A Market Microstructure Rationale for the S&P Game*

Paolo Colla
CORE, Université catholique de Louvain
34, Voie du Roman Pays
1348 Louvain la Neuve, Belgium
colla@core.ucl.ac.be

Abstract

We develop a dynamic trading game in which fundamental insiders coexist with non-fundamental speculators. Non-fundamental speculators possess superior information about the future noise trades and are able to make sharper inference about the fundamental value with respect to the market maker. We show that non-fundamental speculators decrease market depth as well as the insider's ex-ante gains. We study inclusions in the S&P 500 after October 1989 as an example in which non-fundamental speculation may arise due to the preannouncement practice in index replacements. Evidence on the trading activity and the bid-ask spread pattern is consistent with our theoretical analysis.

*We thank Helena Beltran, Andrea Beltratti, Markus Brunnermeier, Antonio Mele, Jean-Charles Rochet, Paolo Vitale and seminar participants at the FMG and CORE for helpful comments. Financial support from the Bank of Italy, Ente 'L. Einaudi', FMG and LSE is gratefully acknowledged. All errors remain ours.

Introduction

The effect on stock prices induced by changes in the composition of broad market indexes has been addressed by many researchers. Most of the empirical work conducted so far focus on the Standard and Poor's 500[1]. There are several reasons behind the attention devoted to the S&P 500. First of all, both investors and institutions can easily trade stocks included in the S&P 500. As a result, index changes are followed by the financial community at large. Secondly, even though Standard and Poor's sets out several criteria for companies to be included in the index, changes to the S&P 500 roster entail some degree of subjectivity. Thus inclusions in the index are unpredictable and cannot be anticipated by the market as a whole. Finally, changes to the S&P 500 are publicly announced usually five business days before they become effective. Different intervals are occasionally used by S&P.

Each year Standard and Poor's publishes a list of the leading S&P 500 passive fund managers together with their assets under management. From the S&P annual survey of indexed assets (2003) it emerges that more that $1.1 trillion dollars were pegged to the S&P 500 at the end of 2003. This figure is possibly a conservative estimate, since Standard and Poor's claims that it captures approximately 90-95% of the total indexed assets in its survey. Index replacements therefore represent a clear rebalancing opportunity for indexers: when a stock is added to the S&P 500, passive funds should buy it. Pruitt and Wei (1989) find changes in institutional investors' holdings to be positively correlated with the abnormal returns experienced by additions to the S&P 500 over the period 1973-1986. This way they establish a direct link between rabalancing by large institutional index funds and price changes subsequent to index replacements.

Within the financial literature, there is general agreement that index changes result in a temporary demand shift represented by index funds' trading activity, which causes prices to increase for included stocks (and to a decrease for deleted companies). Shleifer (1986) does not find any significant price impact over a sample consisting of 144 additions during 1966-76, and relates this evidence to the small value of the S&P 500 owned by index funds –less than 0.5% in 1975. A similar explanation is given in Harris and Gurel (1986) for the evidence that prices for stocks added to the S&P 500 are not significantly affected over the period 1973-77. Along the same lines, Beneish and Gardner (1995) do not find any effect on the price and trading volume of newly included firms in the Dow Jones Industrial Average, and they point at the scarcity of funds pegged to the DJIA as the main reason for this. According to figure 1, inclusion in the S&P 500 during 2003 implies an additional demand due to indexers for about 12% of the outstanding shares. While the role of this demand shift is generally acknowledged in all the studies on list changes, researchers disagree on the temporary/permanent nature of the price impact as well as on the explanation for it [Chen, Noronha and Singal (2004) and Singal (2003) contain a detailed literature survey].

Starting from October 1989 Standard and Poor's preannounces changes in the S&P 500 usually five days before the inclusion. As documented in Beneish and Whaley (1996), Lynch and Mendenhall (1997), and more recently Blume and Edelen (2004), prices increase after the announcement but they do not immediately adjust to the level prevailing upon inclusion. This pattern clearly opens the way to profitable opportunities. In fact, Beneish and Whaley (1996) argue that indexers might enhance their returns buying earlier during the announcement period, thus making the entire price adjustment occur after annoucement. However, Blume and Edelen (2004) show that this early-trading strategy dampens passive managers' performance resulting in higher tracking errors. Looking at the volume pattern around index replacements, they conclude that half of the funds pegged to the S&P 500 submit their orders during the effective day of inclusion. Similarly, Beneish and Whaley (1996) find that prices tend to increase from the open to the close on the effective day over their 1989-1994 sample, supporting last day buying pressure by index funds. Moreover they document a temporary upward shift in the average trade size, which the authors relate to pegged funds waiting until the effective day to rebalance. These findings are consistent with daily tracking error being the driving criterion for indexers' performance evaluation.

The preannouncement practice makes attractive front-running passive assets through the so-called 'S&P game': buy the included stock immediately after the announcement, and sell it at possibly higher prices after the indexers' demand is satisfied. Trading activity dynamics exhibit abnormal average volume following the announcement, which one can attribute to investors –rather than indexers– playing the S&P game. Early-trading profitability is documented in Beneish and Whaley (1996), Blume and Edelen (2004) and Singal (2003) for inclusions between January and July 2002.

Taking the empirical evidence mentioned above as a starting point, this paper contributes to the literature in several ways. In the first place we provide a modelling framework for index replacements: while several studies document returns and trading activity patterns around inclusions, on the theoretical side little work has been done. In Wurgler and Zhuravskaya (2002) demand shifts generate large stock price movements whenever stocks are not perfectly substitutes. Their model is static and, admittedly, cannot be applied to preannounced index changes. Within the market microstructure literature, several authors considered the value of anticipating uninformed trades such as passive funds' demand. Building up on Kyle (1985) some extensions have been proposed addressing this issue. Rochet and Vila (1994) develop a static game in which the insider is aware of both the final liquidation value and the noise traders' demand while submitting his (limit) order. With respect to the static Kyle (1985) equilibrium, they show that an informed investor trades less aggressively on his price signal and in the opposite direction of his volume signal, offsetting half of the uninformed trades. The aggregate order flow and market liquidity decrease, while prices as well as the insider's unconditional profits are unaltered. The latter result seems to preclude any role for profitable speculation based on knowledge of uninformed trades. However in a dynamic setting this is no longer the case, as shown in Yu (1999). At every batch auction the insider's information set –in addition to the final liquidation value– comprises a noisy signal of the uninformed trades. Comparing the insider's expected profits arising from this model to the sequential auction equilibrium in Kyle (1985), it is shown that both the value of knowing (current) noise trades and market liquidity depend on the signal's precision. Our analysis is closely related to the two-period trading model in Madrigal (1996), where a (non-fundamental) speculator profits from privileged information on past uninformed trades he is endowed with. The author shows that this superior knowledge enables the speculator to make sharper forecasts of the final liquidation value with respect to the market maker. The profitability of strategies based on non-fundamental information is analyzed in Foucault and Lescourret (2003) as well. In this paper we explicitly model the preannouncement practice in S&P 500 replacements after October 1989 considering a market in which an insider coexists with a speculator who possesses superior information with respect to *future* uninformed trades, i.e. passive funds' entry at the effective inclusion date. The second contibution of our paper lies in the empirical evidence we provide. We analyze trading volume and bid-ask spreads around index additions between 1989 and 1999. While volume patterns have been extensively documented (and our findings are in line with the existing literature), spread dynamics have received little attention. Edmister, Graham and Pirie (1996) and Erwin and Miller (1998) document improved liquidity, i.e. tighter bid-ask spreads, after inclusion. However both works consider additions before October 1989, thus offering no grounds for studying the S&P game. In our knowledge spreads under the S&P preannouncing policy are analyzed in Beneish and Whaley (1996) only. The authors report a significant spread decrease on the day following the inclusion. On the other hand we find that index additions worsen liquidity. One possible reason for this contrasting evidence is the different sample, since we consider 108 inclusions whereas Beneish and Whaley (1996) deal with 30 companies added to the S&P 500.

The outline of this paper is as follows. The interplay between index funds and S&P 500 replacements is discussed in section 1. The benchmark model is presented in section 2 where the equilibrium in the absence of the speculator is analyzed. Section 3 explicitly introduces a role for non-fundamental speculation based on strategies like the S&P game. We show that front-running index funds is indeed profitable, and results in higher volume and lower liquidity. Section 4 discusses testable implications arising from the theoretical model, while section 5 presents the empirical evidence on S&P 500 inclusions. Finally section 6 concludes.

1 Passive funds, index replacements and the S&P game

The appeal of passive techniques to investors has increased during the last two decades. In 1976, $19 billion out of a total market value of $662 billion were pegged to the S&P 500, which corresponds to 3% of the index capitalization [see Wurgler and Zhuravskaya (2002)]. At the end of 2003 more than $1 trillion were indexed directly or indirectly to the S&P 500, representing roughly 12% of the total index capitalization. Figure 1 presents the indexed assets over the period 1990-2003 as well as the passive industry weight relative to the whole S&P 500 market capitalization. The figure leaves no doubt that passive assets have grown over the last 15 years.

