated with cumulative abnormal returns between the event date and the filing date. For instance, column

(1) shows that one standard deviation increase in an activist's stake size (i.e., 3.5% of outstanding shares) is associated with a 0.75% increase in cumulative abnormal returns. Since the average cumulative abnormal return between the event date and the filing date is 2.6%, the economic magnitude of the effect is significant. We next control for share turnover in order to rule out the possibility that the positive association between abnormal returns and an activist's stake size is driven by share turnover. The evidence reported in column (3) shows that the positive association between an activist's stake size and abnormal returns remains almost unchanged when we control for the daily share turnover. One concern with this regression is that it does not establish causality. While the evidence is consistent with activists creating more value when their stake is higher, it may also simply reflect the fact that activists acquire more shares in companies that more undervalued (ex ante) as in the original Kyle (1985) model. The next section tests evidence more specific to the activism model.

4.3. Trading Strategies of Activist Shareholders

The dataset allows us to test whether trading strategies of activist shareholders are consistent with the proposed model. As we show in Figure 4, the model predicts that the unconditional trading rate is expected to decrease as maturity approaches. This is in contrast to the Kyle (1985) model where, since the terminal value is exogenous, an activist's optimal trading intensity is constant.¹⁸

To test whether this feature of the model is supported by data, we plot trading

¹⁸Another feature that could explain a decreasing trading intensity is risk-aversion (e.g., Baruch (2002))

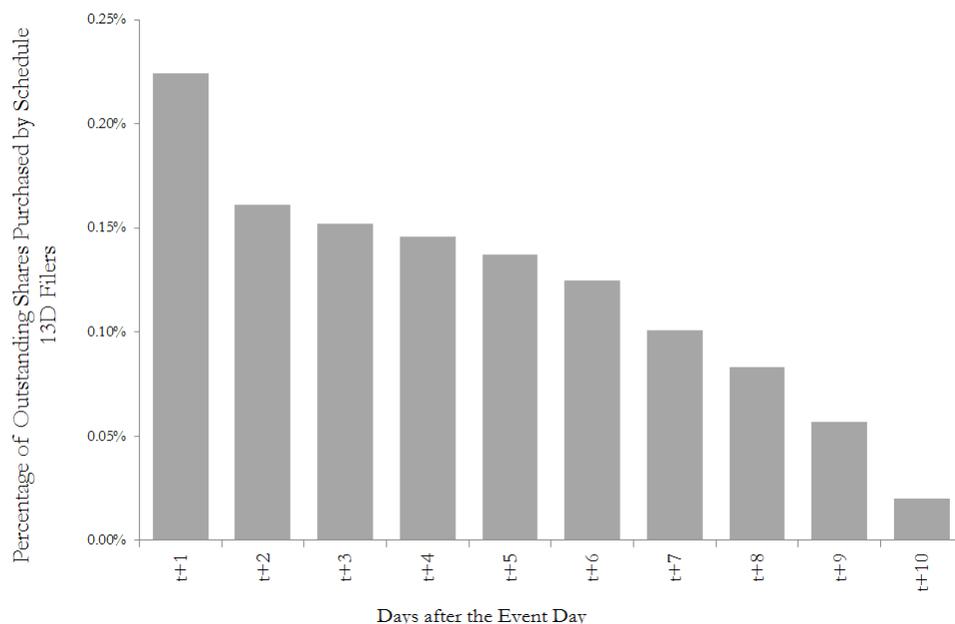


Figure 8: **Trading Strategy of Schedule 13D Filers.** The dark bars plot the average percentage of outstanding shares purchased by Schedule 13D filers after the event day. The event date is the day on which activist’s ownership crosses the 5% threshold.

strategies of Schedule 13D filers after the event day. We restrict the analysis to the post event day period because an activist’s horizon is limited to 10 days after the event day. That is, an activist has up to 10 days to disclose publicly his beneficial ownership and intention. It implies that an activist’s ability to select when and how to trade is limited in this sub-sample (see detailed discussion of this issue in Collin-Dufresne and Fos, 2013a). Figure 8 plots the average percentage of outstanding shares purchased by Schedule 13D filers after the event day.