According to Blume and Edelen (2004) full replication and stratified sampling are the replicating strategies commonly implemented by S&P 500 indexers. Full replication requires holdings in all the 500 stocks in the exact proportion to their weights in the index at all times, while sampling strategies hold less that 500 stocks. When a new stock is added to the S&P 500, passive managers should replicate the weight it has in the index in order to achieve full replication. Strategies based on sampling are likely to result in purchasing the included stock as well, even if the portfolio weight might differ from the one in the index.

Until October 1989, Standard and Poor's announced the inclusion of a new stock in the S&P 500 after the close, the change becoming effective by the following open. After October 1989, Standard and Poor's switched to preannouncing changes in the S&P 500 usually five days before the inclusion. The aim of this new practice was to ease post-announcement order imbalances for companies added to the index. As mentioned in the introduction, trading activity patterns point at index trackers stepping into the market the day after the replacement (before October 1989) or during the effective day of inclusion (after October 1989). While the announcement timing of index replacements does not affect passive managers' behaviour, it has relevant effects on other market participants. Under the old announcement practice, indexers would step into the market at the open immediately after the Standard and Poor's public announcement. It follows that there would not be profitable speculation unless the announcement is anticipated by some traders. However inclusions in the S&P 500 do not seem to be predictable due to the above mentioned Standard and Poor's discretionality in selecting stocks for the index, casting doubts on investors anticipating replacements. Singal (2003) provides anecdotal evidence on failures in predicting index changes. On the other hand the new preannouncement practice makes the S&P game attractive. Beneish and Whaley (1996) show how such a strategy yields significant abnormal returns, even accounting for transaction costs. Blume and Edelen (2004) report a 19.2 basis point yearly return associated with the S&P game over their 1995-2000 sample. Early-trading profitability is also documented in Singal (2003) for inclusions between January and July 2002. These findings support the argument that 'an investor who requires that an indexer maintain tracking errors of just a few basis points a year is giving up additional returns. [...] Forgoing these additional returns can be viewed as an agency cost in delegating investment decisions' [Blume and Edelen (2004), p. 3].

2 The benchmark model for index replacements

2.1 Model setup

2.1.1 Asset markets and changes announcements

We develop a two period sequential trading game along the lines of Kyle (1985). Trading takes place at two dates $t = 1, 2$ and the market operates as a batch auction. There are two traded assets: a riskfree asset whose net payoff is normalized to zero, and a risky asset with final liquidation value $f \sim N\left(p_0, \sigma^2_{f,0}\right)$ whose realization occurs after the second trading round. The trading dates capture the timing in index replacements as follows. Before trading takes place at $t = 1$ the authorities announce the change in the index composition. Further to the stock(s) added to/removed from the index, it is announced that the change is effective after the second trading round.

2.1.2 Agents

There are three types of agents in the market: an insider, a market-maker and noise (or uninformed) traders. Both the insider and the market maker are risk neutral. At each date the trading process is modeled as a two-stage game: in the first stage the insider and the uninformed traders submit their orders to the market maker; in the second stage the market maker determines the price at which the market is cleared. The insider submits his orders $\{x_t\}_{t=1,2}$ at both dates. The noise in the market comes from two different sources: liquidity traders and passive funds. The main difference between these two groups is that passive funds enter the market at date 2 only, while liquidity traders submit their orders $\{u_t\}_{t=1,2}$ at both dates. More specifically we assume that $u_1 \sim N\left(0, \sigma^2_{u_1}\right)$ and $u_2 \sim N\left(0, \sigma^2_{u_2}\right)$, with (f, u_1, u_2) mutually independent. Further to the liquidity traders there are indexers active at the second trading round. Passive trades are denoted by $z_2 \sim N\left(\bar{z}_0, \sigma^2_{z,0}\right)$ and are orthogonal to the other random variables f, u_1 and u_2.[2] The joint distribution of (f, u_1, u_2, z_2) is common knowledge among market participants before the game starts. The date t aggregate order flow $\{\omega_t\}_{t=1,2}$ is given by $\omega_1 \equiv x_1 + u_1$ and $\omega_2 \equiv x_2 + u_2 + z_2$ respectively.

The noise trading specification slightly departs from the standard assumptions, and the way we model z_2 aims at capturing several aspects in passive managers' behaviour. First of all we allow liquidity trades' variance $\left\{\sigma^2_{u_t}\right\}_{t=1,2}$ to vary over time. Later on we compare equilibrium parameters under different market conditions, and use this flexibility in order to keep the overall uninformed variance constant through time. Secondly, replicating strategies are not based on any information related to the asset fundamental value. As mentioned in the introduction, every time the index composition changes, passive managers should rebalance their portfolios. As such indexers can be regarded as uninformed traders submitting orders due to changes in the benchmark they replicate. In the third place, pegged funds' performance is assessed via tracking error procedures, and in our model the index replacement is effective after date 2. Optimizing the fund's performance (relative to the index) on a daily basis thus leads passive managers to rebalance on the inclusion day rather than immediately after the announcement, i.e. at date 2 rather than at date 1 in our model, consistently with the evidence in Beneish and Whaley (1996) and Blume and Edelen (2004). Eventually, we consider a shift in expected uninformed trades between the two dates via the term $\bar{z}_0$, which reflects passive funds stepping into the market at the second round. In general we relate the magnitude of this shift to the weight pegged funds have relative to other liquidity traders.

2.1.3 Information structure

Within our strategic trading setup –as well as in the various extensions to Kyle (1985)– uncertainty among market participants is captured by two random variables: the final liquidation value and uninformed trades. We therefore distinguish the information related to these variables as fundamental and non-fundamental respectively, along the same lines of Madrigal (1996). Let Φ^I_t and Φ^M_t denote the insider's and market maker's information set at time t. At each trading round the market maker observes the aggregate order flow, such that $\Phi^M_t = \{\omega_s, s \leq t\}$. The price in period t is assumed to satisfy the semi-strong efficiency condition:[3]

$$p_t = E\left(f \middle| \Phi^M_t\right) \qquad , \qquad t = 1, 2 \tag{1}$$

After each trading round, the price becomes common knowledge among market participants. The insider possesses superior information regarding both the asset's fundamental value and other non-fundamental aspects of the market. The insider is aware of the final liquidation value before the trading game starts.[4] Further to this fundamental information, the insider knows the quantity submitted by liquidity traders –but not passive funds– at date t before filling his order x_t, i.e. $\Phi_1^I = \{f, u_1\}$ and $\Phi_2^I = \Phi_1^I \cup \{p_1, u_2\}$. Thus the insider possesses long-lived fundamental information as well as short-lived non-fundamental information. The information structure is summarized in table I.

Our information structure departs from the existing literature in the following aspects. As in Foster and Viswanathan (1994) and Kyle (1985) the insider is endowed with long-lived information on the final payoff f. Further, in our game the insider is also aware of the contemporaneous liquidity trades, thus making our setup closer to Rochet and Vila (1994) and Yu (1999). In the absence of date 2 pegged trades our trading game reduces to a two-period version of Rochet and Vila (1994) or, equivalently, to the game in Yu (1999) with non-distorted information on uninformed trades. However the entry of passive funds moves the insider away from complete knowledge about noise trades at the second trading round. Therefore our specification resembles a two-period version of Yu (1999) with time-varying quality of the insider's signal about uninformed trades. The insider can be thought of as a broker possessing both fundamental and non-fundamental information, the latter being captured by the liquidity trades $\{u_t\}_{t=1,2}$ he executes at both dates [an analogous interpretation can be given to the insider in both Rochet and Vila (1994) and Yu (1999)].

2.2 Equilibrium construction and description

We focus on linear equilibria for our trading game. For the insider we denote the period t profit by $\left\{\pi_t^I = \pi_t^I(\omega_s, s \leq t)\right\}$, i.e. $\pi_1^I = x_1(f - p_1(\omega_1))$ and $\pi_2^I = x_2(f - p_2(\omega_1, \omega_2))$. A Bayes-Nash equilibrium (BNE) is defined by a set of linear functions $\{x_t(\cdot), p_t(\cdot)\}_{t=1,2}$ such that the following conditions hold:

1. *insider's profit maximization*: the insider chooses x_1 to maximize total profits

$$E\left[\pi_1^I(\omega_1) + \pi_2^I(\omega_1, \omega_2) \middle| \Phi_1^I\right] , \tag{2}$$

   given that x_2 maximizes second period profits

$$E\left[\pi_2^I(\omega_1, \omega_2) \middle| \Phi_2^I\right] . \tag{3}$$

2. *market efficiency*: the market maker sets prices according to equation (1), i.e.

$$p_1 = E\left(f \middle| \Phi_1^M\right) \tag{4}$$

$$p_2 = E\left(f \middle| \Phi_2^M\right) \tag{5}$$

Proposition 1 *Let the following conditions hold:*

$$a_1 = \frac{2\lambda_2 - \lambda_1}{\lambda_1(4\lambda_2 - \lambda_1)} \quad ; \quad b_1 = a_1\lambda_1 \quad ; \quad \lambda_1 = \frac{a_1\sigma_{f,0}^2}{a_1^2\sigma_{f,0}^2 + (1 - b_1)^2\sigma_{u_1}^2}$$

$$a_2 = \frac{1}{2\lambda_2} \quad ; \quad \lambda_2 = \frac{\sigma_{f,1}}{\left(\sigma^2_{u_2} + 4\sigma^2_{z,0}\right)^{1/2}}$$

where $\sigma_{f,1}$ is the fundamental value residual variance after the first trading round. Then there exists a linear BNE in which strategies and prices are of the form

$$x_1 = a_1 (f - p_0) - b_1 u_1 \tag{6}$$

$$x_2 = a_2 (f - p_1) - u_2/2 \tag{7}$$

$$p_1 = p_0 + \lambda_1 \omega_1 \tag{8}$$

$$p_2 = p_1 + \lambda_2 (\omega_2 - \bar{z}_0) \tag{9}$$

Furthermore, if the following condition holds

$$\lambda_1 (4\lambda_2 - \lambda_1) > 0 \tag{10}$$

the equilibrium is unique.