The plot suggests that an activist’s trading intensity decreases closer to the terminal date. When we regress the percentage of outstanding shares purchased on a given day

on the distance to the terminal date, we find that an activist purchases decrease by 0.02% of outstanding shares on every trading day.¹⁹ Since on an average trading day an activist purchases 0.16% of outstanding shares, the rate at which the trading intensity decreasing is substantial: it decreases by 12.5% on every trading date. Thus, the evidence is consistent with the model’s prediction that the unconditional trading rate is expected to decrease closer to the terminal date.

An additional testable prediction distinguishes the model from the literature. Proposition 2 shows that an activist’s trading strategy depends on the activist’s stake size. This is in contrast to Kyle (1985) and Back (1992), where an activist’s trading strategy depends on the ‘valuation gap’ (i.e., the difference between exogenously given terminal value of the stock and the current stock price) and not on the activist’s stake size. To test whether this prediction is supported by data, we estimate the following regression for every event:

$$\theta_{it} = a_{0i} + a_{1i}X_{it-1} + a_{2i}P_{it} + \varepsilon_{it}, \tag{59}$$

where θ_{it} is the number shares purchased by a Schedule 13D filer in company i on date t , X_{it-1} is the number shares owned by the Schedule 13D filer on date $t - 1$, and P_{it} is the closing price of the stock i on date t . The analysis is based on daily observations from 60 days before the filing date to the filing date.

¹⁹Specifically, we estimate the following regression: $w_{it} = \alpha + \beta(T - t)_i + \varepsilon_{it}$, where w_{it} is the percentage of outstanding shares purchased on day t and $(T - t)$ is the distance to the terminal day T . The estimate of β is 0.02% with t-stat of 22.52, calculated using heteroscedasticity robust standard errors.

Consistently with the proposed model, we find that an activist's trading strategy is positively associated with the stake size in 97% of events and is negatively associated with the stock price in 91% of events. The cross-sectional mean of a_1 is 4.5 with t-stat of 41.05. The cross-sectional mean of a_2 is -18.4 with t-stat of -14.66. Thus, an activist's stock ownership seems to have a significant incremental effect on the activist's trading strategy.

4.4. Characteristics of Targeted Companies

We conclude the empirical analysis by studying characteristics of targeted companies. In order to perform the analysis, we merge the sample of Schedule 13D filings with CRSP, Compustat, and Thomson Reuters databases. We begin the analysis from reporting average levels of firm characteristics in Table 2. Schedule 13D filers are less likely to target companies with illiquid stocks. In addition, stocks of targeted firms have higher idiosyncratic volatility, higher proportion of shares owned by institutional investors, and higher proportion of shares owned by activist hedge funds.²⁰ This evidence is consistent with the proposed model (see Section 3). In addition, targeted companies have lower market cap and higher book-to-market ratio when compared to non-targeted companies (similar evidence is documented by Brav et al., 2008; Klein and Zur, 2009; Fos, 2012).

[Insert Table 2 here]

²⁰We thank Wei Jiang for providing data on beneficial ownership of activist hedge funds.

We estimate the following linear probability model in order to further investigate characteristics of Schedule 13D targets:

$$Targeted_{it} = X_{it-1}\alpha_1 + \zeta_t + \zeta_j + \varepsilon_{it}, \quad (60)$$

where *Target* is an indicator of firm *i* being targeted in year *t*, X_{it-1} is the vector of firm characteristics (defined in Table 2) lagged by one year, ζ_t are firm fixed effects, and ζ_j are industry fixed effects (2-digit SIC). Since the regression includes year and industry fixed effects, the reported evidence is based on within year and within industry differences in firm characteristics. Table 3 reports the results.

[Insert Table 3 here]

Overall, the results are consistent with the univariate analysis with one exception: the number of years on CRSP is now significantly associated with the likelihood of a Schedule 13D filing. Specifically, ‘older’ companies are more likely to experience a Schedule 13D filing. This evidence is consistent with ‘older’ companies being associated with a more severe agency problem.

5. Conclusion

We have proposed a model of activist share-holder that extends Kyle (1985) and Back (1992) to allow for an endogenous liquidation value that is determined by the effort level chosen by the informed activist. In equilibrium, price impact reflects two sources of information asymmetry: one related to the insider’s pure informational advantage (his

stock-picking ability) as in the original Kyle model, and one that is related to moral hazard (his share-holder activism).

We find that while, in equilibrium, prices eventually reveal the total value of the firm, in many cases, the market cannot identify the actual source of value ‘creation.’ Forcing activists to disclose information on their position helps in separating stock-pickers from activists. However, more disclosure also leads to less share-accumulation by the activist who then exerts less effort and creates less value. Indeed, in general the model shows that less price efficiency (i.e., higher noise-trading volatility) allows the insider to accumulate more shares and thus to exert more effort to generate more share-holder value.

Using a data-set on trades by activist share-holders built from 13D SEC filings (first used by Collin-Dufresne and Fos (2013a)) we document several stylized facts that are broadly in line with the predictions of the model. Indeed, we find that the value created by an activist and his trades are empirically significantly positively related to his accumulated stake size. We also confirm that activism tends to target stocks that have both relatively higher trading liquidity and higher idiosyncratic volatility.

References

- Admati, A. R., Pfleiderer, P., Zechner, J., 1994. Large shareholder activism, risk sharing, and financial market equilibrium. *Journal of Political Economy* 102 (6), 1097–1130.
- Agarwal, V., Jiang, W., Tang, Y., Yang, B., 2013. Uncovering hedge fund skill from the portfolio holdings they hide. *The Journal of Finance* 68 (2), 739–783.
- Amihud, Y., January 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5 (1), 31–56.
- Back, K., 1992. Insider trading in continuous time. *Review of Financial Studies* 5, 387–409.
- Baruch, S., 2002. Insider trading and risk aversion. *Journal of Financial Markets* 5 (4), 451 – 464.
- Bebchuk, L., Brav, A., Jackson, R., Jiang, W., 2013. Pre-disclosure accumulations by activist investors: Evidence and policy. *Journal of Corporation Law* 39.
- Brav, A., Jiang, W., Partnoy, F., Thomas, R., August 2008. Hedge fund activism, corporate governance, and firm performance. *The Journal of Finance* 63 (4), 1729–1775.
- Collin-Dufresne, P., Fos, V., July 2013a. Do prices reveal the presense of informed trading?, working paper.
- Collin-Dufresne, P., Fos, V., March 2013b. Insider trading, stochastic liquidity, and equilibrium prices, working paper.

- Easley, D., O'Hara, M., 1987. Price, trade size, and information in securities markets. *Journal of Financial Economics* 19 (1), 69 – 90.
- Edmans, A., February 2013. Blockholders and corporate governance, working paper.
- Fos, V., February 2012. The disciplinary effects of proxy contests, working paper.
- Glosten, L. R., Milgrom, P. R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14 (1), 71–100.
- Grossman, S., Hart, O., 1980a. Takeover bids, the free-rider problem, and the theory of the corporation. *The Bell Journal of Economics* 11 (1), 42–64.
- Grossman, S. J., Hart, O. D., 1980b. Disclosure laws and takeover bids. *The Journal of Finance* 35 (2), 323–334.
- Grossman, S. J., Stiglitz, J. E., 1980. On the impossibility of informationally efficient markets. *The American Economic Review* 70 (3), 393–408.
- Klein, A., Zur, E., January 2009. Entrepreneurial shareholder activism: Hedge funds and other private investors. *The Journal of Finance* 64 (1), 187–229.
- Kyle, A., November 1985. Continuous auctions and insider trading. *Econometrica* 53 (6), 1315–1335.
- Kyle, A., Vila, J.-L., 1991. Noise trading and takeovers. *RAND Journal of Economics* 22 (1), 54–71.