The equilibrium strategies in Proposition 1 have the following interpretation. Before trading takes place at date 1, the market maker's forecast of the random variables (f, u_1) coincides with the unconditional means $(p_0, 0)$. Thus at time 1 the insider trades on the market maker's misperception of the final liquidation value $(f - p_0)$, and current liquidity trades u_1. The insider places a positive weight $(a_1 > 0)$ and a negative one $(-b_1 < 0)$ respectively on the former and the latter forecast error. After the first trading round, the market maker updates his beliefs about the liquidation value to p_1. Since the first period aggregate order flow does not contain any information about z_2, the market maker doesn't learn anything about passive trades. As a consequence, date 2 passive funds' conditional mean coincides with its unconditional counterpart $\bar{z}_0$. In section 3 we discuss how non-fundamental speculation arising from the S&P game modifies the latter feature. Thus at date 2 the insider trades on the market maker's misperception of the final liquidation value $(f - p_1)$, and current liquidity trades u_2. The weights on the market maker's errors are consistent with the ones prevailing during the first trading round: positive on $(f - p_1)$ –since $a_2 > 0$–, and negative on u_2. The insider's trading intensities in eqs. (6) and (7) are consistent with the previous literature: date t trading aggressiveness on the fundamental information –as captured by a_1 and a_2– is positive like in Kyle (1985). Moreover at every batch auction the insider trades against current uninformed orders like in Yu (1999), given that the intensities on u_t are negative. Finally at date 2 the insider offsets half of the (contemporaneous) liquidity trades like in Rochet and Vila (1994). Equilibrium prices have the usual linear form with λ_t capturing the price response induced by unit-size changes in the aggregate order flow. Equivalently, $1/\lambda_t$ is the date t market depth (or liquidity[5]): large values for λ_t imply that prices are extremely sensitive to changes in the order flow, which occurs in illiquid markets. Finally, we define the fundamental value residual variance as $\sigma^2_{f,t} = var\left(f | \Phi^M_t\right)$, i.e. the final payoff variance after t rounds of trading. $\left\{1/\sigma^2_{f,t}\right\}_{t=1,2}$ therefore gives the speed at which private information about f is revealed to the market, and it can be thought of as measuring market efficiency.

The equilibrium for our trading game is investigated in figures 2–6. We normalize the initial fundamental volatility setting $\sigma^2_{f,0} = 1$, and consider uninformed trades' uncertainty at both dates to be equal to the fundamental variance, i.e. $\sigma^2_{u_1} = 1$ and $\sigma^2_{u_2} + \sigma^2_{z,0} = 1$. Further we define $k_I \equiv \sigma^2_{u_2}/\sigma^2_{z,0}$, and refer to k_I as the quality of the insider's non-fundamental information (or equivalently the insider's informational advantage[6]). In fact $\sigma^2_{u_2} = \sigma^2_{u_2}\left(\sigma^2_{u_2} + \sigma^2_{z,0}\right)^{-1}$ can be interpreted as the share of date 2 uninformed orders channeled by the insider to the market maker. Therefore k_I denotes the insider's informational advantage (relative to the market maker) with respect to date 2 noise trades: high values for k_I correspond to small passive funds' volatility, which in turn implies that $\sigma^2_{u_2}$ captures most of the noise trading volatility at date 2. For example if $k_I = 1$, the variance of the uninformed trades observed by the insider is half of the entire noise trading variance faced by the market maker at date 2. We consider several values[7] for k_I and plot the parameters in Proposition 1 in figures 2–6 (solid line). The dashed line corresponds to a two-period Rochet and Vila (1994) trading game (henceforth RV) in which passive funds are absent at date 2, such that the insider is aware of current liquidity trades at both dates. Clearly, our trading game resembles RV when k_I is large, or equivalently when $\sigma^2_{z,0}$ is negligible relative to $\sigma^2_{u_2}$. [8]

We plot insider's intensities a_1, b_1 and a_2 in figure 2, while values for λ_1 and λ_2 are reported in figure 3. Since λ_t measures the adverse selection costs faced by the market maker at round t, it is not surprising that λ_2 increases in the insider's advantage k_I (figure 3-panel B). Recall from the equilibrium strategy (7) that the second period trading intensity on current liquidity orders does not depend on k_I, and is equal to $-1/2$ as in RV and Yu(1999). Therefore at date 2 the insider's advantage $\sigma^2_{u_2}/\sigma^2_{z,0}$ affects the trading aggressiveness a_2 only, which is shown to be decreasing in k_I (figure 2-panel C). This is due to the mentioned finding that date 2 liquidity –as measured by $1/\lambda_2$– decreases with k_I.

Turning to date 1 parameters, we note that both the trading intensities a_1 and b_1 increase in k_I (figure 2-panel A and B respectively). The bottom line of figure 2 is that the insider increases his trading intensity with respect to *both* sources of information together with his informational advantage. This means that the insider incorporates more information on both f and u_1 in his trade x_1 as k_I increases: since the insider anticipates the negative relationship between k_I and date 2 liquidity, he increases his aggressiveness with the information quality during the first trading round. The market maker's reaction is to make date 1 liquidity decreasing with k_I as well (figure 3-panel A).

Following Admati and Pfleiderer (1988), we decompose date t trading volume into its components. The contribution of the insider to the expected total volume is therefore given by (see section 4 and the appendix for further details[9])

$$V_t^I \equiv \sqrt{\frac{var\left(x_t\right)}{2\pi}}, \quad t = 1, 2$$

We plot V_1^I and V_2^I in figure 4 (panel A and B respectively). Consistently with the previous analysis for the trading intensities, V_1^I increases (resp. V_2^I decreases) with the informational advantage. The residual variances $\sigma^2_{f,1}$ and $\sigma^2_{f,2}$ are depicted in figure 5 (panel A and B respectively), as well as the ratio $\sigma^2_{f,1}/\sigma^2_{f,2} = \left(\frac{1/\sigma^2_{f,2}}{1/\sigma^2_{f,1}}\right)$ which captures the market efficiency dynamics through time (panel

C). As shown in the appendix, $\sigma^2_{f,t}$ is negatively related to the insider intensity a_t and the price sensitivity λ_t. Therefore, date 1 efficiency $1/\sigma^2_{f,1}$ increases in k_I, since both a_1 and λ_1 increase with the informational advantage. Furthermore $\sigma^2_{f,1}/\sigma^2_{f,2}$ does not depend on k_I, which means that the positive relation between k_I and date 2 efficiency in panel B is entirely due to the increase in date 1 market efficiency.

The insider's unconditional expected profits are depicted in figure 6 (panel A). Unlike other variables, ex-ante gains are non-monotonic in k_I. At a first sight this might seem surprising, as one would expect insider's profits to increase together with the information quality $\sigma^2_{u_2}/\sigma^2_{z,0}$. On the other hand figure 6 suggests that the insider is (ex-ante) worse off with more precise information whenever k_I is below some threshold value (in figure 6 the minimum value is 0.8527 corresponding to $k_I = 1.4$). Yu (1999) (figure 2, p. 92) documents a similar behaviour and notes that the insider is not necessarily better off with more precise non-fundamental information. Furthermore he shows that a U-shaped curve for ex-ante gains is more likely to emerge when the number of batch auctions is small, like in our model. Therefore the pattern in figure 6 is in line with results in Yu (1999). Comparison between figures 3 and 6 suggests that the insider expects to lose out to a lower date 1 market depth as his information becomes more precise until a threshold level: within this region, liquidity decreases very rapidly with k_I such that profits decrease with the informational advantage.

3 Non-fundamental speculation and the S&P game

3.1 Model setup

In what follows we explicitly introduce a role for purely non-fundamental speculation (the S&P game) within the setup outlined in section 2. As explained in the introduction, what lies behind this speculative opportunity is privileged information about passive assets, together with index changes preannouncement. For what is not mentioned in this subsection, we maintain the assumptions in subsection 2.1.