- Liptser, R., Shiryaev, A., 2001. *Statistics of Random Processes II: Applications*, 2nd Edition. Springer Verlag.
- Maug, E., February 1998. Large shareholders as monitors: Is there a trade-off between liquidity and control? *The Journal of Finance* 53 (1), 65–98.
- Noe, T. H., 2002. Investor activism and financial market structure. *The Review of Financial Studies* 15 (1), 289–318.
- Norli, O., Ostergaard, C., Schindele, I., 2010. Liquidity and shareholder activism, working paper, EFA 2009 Bergen Meetings.
- Shleifer, A., Vishny, R. W., 1986. Large shareholders and corporate control. *Journal of Political Economy* 94 (3), 461–488.

Table 1: **Value Creation by Activist Shareholders.** This table studies value creation by Schedule 13D filers. We estimate the following cross-sectional regression: $car_i = a + bX_i + \gamma to_i + \varepsilon_i$, where car_i is the cumulative abnormal return between the event date and the filing date, X_i is the percentage of outstanding shares of company i owned by the Schedule 13D filer on the filing date, and to_i is the average daily turnover between the event date and the filing date (daily volume divided by the number of shares outstanding). The event date is the day when Schedule 13D filer's ownership in stock i crosses 5% threshold and the filing date is the day on which the Schedule 13D filing is submitted to the SEC. In each column, we report estimated coefficients and their t -statistics. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels.

	(1)	(2)	(3)
X_i	0.2134*** [2.90]		0.1899*** [2.58]
to_i		0.8175*** [4.38]	0.7528*** [3.93]
Constant	0.0093 [1.54]	0.0173*** [5.43]	0.0029 [0.47]
Observations	1,854	1,702	1,702

Table 2: **Firm Characteristics.** This table compares firm characteristics of targeted and non-targeted companies. *Illiquidity* is Amihud (2002) illiquidity measure, defined as the yearly average (using daily data) of $1000\sqrt{|Return|/DollarTradingVolume}$. *Idiosyncratic Volatility* is the ratio of idiosyncratic to total volatility from Fama=French three-factor model (estimated using daily data). *Institutional Ownership* is the proportion of shares held by institutions. *Activist Hedge Fund Ownership* is the proportion of shares held by activist hedge funds. *Market Cap* is market capitalization in millions of dollars. *Book-to-market* is the ratio of the book value of equity to the market value of equity. *Number of Years on CRSP* is the number of years since first appearance on CRSP. Column (1) reports the average level of each characteristic of targeted companies during the pre Schedule 13D filing year. Column (2) reports the average level of each characteristic of non-targeted companies. Column (3) reports the difference between columns (1) and (2). Column (4) reports t-stat of the difference. *** indicates statistical significance at the 1% level.

	targeted companies (1)	non-targeted (2)	difference (3)	t-stat (4)
Illiquidity	0.488	0.534	-0.047***	-3.07
Idiosyncratic Volatility	0.949	0.922	0.027***	13.45
Institutional Ownership	0.433	0.347	0.085***	13.39
Activist Hedge Fund Ownership	0.058	0.025	0.033***	32.12
Market Cap (\$m)	592	2,541	-1,949***	7.42
Book-to-market	0.744	0.558	0.186***	9.15
Number of Years on CRSP	13.621	13.767	-0.146	-0.49

Table 3: **Linear Probability Model.** This table reports estimates of the linear probability model: $Targeted_{it} = X_{it-1}\alpha_1 + \zeta_t + \zeta_j + \varepsilon_{it}$, where $Target$ is an indicator of firm i being targeted in year t , X_{it-1} is vector of firm characteristics (defined in Table 2) lagged by one year, ζ_t are firm fixed effects, and ζ_j are industry fixed effects (2-digit SIC). In each column, we report estimated coefficients and heteroscedasticity robust standard errors clustered by industry (2-digit SIC). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels.

	(1)	(2)	(3)	(4)
Illiquidity	-0.002* [0.001]		-0.006*** [0.001]	-0.012*** [0.002]
Idiosyncratic Volatility		0.058*** [0.008]	0.082*** [0.009]	0.041*** [0.008]
Institutional Ownership				0.034*** [0.005]
Activist Hedge Fund Ownership				0.180*** [0.023]
Market Cap (log)				-0.009*** [0.001]
Book-to-market				0.005*** [0.001]
Number of Years on CRSP				0.000*** [0.000]
Constant	0.056*** [0.002]	-0.003 [0.009]	-0.021** [0.009]	0.061*** [0.010]
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
R^2	0.008	0.009	0.010	0.021
N	89,799	89,117	89,064	68,597

6. Appendix

6.1. Proof of Lemma 1

Here we give a heuristic derivation of the filtering equations given in lemma 1 based on the Gaussian projection theorem, a discrete time approximation of the continuous time model and taking the limit as the time step dt goes to zero.

$$\begin{aligned}
P_{t+dt} &= \mathbb{E}[v | Y^t, Y_{t+dt}] \\
&= \mathbb{E}[v | Y^t] + \frac{\text{Cov}(v, Y_{t+dt} - Y_t | Y^t)}{V(Y_{t+dt} - Y_t | Y^t)} (Y_{t+dt} - Y_t - \mathbb{E}[Y_{t+dt} - Y_t | Y^t]) \\
&= V_t + \frac{\beta \Sigma_v dt + \gamma \Sigma_{Xv} dt}{\beta^2 \Sigma_v dt^2 + \gamma^2 \Sigma_X dt^2 + 2\beta\gamma \Sigma_{Xv} dt^2 + \sigma^2 dt} (Y_{t+dt} - Y_t) \\
&\approx V_t + \frac{\beta \Sigma_v + \gamma \Sigma_{Xv}}{\sigma^2} (dY_t).
\end{aligned}$$

The third line uses the fact that the expected change in order flow is $\mathbb{E}[Y_{t+dt} - Y_t | Y^t] = 0$ for the conjectured policy. The last line follows from going to the continuous time limit (with $dt^2 \approx 0$).

Thus we also find:

$$\lambda = \frac{\beta \Sigma_v + \gamma \Sigma_{Xv}}{\sigma^2}$$

Similarly, by the projection theorem, we have:

$$\text{Var}[v | Y^t, Y_{t+dt}] = \text{Var}[v | Y^t] - (\lambda_t)^2 \text{Var}[Y_{t+dt} - Y_t | Y^t], \quad (61)$$

which gives (when keeping only order dt terms or lower):

$$\Sigma_{t+dt} \approx \Sigma_t - \lambda_t^2 \sigma^2 dt. \quad (62)$$

A similar ‘proof’ applies for the optimal filter for X_t .

$$\begin{aligned} Q_{t+dt} &= \mathbb{E} [X_{t+dt} | Y^t, Y_{t+dt}] \\ &= \mathbb{E} [X_{t+dt} | Y^t] + \frac{\text{Cov}(X_{t+dt}, Y_{t+dt} - Y_t | Y^t)}{V(Y_{t+dt} - Y_t | Y^t)} (dY_t) \\ &= Q_t + \frac{\beta \Sigma_{Xv} dt + \gamma \Sigma_X dt}{\beta^2 \Sigma_v dt^2 + \gamma^2 \Sigma_X dt^2 + 2\beta\gamma \Sigma_{Xv} dt^2 + \sigma^2 dt} dY_t \\ &\approx Q_t + \frac{\beta \Sigma_{Xv} + \gamma \Sigma_X}{\sigma^2} dY_t. \end{aligned}$$

since

$$\begin{aligned} \mathbb{E} [X_{t+dt} | Y^t] &= \mathbb{E} [X_t + (\beta(v - V_t) + \gamma_t(X_t - Q_t)) dt | Y^t] \\ &= Q_t \end{aligned}$$

and

$$\begin{aligned} \text{Cov}(X_{t+dt}, Y_{t+dt} - Y_t | Y^t) &= \text{Cov}(X_t + (\beta(v - V_t) + \gamma_t(X_t - Q_t)) dt, (\beta(v - V_t) + \gamma_t(X_t - Q_t)) dt | Y^t) \\ &= \beta_t \Sigma_{Xv} dt + \gamma_t \Sigma_X dt + [\dots] dt^2 \\ &\approx (\beta_t \Sigma_{Xv} + \gamma_t \Sigma_X) dt. \end{aligned}$$