3.1.1 Agents

We introduce another risk-neutral informed trader, the (non-fundamental) speculator. While the insider receives both fundamental and non-fundamental information, the speculator is endowed with non-fundamental information only, as specified later in this subsection. At both trading dates the speculator submits orders $\{y_t\}_{t=1,2}$ to the market maker. The aggregate order flow therefore becomes $\omega_1 \equiv x_1 + y_1 + u_1$ and $\omega_2 \equiv x_2 + y_2 + u_2 + z_2$.

3.1.2 Information structure

Let Φ^S_t denote the speculator's information set at time t. The speculator knows the demand submitted by a subset of passive funds before the game starts. As a consequence we decompose the passive industry demand z_2 into two components v and w: the former aggregates the trades known by the speculator, while the latter groups the demand submitted by other passive funds:

$$z_2 = v + w \sim N\left(\bar{v}_0 + \bar{w}, \sigma_{v,0}^2 + \sigma_w^2\right)$$

where $\bar{v}_0 = E(v)$, $\bar{w} = \bar{z}_0 - \bar{v}_0$, $\sigma_{v,0}^2 = var(v)$ and $\sigma_w^2 = \sigma_{z,0}^2 - \sigma_{v,0}^2$. The speculator's information sets are given by $\Phi_1^S = \{v\}$ and $\Phi_2^S = \Phi_1^S \cup \{p_1\}$, implying that the speculator is endowed with (long-lived) non-fundamental information.

The focus on the role of purely non-fundamental speculation [not considered in Kyle (1985), Foster and Viswanathan (1994), RV and Yu (1999)] closely resembles the analysis in Madrigal (1996). However we depart from Madrigal (1996) in several aspects. First of all the speculator is endowed with superior knowledge about a fraction of *future* –rather than past– uninformed trades. It follows that in our model the speculator exploits his advantage trading at both dates, while in Madrigal (1996) he enters the picture at date 2 only. In the second place the speculator acts as a *monopolist* on his privileged information in the first trading round, and competes with the insider at the date 2, whereas Madrigal (1996) focuses on the latter feature only.

Most of the literature on asymmetries in financial markets is concerned with fundamental information, and knowledge about the final payoff is widely accepted as arising from analysts' research activity as well as confidential discussions. On the other hand informational advantages on uninformed orders can be traced to brokers engaging in proprietary –or dual– trading. Brokers both execute trades on behalf of their (liquidity) customers and fill in orders on their own account. As a consequence, brokers can engage in dual-trading based on the ability to observe their clients' orders. In Madrigal (1996) the speculator channels liquidity orders in the first round and then uses this information (together with the price set by the market maker) to forecast the final liquidation value.[10] Similarly in Foucault and Lescourret (2003) the speculator is not endowed with fundamental information, but he observes contemporaneous liquidity trades before submitting his order. This leaves open the question as how our speculator gathers *more precise* information about indexed assets *ahead* of other market participants. As a matter of fact one might object that preannouncing index changes conveys information to the whole market about passive funds' entry at the inclusion. For example one might use publicly available data on pegged funds capitalization [like the S&P survey (2003)] and infer the realization of z_2. However this estimate would be accurate only in case passive funds track the index via full replication, i.e. buy all the stocks in the index and in the same proportion, and if funds do not experience inflows and outflows during the year –which is rather unlikely.[11] Even though in principle full replication allows to track the index very closely, it entails substantial administrative costs due to the number of stocks to be bought/sold and, consequently, the number of dividends to be handled. Given that these costs might dampen passive funds' performance and result in larger tracking errors, indexers can resort to other strategies such as stratified sampling or optimization techniques. Based on the Morningstar database, Blume and Edelen (2004) report that the vast majority of funds indexed to the S&P 500 hold roughly all the stocks included in the index. However, as the authors suggest, this does not necessarily imply that all the funds implement full replication techniques. For instance Blume and Edelen (2004) argue that the increase in the tracking error for the Vanguard 500 Index Fund –one of the largest passive funds– after 1998 is inconsistent with full replication. This example suggests that knowledge about the tracking procedures actually implemented by individual passive managers is inherently difficult to gather, and as a consequence the realization of z_2 cannot be regarded as public information. Thus we consider public data on passive industry capitalization as providing the *expected* passive funds' orders $\bar{z}_0 = \bar{v}_0 + \bar{w}$ to the whole market, and reasonably conceive that some traders are endowed with superior information about z_2. For example, a broker might learn something about the replication technique implemented by a given fund manager because he previously executed his trades. Alternatively, an indexer can direct his order to a broker under the agreement that execution occurs at a specified future date. Both these cases would generate non-fundamental informational advantages consistent with our speculator's information sets. Confidential discussions with passive fund managers would fit into the same specification and result in long-lived information on future uninformed trades as well.

As a consequence of these assumptions, our trading game inherits several interesting features. When trading at date 1 both the insider and the speculator impound their information into orders x_1 and y_1. Time 1 noise trades u_1 keep the aggregate order flow away from fully revealing both the insider's information (f, u_1) as well as the speculator's information v. After observing the aggregate order flow ω_1, the market maker forms an estimate $\bar{v}_1$ of future passive trades:

$$\bar{v}_1 = E\left(v \middle| \Phi_1^M\right) \tag{11}$$

Note that our trading game allows the market maker to update his beliefs on (a fraction of) the second period uninformed trades as well as on the final liquidation value –through the price p_1– and to use these updates when setting the market clearing price at date 2. The existing literature concentrates on the market maker's inference on the final payoff only: posteriors on noise trades are not considered, since informed agents are endowed with signals on either current or past uninformed orders. In Yu (1999) the insider receives at each date t (a signal of) contemporaneous noise trades. Nonetheless the independence through time of liquidity-motivated orders prevents the market maker from extracting any signal on time $t+1$ noise trading based on the order flow received at time t. A similar argument holds for both Foucault and Lescourret (2003) and Madrigal (1996).

The information on v, together with the price realization p_1, allows the speculator to form a superior estimate of f relative to the market maker. After the first trading round, the speculator nets out the insider's and liquidity traders' demand out of the aggregate order flow –due to price linearity in ω_1– and extracts a signal s of the fundamental value that is more precise than the market maker's expectation:

$$s = E\left(f \middle| \Phi_2^S\right) = E\left(f \middle| x_1 + u_1\right) \tag{12}$$

Therefore the speculator can profit on the difference $(s - p_1)$ because noise trades in period 2 will prevent the order flow from revealing the speculator's information. The insider reacts to the speculator's presence incorporating an estimate of s when trading in the first round. A similar signal extraction problem and the incentives for the insider to manipulate the first period price are analyzed in Madrigal (1996).

On the other hand the insider infers the speculator's information about v after observing the first period price. Note, however, that the information structure enables the insider to know the realization (of a fraction) of passive funds' trades v, while the speculator extracts only a signal s of the fundamental value f. As such our model displays a hierarchical information structure[12]during the second trading round. Borrowing the terminology in Foster and Viswanathan (1994) the insider is the 'better informed trader' and the speculator is the 'lesser informed trader' at date 2[13]. The information structure is summarized in table I, while figure 10 presents the time line underlying our trading game.

3.2 Equilibrium construction and description

Let date t speculator's profits be defined along the same lines as in section 2, i.e. $\pi_1^S = y_1 (f - p_1 (\omega_1))$ and $\pi_2^S = y_2 (f - p_2 (\omega_1, \omega_2))$. A BNE for our trading game is given by a set of linear functions $\{x_t (\cdot), y_t (\cdot), p_t (\cdot)\}_{t=1,2}$ satisfying the insider's profit maximization [see conditions $(2, 3)$], market efficiency [see conditions $(4, 5)$] and the following:

speculator's profit maximization: the speculator chooses y_1 to maximize total profits

$$E\left[\pi_1^S (\omega_1) + \pi_2^S (\omega_1, \omega_2) \middle| \Phi_1^S\right] \ , \tag{13}$$

given that y_2 maximizes second period profits

$$E\left[\pi_2^S (\omega_1, \omega_2) \middle| \Phi_2^S\right] \ . \tag{14}$$

Requirements $(13, 14)$ amount to look for a pair of linear functions $y_1 (\cdot)$ and $y_2 (\cdot)$ such that $y_1 = y_1 (v)$ and $y_2 = y_2 (s, v)$, where s is defined in eq. (12). Recall that within our informational structure the insider knows –prior to trading at time 2– the signal s that the speculator extracts from p_1. As a consequence, when trading at date 1 the insider keeps into account the effect of his order on the speculator's estimate of the final liquidation value $E\left(s \middle| \Phi_1^I\right) = E\left(E\left(f \middle| \Phi_2^S\right) \middle| \Phi_1^I\right)$. The insider's trading strategies are therefore given by a pair of linear functions $x_1 (\cdot)$ and $x_2 (\cdot)$ such that $x_1 = x_1 \left(f, u_1, E\left(s \middle| \Phi_1^I\right)\right)$ and $x_2 = x_2 (f, u_2, s, v)$. At $t = 1$ the insider trades on the speculator's (expected) mispricing, i.e. the difference between $E\left(s \middle| \Phi_1^I\right)$ and the true liquidation value, as well as on the market maker's mispricing, i.e. the difference between the realization f and p_0. This amounts to conjecture the following form for x_1:

$$x_1 = \alpha (f - p_0) + \beta u_1 + \gamma \left(E\left(s \middle| \Phi_1^I\right) - f\right) \tag{15}$$

Note that the insider's date 1 trade depends on the (estimate) of the speculator's conjecture of the final liquidation value, which depends itself on the insider's first period trade. Thus one needs to solve for $E\left(s \middle| \Phi_1^I\right)$ and then verify the consistency between the resulting expression for x_1 and the speculator's belief s.[14]

Proposition 2 *Let the following conditions hold:*

$$a_1 = d^{-1} \left(1 - \frac{2\lambda_1 + \phi - 2\lambda_2 \mu}{6\lambda_2}\right) \quad ; \quad b_1 = 1 - d^{-1}\lambda_1 \quad ; \quad d = 2\lambda_1 - \frac{(2\lambda_1 + \phi - 2\lambda_2 \mu)^2}{18\lambda_2}$$

$$C_1 = D^{-1}\left(\frac{\lambda_1 - \lambda_2\mu}{9}\right) \quad ; \quad D = \lambda_1 - \frac{(\lambda_1 - \lambda_2\mu)^2}{9\lambda_2}$$

$$a_2 = \frac{1}{2\lambda_2} \quad ; \quad C_2 = \frac{2a_2}{3}$$

$$\lambda_1 = \frac{a_1\sigma_{f,0}^2}{a_1^2\sigma_{f,0}^2 + (1-b_1)^2\sigma_{u_1}^2 + C_1^2\sigma_{v,0}^2}$$

$$\lambda_2 = \Sigma^{-1}\left(-2\sigma_{fv,1} + \left[4\sigma_{fv,1}^2 + \left(9\sigma_{f,1}^2 - \sigma_{s,1}^2\right)\Sigma\right]^{1/2}\right) \quad ; \quad \Sigma = 9\sigma_{u_2}^2 + 36\sigma_w^2 + 4\sigma_{v,1}^2$$

where $\sigma_{f,1}^2$*,* $\sigma_{s,1}^2$ *and* $\sigma_{fv,1}$ *are residual variances after the first trading round (defined in appendix B). Then there exists a linear BNE in which trading strategies and prices are of the form*

$$x_1 = a_1(f - p_0) - b_1u_1 \tag{16}$$

$$x_2 = a_2(f - p_1) - u_2/2 - C_2(s - p_1)/2 - (v - \bar{v}_1)/3 \tag{17}$$

$$y_1 = C_1(v - \bar{v}_0) \tag{18}$$

$$y_2 = C_2(s - p_1) - (v - \bar{v}_1)/3 \tag{19}$$

$$p_1 = p_0 + \lambda_1\omega_1 \tag{20}$$

$$p_2 = p_1 + \lambda_2(\omega_2 - \bar{z}_1) \tag{21}$$

where s *is the speculator's belief on* f *conditional on* Φ_2^S*, and* $\bar{z}_1 = \bar{v}_1 + \bar{w}$ *is the market maker's belief on* v *conditional on* Φ_1^M

$$s = p_0 + \phi(x_1 + u_1) \tag{22}$$

$$\bar{v}_1 = \bar{v}_0 + \mu(x_1 + y_1 + u_1) \tag{23}$$

The updating coefficients in (22),(23) *are given by*

$$\phi = \frac{a_1\sigma_{f,0}^2}{a_1^2\sigma_{f,0}^2 + (1-b_1)\sigma_{u_1}^2} \tag{24}$$

$$\mu = \frac{C_1\sigma_{f,0}^2}{a_1^2\sigma_{f,0}^2 + (1-b_1)\sigma_{u_1}^2 + C_1\sigma_{f,0}^2} \tag{25}$$

Furthermore, if the following condition holds

$$2\lambda_1 - \frac{(2\lambda_1 + \phi - 2\lambda_2\mu)^2}{18\lambda_2} > 0 \tag{26}$$

the equilibrium is unique.

The equilibrium strategies in Proposition 2 have the following interpretation. Before trading takes place at date 1, the market maker's forecast of the random variables (f, u_1, z_2) coincides with their unconditional mean $(p_0, 0, \bar{z}_0)$. Thus at time 1 the insider trades on (1) the market maker's misperception of the fundamental value $(f - p_0)$ as well as current liquidity trades u_1, and (2) the speculator's (expected) forecast error $E\left(s|\,\Phi_1^I\right)-f$. Therefore the presence of the speculator results in the insider manipulating his trades (relative to the equilibrium in Proposition 1) as a reaction to the speculator's extraction of the signal s. On the other hand the speculator trades on his informational advantage –with respect to both the market maker and the insider– $(v-\bar{v}_0)$. After observing ω_1, the market maker forms the posterior $\bar{v}_1$ according to (11), such that the passive industry conditional mean is given by $\bar{z}_1=\bar{v}_1+\bar{w}$. Therefore, after the first trading round, the market maker's update on (f,z_2) is given by $(p_1,\bar{z}_1)$.[15] A similar argument to the one used for first period orders ensures that at date 2 the insider trades on (1) the market maker's misperception of the final liquidation value $(f-p_1)$ as well as current liquidity trades u_2 and (2) the speculator's forecast error $(s-f)$. The linearity assumption further implies that the insider trades on the market maker's misperception of the passive orders $(v-\bar{v}_1)$. Similarly, the speculator trades on his informational advantage (with respect to the market maker only) captured by the terms $(v-\bar{v}_1)$ and $(s-p_1)$.

As mentioned in the introduction, the S&P game consists in front-running index funds. In our trading framework, the S&P game would translate into the speculator buying at date 1 whenever the market underestimates the realization of v, and subsequently selling at date 2. Conversely, the speculator should sell at date 1 whenever the market overestimates pegged trades, i.e. $v<\bar{v}_0$, and buy back at date 2. Since the speculator trades against $(v-\bar{v}_1)$ [see equation (19)] during the second trading round, the occurrence of the S&P game depends on the sign of the coefficient C_1 in (18). In fact, we show the following:

Corollary 1 *In equilibrium the speculator plays the S&P game, i.e.* $C_1>0$.

The effect of the speculator on equilibrium parameters is analyzed in figures 2–9. Along the same lines we used in section 2 for the insider's non-fundamental advantage k_I, we let $k_S\equiv\sigma_{v,0}^2/\sigma_w^2$ be the precision of the speculator's information. Higher values for k_S imply that the speculator is able to make sharper inference relative to the rest of the market. We set k_S equal to 0.33, 1 and 3, and refer to these three cases as low, medium and high (speculator's) informational advantage respectively. In figures 2–9 we plot equilibrium values corresponding to $k_S=0.33$, 1 and 3 with circles, squares and triangles respectively. The other underlying parameters are set accordingly to section 2.

Trade aggressiveness and market liquidity

When trading at date 1, the speculator places a positive weight ($C_1>0$ from Corollary 1) on the market maker's initial forecast error $(v-\bar{v}_0)$ and then reverses his strategy at date 2 (with intensity equal to $-1/3$). From the expression for y_2 the speculator offsets one third of the (current) passive trades at date 2. This finding is consistent with RV and the trading game in Proposition 1, keeping into account that in Proposition 2 both the speculator *and* the insider trade on the same information v at the second round, thus offsetting 2/3 of the date 2 indexed trades. As is known, Cournot competition on the indexers' order v between the two informed agents results in higher aggregate intensity on v. Furthermore from eq. (17) –as well as from RV and eq. (7)–, the insider acts as a monopolist on date 2 liquidity trades, thus offsetting 50% of u_2. Consider the case in which the demand from passive funds known by the speculator is large relative to its unconditional value, i.e. $(v - \bar{v}_0) >> 0$. Other things equal, this would determine an increase in p_2 thanks to the linear pricing rule (21) and $\lambda_2 > 0$. Since the speculator observes the realization v one period ahead of the market maker, he forecasts the increase in p_2 induced by unexpectedly large passive trades v. As a reaction the speculator trades positively on $(v - \bar{v}_0)$ at the first date and profits from the price difference between the two trading rounds. Note the difference in the trading intensity on v between the two dates (figure 7): the trading aggressiveness on v increases through time, since C_1 is always smaller than 1/3. This arises from the fact that the information about v impounded by the speculator's trades at date 1 allows the market maker[16] to make a sharper inference about the second period indexed assets via the posterior $\bar{v}_1$. Furthermore C_1 decreases with k_S and increases with k_I. The first finding is consistent with the speculator trying to keep his advantage in forecasting the final liquidation value with respect to the market maker, thus incorporating less information whenever k_S is large. The second finding is related to the insider's behaviour at the first trading round. From Proposition 1 we have that the insider trades more aggressively on both $(f - p_0)$ and u_1 the larger is his advantage k_I. Thus the speculator can hide more of his information to the market maker as k_I increases, and as a consequence C_1 increases in the insider's advantage.