We thus obtain:

$$\Lambda = \frac{\beta\Sigma_{Xv} + \gamma\Sigma_X}{\sigma^2}.$$

Similarly, we have:

$$\text{Var} [X_{t+dt} | Y^t, Y_{t+dt}] = \text{Var} [X_t + \beta_t(v - V_t)dt + \gamma_t X_t dt | Y^t] - \Lambda_t^2 \text{Var} [Y_{t+dt} - Y_t | Y^t], \quad (63)$$

which gives:

$$\Sigma_X(t + dt) \approx \Sigma_X(t) + 2\beta_t \Sigma_{Xv} dt + 2\gamma_t \Sigma_X dt - \Lambda_t^2 \sigma^2 dt \quad (64)$$

$$= \Sigma_X(t) + \sigma^2 \Lambda(2 - \Lambda) dt. \quad (65)$$

Lastly, we can compute the covariance between the two filters from:

$$\text{Cov} [X_{t+dt}, v | Y^t, Y_{t+dt}] = \text{Cov} [X_{t+dt}, v | Y^t] - \lambda_t \Lambda_t \text{Var} [dY | Y^t] \quad (66)$$

$$\approx \Sigma_{Xv} + \sigma^2 \lambda(1 - \Lambda) dt. \quad (67)$$

Since

$$\text{Cov} [X_{t+dt}, v | Y^t] = \text{Cov} [X_t + \beta_t(v - V_t)dt + \gamma_t(X_t - Q_t)dt, v | Y^t] \quad (68)$$

$$\approx \Sigma_{Xv} + \beta \Sigma_v dt + \gamma \Sigma_{Xv} dt \quad (69)$$

$$= \Sigma_{Xv} + \sigma^2 \lambda dt \quad (70)$$

6.2. One period Kyle model with moral hazard

To show our results are not specific to continuous time setup we derive the optimal trading strategy and equilibrium price dynamics in a simple one-period setup.

We assume as in the main part of the paper, that the activist is endowed with an initial number of shares X_0 and that he will purchase θ shares in the market by submitting a market order to the market maker. The market maker will set prices so as to break even, i.e., such that his expected profits are zero. The market maker only observes total order flow $Y = \theta + u$ where $u \sim N(0, \sigma_u^2)$ is uninformed noise trading which we assume to be normally distributed. After the trades are settled, the activist can choose his level of activism w by paying a cost $C(w) = \frac{w^2}{2\psi}$. As a result of this action the terminal payoff of the firm is realized and equal to $v + w$ where v is a constant known only to the activist at date 0.

We assume the market maker's initial prior about X_0 and v is Gaussian: $v \sim N(V_0, \sigma_v^2)$ and $X_0 \sim N(Q_0, \sigma_Q^2)$ and that their covariance is σ_{Xv} .

To summarize the activist's problem is:

$$\max_{\theta, w} E[(v + w - P_1)\theta + (w + v)X_0 - C(w) | v, X_0] \quad (71)$$

Because of the assumed timing, the activist's optimal choice of w will maximize:

$$\max_w w(\theta + X_0) - C(w), \quad (72)$$

With our choice of cost function this leads to:

$$w^* = \psi(\theta + X_0). \quad (73)$$

Plugging back into his objective function we see that the activist is maximizing:

$$\max_{\theta} \mathbb{E} \left[vX_0 + (v - P_1)\theta + \frac{\psi(\theta + X_0)^2}{2} \mid v, X_0 \right]. \quad (74)$$

We look for a linear equilibrium where price responds linearly to order flow:

$$P_1 = P_0 + \Delta Y$$

we obtain (after dropping the constant term and taking expectation):

$$\max_{\theta} v\theta - (P_0 + \Delta\theta)\theta + \frac{\psi(\theta + X_0)^2}{2} \quad (75)$$