The insider reacts to the presence of the speculator increasing his trading intensities at date 1 (figure 2-panel A and B) relative to Proposition 1. It is worth noting that while a_1 monotonically increases in k_I –as it happens without the speculator– b_1 decreases with k_I when the speculator's advantage is relatively high ($k_S = 3$), while it increases in k_I in the absence of the speculator as well as for low values of k_S. In order to understand this finding note from figure 3 that the speculator's entry decreases liquidity at both dates. The market maker faces more severe information asymmetries than in the absence of the speculator, and market depth is reduced: λ_t increases with both k_I and k_S in all cases but when the speculator's advantage is high (panel A). When this occurs, both λ_1 and b_1 decrease in k_I. Other things equal, the speculator decreases date 1 liquidity. The insider reacts trading less aggressively on u_1 in order to counterbalance the negative effect on liquidity due to the speculator's trades. The net result is that date 1 liquidity improves with k_I due to the insider's reaction when the speculator's advantage is high. Finally, note from equation (17) that the insider trades at date 2 in the opposite direction of the signal extracted by the speculator as it occurs in Madrigal (1996).

Trading volume

The insider's contribution to the total trading volume is defined along the same lines as in section 2, and is shown in figure 4. V_1^I increases due to the speculator's entry. This is due to the externality imposed by the speculator, which results in the insider trading more aggressively in order to exploit his fundamental advantage before the speculator makes his superior inference. As a consequence V_1^I increases with the speculator's advantage k_S. At the second round, the insider's intensity on his fundamental information decreases with k_S as in figure 2. Furthermore the insider trades in the opposite direction of the speculator's aggressiveness on his misperception $(s - p_1)$, which is again decreasing in k_S (see figure 7). The overall effect on trading volume is that V_2^I decreases with k_S as in figure 4 (panel B).

Similarly to V_t^I, the speculator's volume V_t^S is defined as

$$V_t^S \equiv \sqrt{\frac{var\,(y_t)}{2\pi}}, \quad t = 1, 2$$

and it is shown to be increasing in the informational advantage k_S in figure 8. At a first glance this finding might seem inconsistent with the pattern for the trading intensities C_1 and C_2 in figure 7. However in the appendix we show that V_t^S depends positively on both the trading intensity C_t and the passive funds' variance $\sigma_{v,0}^2$, the latter dependence implying that the quantity traded by the speculator increases in $\sigma_{v,0}^2/\sigma_w^2$. This means that larger uncertainty on the passive trades' volatility offers more opportunities to hide the speculator's informational advantage, thus justifying the pattern in figure 8.

Market efficiency

Market efficiency improves due to the speculator's entry: both $1/\sigma_{f,1}^2$ and $1/\sigma_{f,2}^2$ go up with respect to Proposition 1 due to the increased trading aggressiveness of the insider in the presence of the speculator. The insider impounds more information about the final liquidation value in order to anticipate the speculator's signal extraction, and as a result residual variances are lower in the presence of the speculator. From panel C in figure 5 it emerges that improvements in market efficiency come mainly from the first trading round. Recall that in the absence of the speculator, efficiency gains $\sigma_{f,1}^2/\sigma_{f,2}^2$ are not affected by k_I, while they depend on both informational advantages in Proposition 2: the higher is the quality of the speculator's information, the lower is the efficiency ratio. In other words the speculator's precision reduces efficiency gains over time.

Ex-ante incentives

As for the insider's expected profits (figure 6-panel A) the speculator reduces the insider's motives to trade, like in Madrigal (1996). We note however that in Madrigal (1996) the speculator acts as a free-rider on the insider's information extracting the signal s, which is a better forecast of the final liquidation value than the price set by the market maker. In our game the speculator is able to extract the signal s *at the additional cost* of revealing his privileged information both to the insider (which knows the realization v after the first trading round) and to the market maker (which forms the posterior $\bar{z}_1$ on passive trades after observing the order flow ω_1). Therefore one might expect that the insider makes higher profits in the presence of the speculator due to the additional information about v. However figure 6 shows that this is not the case: the negative externality imposed by the speculator on the insider, i.e. the loss due to the speculator's signal gathering activity, dominates the benefit of knowing v in addition to liquidity trades u_2.

As for the speculator's ex-ante incentives, they increase with his own advantage k_S and decrease with k_I (figure 6-panel B). This behaviour hinges on the very same trading motives for the insider. Whenever the quality of the insider's information is relatively high, the insider impounds more information on the fundamental value into his orders. As a result p_1 improves its precision as a forecasting tool for the final payoff. The speculator's inferential ability in extracting the signal s reduces relative to the improved market maker's forecast, and speculator's ex-ante gains drop.

In summary the effects of the S&P game are as follows. The presence of the speculator makes the insider trade more aggressively on both the fundamental value and the current liquidity trades at date 1. This arises from the externality imposed by the speculator on the insider via the signal s extracted from $x_1 + u_1$. The insider has an incentive to tilt his trades at date 1 and manipulate p_1 in order to avoid the speculator's inference. Market efficiency improves thanks to the increase in the insider's trading intensity following the speculator's entry. However relative market efficiency is worsened by the speculator. Market depth is lowered by the speculator's entry at both dates. This is due to the higher adverse selection costs faced by the market maker in the presence of the speculator. Finally, speculator's profits increase with the quality of his non-fundamental information, while the insider's ex-ante gains are reduced. The consequences on market volume and liquidity are further investigated in the following section.

4 Testable implications

When bringing our model in section 3 to the data, we interpret days as rounds. This way the first date coincides with the day following the announcement, while the second date is the inclusion day.

4.1 Trading volume

We follow Admati and Pfleiderer (1988) and decompose the expected total volume into the contribution of each group of traders. For the model in section 2 one has:

$$V_1 \equiv V_1^I + V_1^L + V_1^M = \frac{1}{2}\left(E_0\left|x_1\right| + E_0\left|u_1\right| + E_0\left|\omega_1\right|\right) \tag{27a}$$

$$V_2 \equiv V_2^I + V_2^L + V_2^P + V_2^M = \frac{1}{2}\left(E_1\left|x_2\right| + E_1\left|u_2\right| + E_1\left|z_2\right| + E_1\left|\omega_2\right|\right) \tag{27b}$$

where $E_t(\cdot)$ denotes the expectation conditional on time $t-1$ public information, and superscripts L and P refer respectively to liquidity and passive traders. Since all orders but z_2 and ω_2 are conditionally normal with mean zero, the contributions to date 1 total trading volume follow from Admati and Pfleiderer (1988). On the other hand, in the absence of the speculator one has $E_1(z_2) = E_1(\omega_2) = \bar{z}_0 \neq 0$, which implies that volume at the inclusion depends on the (unconditional) expectation of the passive trades (we derive expressions for V_2^P and V_2^M in the appendix). We plot V_1 and V_2 in figure 9 setting $\bar{z}_0 = 2$ as a representative case. Note that we do not consider volume in RV in figure 9, since passive trades are absent in this model. During the first trading round V_1 increases in k_I due the insider's contribution V_1^I.[17] The fact that V_2 increases with the insider's informational advantage as well (panel B) might seem in contrast with the analysis for V_2^I, which was shown to be decreasing in k_I. In fact, one can show that V_2^P decreases with k_I as well, since passive volume is proportional to $\sigma_{z,0}^2$. However, liquidity trades V_2^L increase in their own variance $\sigma_{u_2}^2$, or equivalently in the informational advantage k_I. The latter dependence dominates the other two effects, and as a result V_2 increases in k_I. Moreover we note that an increase in the mean passive trades $\bar{z}_0$ (not reported for reasons of space) would move V_2 further up. Finally, the ratio V_2/V_1 (panel C) is always above unity, as to say that volume is expected to be higher upon inclusion.

For the model developed in section 3, the time t market volume is decomposed similarly to $(27a - 27b)$ as:

$$V_1 \equiv V_1^I + V_1^S + V_1^L + V_1^M = \frac{1}{2}\left(E_0 |x_1| + E_0 |y_1| + E_0 |u_1| + E_0 |\omega_1|\right)$$
$$V_2 \equiv V_2^I + V_s^S + V_2^L + V_2^P + V_2^M = \frac{1}{2}\left(E_1 |x_2| + E_1 |y_2| + E_1 |u_2| + E_1 |z_2| + E_1 |\omega_2|\right)$$

The presence of the speculator increases volume after the announcement (figure 9-panel A). This stems from the volume generated by the speculator (figure 8-panel A) as well as from the insider's manipulative incentives (figure 4-panel A). Given that both V_1^I and V_1^S increase in the speculator's advantage, it is not surprising that V_1 increases with k_S. Again, while date 1 trades are centered around zero, volume at the inclusion depends on the posterior $\bar{z}_1 = \bar{z}_0 + \bar{v}_1$ in the presence of the speculator. Hence the (conditional) expectation of passive trades after the first round plays a role in determining both V_2^P and V_2^M. Moreover from eq. (23) the posterior $\bar{z}_1$ depends on the realization of the first period aggregate order flow (as well as on $\bar{z}_0$). This implies that every realization of the first period trades $\omega_1 = x_1(f, u_1) + y_1(v) + u_1$ generates a different date 2 expected volume. In order to assess the effect of non-fundamental speculation on market volume, we therefore replace $\bar{v}_1$ by its estimate $\hat{v}_1$ using Monte Carlo simulation with 1000 draws for f, v and u_1, and then use $\bar{z}_0 + \hat{v}_1$ instead of $\bar{z}_1$ in the expression for V_2. Volume at the inclusion increases in k_I along the same lines as V_2 in $(27b)$ (see figure 9-panel B). Note that V_2 (as well as the volume ratio V_2/V_1 in panel C) is inversely related to k_S. In fact, while the speculator generates more volume when his advantage is sharp (figure 8-panel B), for the insider the opposite holds true (figure 4-panel B). The net result is that the latter effect offsets the former. Finally, an increase in expected passive trades $\bar{z}_0$ (not reported for reasons of space) increases volume at the inclusion, like in the absence of the speculator.

4.2 Market liquidity

The impact of the S&P game on market liquidity pattern is analyzed in figure 3-panel C, which plots the ratio λ_1/λ_2. Since date t market liquidity is given by $1/\lambda_t$, the ratio λ_1/λ_2 gives the evolution of market liquidity though time: for example $\lambda_1/\lambda_2 = \left(\frac{1/\lambda_2}{1/\lambda_1}\right) > 1$ means that the market is deeper at date 2 than at date 1. In the absence of the speculator, liquidity decreases over time with the insider's non-fundamental information quality: in fact, large values for k_I imply that the information asymmetry faced by the market maker is relatively severe. Thus the market maker's reaction to large values of k_I is to decrease market liquidity at both dates. Note that it takes a rather precise non-fundamental information ($k_I > 8$) in order to observe more illiquid markets at date 2. This means that depth decreases upon inclusion whenever the (volatility of the) noise coming from the passive industry is extremely small relative to other liquidity traders, i.e. $\sigma_{z,0}^2 < \sigma_{u_2}^2/8$. Hence the reduction in spreads before October 1989 [see Beneish and Whaley (1996), Edmister, Graham and Pirie (1996) and Erwin and Miller (1998)] suggests that $k_I < 8$ is in fact a reasonable bound for the insider's advantage.

The speculator's entry causes liquidity to decrease at both date (figure 3-panel A and B), since informational asymmetries are now more severe. The stock becomes more illiquid the higher is the speculator's advantage k_S. Moreover market depth reduces at the inclusion relative to the previous day (figure 3-panel C), and again this reduction is positively related to the speculator's informational advantage. In particular note that $\lambda_1/\lambda_2 < 1$ when $k_S = 3$ *regardless of the insider's informational advantage.* This means that –irrespective of the passive trades volatility relative to other liquidity-motivated orders– liquidity decreases at date 2 whenever the speculator is aware of at least 3/4 of the indexers' trades.

Recall from the introduction that the main reason for moving to preannouncing index changes hinges on the attempt to reduce trading imbalances after the announcement. The volume ratio seems to confirm this, since non-fundamental speculation reduces V_2/V_1. On the other hand, the S&P game reduces market liquidity at the inclusion. These two opposite effects might allow to cast some doubts on the effectiveness of the S&P change in the announcement practice.

5 Empirical study

5.1 Data set description

Between October 1989 and December 1999 there have been 248 replacements in the S&P 500.[18] As in the previous literature, we concentrate on market additions due to the fact that stocks removed from the S&P 500 often do not trade after the list change, or the announcement of deletion is confounded by firm-specific information [see Chen, Noronha and Singal (2004) and the references therein for empirical studies on deletions]. For notational convenience let AD denote the announcement day (i.e. the day in which after the close the announcement is made) and CD the effective day (i.e. the day in which after the close the change is effective). As previously noted, after October 1989 the replacement is effective at least one day after AD. From the total sample we removed some stocks. First of all we drop companies added and deleted from the index due to name changes (33 stocks) as well as stocks included due to merger (20) or spin-off (17) with another S&P 500 company. In all of these cases we would not observe the demand shock arising from passive traders which is the driving force for non-fundamental speculation in the model developed in section 3. In the second place we exclude companies for which we are not certain about the announcement date and/or the effective date (19) as well as stocks for which the inclusion occurs the day after the announcement (30). This latter requirement arises naturally from the time line underlying our theoretical model, since whenever AD+1 coincides with CD one cannot disentangle the effect of non-fundamental speculation from indexers' demand. For each company we collect daily (closing) data from CRSP on (1) bid price, (2) ask price, (3) volume (number of shares traded) and (4) outstanding shares. Eventually we require stock data availability for a period ranging from 250 days before to 40 days after the announcement, which resulted in dropping 21 companies. The final data set comprises 108 stocks. Figure 11 shows the frequency distribution of the number of trading days between AD and CD for the inclusions occurred under the preannouncement practice. The support ranges from one to sixteen trading days and the mode (resp. mean) is five (resp. 4.43), documenting the S&P common practice to preannounce changes five business days beforehand. This evidence is consistent with Beneish and Whaley (1996) for announcements between October

1989 and June 1994.

5.2 Trading volume

As explained in the introduction, the appeal to investors of passive techniques is widely documented by the growth in the net asset value experienced by the major funds pegged to the S&P 500 in the last two decades [see Beneish and Whaley (1996) and Wurgler and Zhuravskaya (2002) among others]. The widespread use of indexed funds can be assessed by looking at the trading volume pattern around AD and CD, since in section 4 we have shown that an increase in average passive trades –captured by $\bar{z}_0$– results in higher volume at the inclusion. Given that indexers' performance is assessed by daily tracking error minimization, pegged funds' rebalancing should occur at CD. Moreover the presence of risk arbitrageurs (i.e. the speculator in our model in section 3) increases volume after the announcement, i.e. over the window AD+1,...,CD-1. Finally, abnormal trading volume on AD may provide evidence that leakage of information regarding index inclusion has occurred.

Let $V_{i,t}$ denote the daily turnover for stock i on day t as measured by the ratio between the number of shares traded and the number of outstanding shares for company i during day t. We use daily turnover as a measure of the daily trading volume since it accounts for splits experienced by the stock, thus making turnover[19] preferred to raw volume. Therefore the abnormal trading volume on day t is the ratio between $V_{i,t}$ and the average trading volume in the 40 days[20] preceding the announcement day $\bar{V}_i \equiv \left(\sum_{t=AD-40}^{AD-1} V_{i,t}\right)/40$. Eventually we let $MAVR_t$ denote the cross-section average for the abnormal trading volume over a sample of size N_t:

$$AVR_{i,t} = V_{i,t}/\bar{V}_i \quad ; \quad MAVR_t = \frac{1}{N_t}\sum_{i=1}^{N_t} AVR_{i,t} \tag{29}$$

Results from inclusions in the S&P 500 are summarized in table II and figures 12–13, taking into account the number of trading days between AD and CD. Under the assumption that individual abnormal volume ratios are (cross-sectionally) independently and identically normally distributed, the resulting statistic for $MAVR_t$ follows a Student-t distribution with $N_t - 1$ degrees of freedom. Moreover, in order to assess the impact of outliers in our analysis, we perform a binomial test for the null hypothesis that the percentage of companies with $AVR_{i,t} > 1$ is different from 50%. [21] Table II reports sample size, mean abnormal volume ratio ($MAVR_t$), cross-sectional t-ratio ($t\,(MAVR)$) and the percentage of companies for which $AVR_{i,t}$ is greater than one over the window $AD-10, ..., CD+10$. Since the number of trading days between AD and CD varies across companies (see figure 11), the column labeled N_t in each panel in table II reports the number of stocks included in the sample. For each of the ten days after AD in panel A, only those firms for which CD has not yet occurred are included. This is why the sample size in the second column decreases over the days after AD in panel A. Similarly, for the ten days preceding CD, only firms for which AD has not yet occurred are included, such that the sample size increases over the ten days before CD in panel B. Figure 12 (resp. figure 13) plots $MAVR_t$ and its 95% confidence interval around AD (resp. CD).

On the day after the announcement trading volume is more than 4 times larger than the average daily volume over the 8 weeks base period, and is statistically significant at 5% level. Abnormal volume appears to be persistent in that $MAVR_t$ is greater than one for the whole week after AD+1, even though its magnitude is far from the increase experienced during AD+1, and $MAVR_t$ is not significantly different from one after AD+5. This evidence is consistent with the presence of non-fundamental speculators stepping into the market after the announcement, and diluting their orders over the days preceding the effective change. On AD the estimated mean abnormal volume is roughly 25% above the level in the 40 days preceding the announcement. Further, the t-statistic rejects $MAVR_{AD} = 1$ at 5% significance level. The latter finding is in line with all the above mentioned empirical studies on S&P inclusions after October 1989, and might suggest leakage of information about index replacements before announcement. Results from the ten days preceding AD do not detect abnormal trading activity, with the only exception of mean abnormal volume significantly greater than unity documented for AD-2. Comparing the percentage of individual firms whose $AVR_{i,t}$ is greater than one is useful to determine whether the $MAVR$'s are driven by outliers. More than 95% of the cross-section have individual $AVR_{i,AD+1}$ greater than one, this percentage being statistically different from 50% at 5% significance level. Over the window (AD+2,AD+5) more than 70% of the stocks in our sample display abnormal trading volume, which we regard as strengthening the evidence in favour of front-running strategies implemented after the announcement. On the other hand, the percentage of stocks with $AVR_{i,t} > 1$ on both AD-2 and AD is not statistically different from 50%, and we conclude that the abnormal trading volume documented for these two dates is due to a small number of companies.