The FOC w.r.t. θ is:

$$v - P_0 - 2\Delta\theta + \psi(\theta + X_0) = 0. \quad (76)$$

Thus:

$$\theta^* = \frac{v - P_0 + \psi X_0}{2\Delta - \psi}. \quad (77)$$

We see that the optimal trading strategy (assuming a linear price order flow relation) is

linear of the form:

$$\theta = \alpha + \beta v + \gamma X_0 \quad (78)$$

$$\alpha = -\frac{P_0}{2\Delta - \psi} \quad (79)$$

$$\beta = \frac{1}{2\Delta - \psi} \quad (80)$$

$$\gamma = \frac{\psi}{2\Delta - \psi} \equiv \psi\beta. \quad (81)$$

Next we show that if the trading strategy is linear of the form (78), then the price order flow relation is indeed linear. Recall that the market maker is risk-neutral and sets prices such that:

$$\begin{aligned} P_1 &= \mathbb{E}[v + w \mid Y] \quad (82) \\ &= \mathbb{E}[v + \psi(\theta + X_0) \mid Y] \\ &= \mathbb{E}[\psi\alpha + v(1 + \psi\beta) + \psi(\gamma + 1)X_0 \mid Y] \end{aligned}$$

Using normality of the random variables we obtain:

$$P_1 = \mathbb{E}[\psi\alpha + v(1 + \psi\beta) + \psi(\gamma + 1)X_0] + \frac{\text{Cov}(v(1 + \psi\beta) + \psi(\gamma + 1)X_0, Y)}{\text{Var}(Y)} (Y - \mathbb{E}[Y])$$

Now, note that

$$\mathbb{E}[\psi\alpha + v(1 + \psi\beta) + \psi(\gamma + 1)X_0] = \psi\alpha + V_0(1 + \psi\beta) + \psi(\gamma + 1)Q_0$$

$$\text{Cov}(v(1 + \psi\beta) + \psi(\gamma + 1)X_0, Y) = (1 + \psi\beta)\beta\sigma_v^2 + \psi(\gamma + 1)\gamma\sigma_X^2 + (\gamma + \psi\beta(2\gamma + 1))\sigma_{Xv}$$

$$\text{Var}(Y) = \beta^2\sigma_v^2 + \gamma^2\sigma_X^2 + 2\beta\gamma\sigma_{Xv} + \sigma_u^2$$

Thus we have shown that if the optimal strategy θ is linear then the price-order flow is indeed linear and of the form $P_1 = P_0 + \Delta Y$ with:

$$\Delta = \frac{(1 + \psi\beta)\beta\sigma_v^2 + \psi(\gamma + 1)\gamma\sigma_X^2 + (\gamma + \psi\beta(2\gamma + 1))\sigma_{Xv}}{\beta^2\sigma_v^2 + \gamma^2\sigma_X^2 + 2\beta\gamma\sigma_{Xv} + \sigma_u^2} \quad (83)$$

and

$$P_0 = \psi\alpha + V_0(1 + \psi\beta) + \psi(\gamma + 1)Q_0 - \Delta(\alpha + \beta V_0 + \gamma Q_0) \quad (84)$$

It remains to find the fixed point solution to the system of equations (79), (80), (81), (83), (84), if it exists.

Using $\gamma = \psi\beta$ and defining $\omega^2 \equiv \sigma_v^2 + 2\psi\sigma_{Xv} + \psi^2\sigma_X^2$ we obtain:

$$\Delta = \frac{\beta\omega^2(1 + \psi\beta)}{\beta^2\omega^2 + \sigma_u^2}. \quad (85)$$

From equation (80) we obtain $2\beta\Delta = 1 + \beta\psi$. After substituting it in equation (85) we find:

$$\beta = \frac{\sigma_u}{\omega}. \quad (86)$$

Then we immediately obtain:

$$\Delta = \frac{1}{2} \left(\psi + \frac{\omega}{\sigma_u} \right) \quad (87)$$

$$\gamma = \frac{\psi\sigma_u}{\omega} \quad (88)$$

$$P_0 = V_0 + \psi Q_0 \quad (89)$$

$$\alpha = -\beta P_0 \tag{90}$$

Note that, as in the continuous time solution we obtain $\theta = \beta(v + \psi X_0 - P_0)$ and $E[\theta^*] = 0$. Further, similarly to the continuous time solution we see that the total price impact is increasing in the productivity of the activist (ψ) as well as in the signal to noise ratio $\frac{\omega}{\sigma_u}$, which implies that it is increasing in the uncertainty about his position (σ_X).