Trading volume on CD is roughly 15 times higher than the base period and is statistically significant at 5% level. This suggests that passive managers actually wait until the effective day to rebalance their portfolios. It is noteworthy that virtually all of the companies experience an increase in trading volume upon inclusion. Trading activity for the five days before CD is at least twice the average volume during the 8 weeks preceding the announcement, and can be attributed to risk-arbitrageurs' activity. The increase in volume tends to be permanent, in that $MAVR_t$ is significantly different from one in all the days from CD+1 to CD+10, even though it steadily decreases after CD. The percentage of companies with $AVR_{i,t}$ different from one around CD shows that these findings do not appear to be driven by outliers over the fourteen days ranging from CD-4 to CD+9.

5.3 Liquidity

While several authors focused on trading volume around inclusions in the S&P 500, market depth has received little attention. Beneish and Whaley (1996) analyze the bid-ask spread for inclusions: as mentioned in the introduction they find reductions in the spread after the effective date. Furthermore, the authors report a significant 13% liquidity improvement for stocks included between 1986 and 1989 as well. Erwin and Miller (1998) focus on additions between 1984 and 1988, and document a significant spread decrease over the 30 days following the index change. Similarly, included companies experience liquidity improvements between 1983 and 1989 according to Edmister, Graham and Pirie (1996). In what follows we employ the relative bid-ask spread as a proxy for market liquidity. A measure for abnormal depth can be constructed along the same lines used for the trading volume analysis. Let $A_{i,t} - B_{i,t}$ be stock i's absolute bid-ask spread during day t, and $Q_{i,t}$ the quote midpoint, i.e. $Q_{i,t} = (A_{i,t} + B_{i,t})/2$. The relative bid-ask spread is $S_{i,t} = (A_{i,t} - B_{i,t})/Q_{i,t}$ and $\bar{S}_i \equiv \left(\sum_{t=AD-40}^{AD-1} S_{i,t}\right)/40$ denotes the average relative bid-ask spread over the base period.[22] The abnormal spread ratio $ASR_{i,t}$ and its cross-section counterpart $MASR_t$ are defined as follows:

$$ASR_{i,t} = S_{i,t}/\bar{S}_i \quad , \quad MASR_t = \frac{1}{N_t}\sum_{i=1}^{N_t} ASR_{i,t} \tag{30}$$

For the announcement and the inclusion not to affect market depth one should observe $MASR_t$ close to one both around AD and CD. On the other hand, a situation in which $MASR_t$ is less (resp. greater) than one detects a reduction (raise) in the average spread during day t relative to the base period, i.e. an increase (decrease) in market depth on day t relative to the 8 weeks preceding the announcement.

Table III and figures 14–15 report the mean abnormal spread over the window $AD-10, ..., CD+10$. Stocks experience a statistically significant 35% increase in the bid-ask spread the day after the announcement. No clear pattern emerges from our sample for the other days in the event window, with the exception of $MASR_{AD-2}$ being statistically different from unity. Before AD the market is more liquid over the week AD-10 to AD-6, while liquidity decreases over the week following the announcement. However one cannot reject the null hypothesis that liquidity is statistically different from the base period over both weeks. The percentage of firms experiencing wider spreads on AD+1 is significantly greater than 50%, while this does not occur on AD-2. We conclude that the latter reduction in liquidity is affected by few observations, while the former is not.

Market liquidity significantly decreases around inclusion. There is an average 20% increase in the spread during the day preceding the index change, which rises to almost 70% upon inclusion. The abnormal spread ratio is statistically different from one (5% significance level) on the effective day, as well as on CD-1 and CD-4. Notice, however, that the reduction in market liquidity during CD is not driven by outliers (approximately 80% stocks of the cross-section experience abnormal spreads), while this is not the case for both $MASR_{CD-1}$ and $MASR_{CD-4}$. Our findings are in contrast with Beneish and Whaley (1996), even though their results might be affected by their small sample size. In fact, their data set comprises 30 index inclusions from October 1989 through June 1994: the authors report a spread decrease during CD and the following days, even though the spread is significantly below normal only for CD+1. Recall from subsection 5.2 that stocks experience a significant increase in trading volume upon inclusion, which is driven by passive traders' demand. Beneish and Whaley (1996) argue that the specialist might temporarily charge a lower spread, given that the increase in trading volume would cover the operation costs. On the other hand the trading game in section 3 is consistent with an increase in the spread during the day of inclusion: liquidity should decrease over time as a consequence of the higher adverse selection costs faced by the market maker in the presence of (sufficiently accurate) non-fundamental information. Our sample seems to confirm this implication, and we argue that any reduction in operational costs arising from greater volume is more than offset by the asymmetric information costs faced by the market maker in the presence of non-fundamental speculators.

5.4 Assessing the importance of the S&P game

The findings in the previous two subsections point at (1) a significant increase in trading volume following the announcement and upon inclusion (2) a significant decrease in liquidity both after the announcement and upon inclusion. The empirical implications arising from our theoretical model(s) stand on the comparison of both volume and liquidity between the day after the announcement and the effective day (respectively the first and second trading round in the models developed in sections 2 and 3). However for most of the companies in our sample there is more than one day between AD and CD. Therefore, for each of these stocks we define $\overline{AVR_i}$ as the average abnormal volume over the window (AD+1,CD-1), i.e. $\overline{AVR_i} \equiv \sum_{t=AD+1}^{CD-1} AVR_{i,t}$, and then average the $\overline{AVR}$s across stocks to get $M\overline{AVR} = N^{-1} \sum_{i=1}^{N} \overline{AVR_i}$, where N denotes the sample size ($\overline{ASR_i}$ and $M\overline{ASR}$ are defined along the same lines). In order to measure the change in volume we perform a two-sided test for the null $M\overline{AVR} = MAVR_{CD}$ (the change in liquidity is assessed in the same way).

Results on the entire dataset are summarized in the first row of table IV. Both volume and bid-ask spread significantly increase upon inclusion relative to the window (AD+1,CD-1). Our findings appear robust, since the binomial test rejects the hypothesis that the percentage of firms experiencing this increase is equal to 50% at 5% significance level. The increase in volume is consistent with both models in sections 2 and 3, since it may simply reflect the presence of passive funds stepping into the market on the effective day. Similarly –as explained in subsection 4.2– the decrease in liquidity might be due to an insider with highly accurate non-fundamental information in the absence of the speculator ($k_I > 8$). Alternatively, a speculator endowed with relatively precise information ($k_S > 3$) might be responsible of the spread increase. We tend however to disregard the first explanation given that all the authors focusing on inclusions before October 1989 document decline in spreads for stocks included in the S&P 500. During this period, the simultaneous occurrence of announcement and inclusion *de facto* rules out the S&P game. As previously noted, taking the model in section 2 as a reference, the findings in Edmister, Graham and Pirie (1996) and Erwin and Miller (1998) are consistent with a relatively poor quality of the insider's signal ($k_I < 8$). We therefore attribute the worsening in liquidity documented in table III to non-fundamental speculators front-running index funds after October 1989. In order to assess the robustness of the S&P game we test for significant changes in volume and liquidity splitting our dataset in three subsamples: 1989-1994, 1995-1997 and 1998-99. Reasonable sample size is one of the criteria we used in choosing these subsamples. Moreover the average net asset value for passive funds over these periods is equal to 287, 606 and 1117.5 billion USD respectively. Expected passive trades thus increase across the subsamples, but are relatively stable within each subsample (see figure 1). It is further reasonable to conjecture that the shift towards passive strategies resulted in an increase in the number of indexers, which in turn implies that passive trades' variance increases over time. This is particularly true for the last subsample, given that funds moved away from full replication in recent years [see Blume and Edelen (2004)]. Recall from section 4 that, absent the speculator, the combined effect of an increase in both $\bar{z}_0$ and $\sigma^2_{z,0}$ is higher volume and tighter spreads upon inclusion. The former is due to larger orders coming from indexers, while the latter stems from a reduction in the insider's advantage $k_I = \sigma^2_{u_2}/\sigma^2_{z,0}$. On the other hand, an increase in both $\bar{z}_0$ and $\sigma^2_{z,0}$ is compatible with higher volume and lower liquidity in the presence of a speculator with highly accurate information.

From table IV it emerges that volume significantly increases upon inclusion in all subsamples, while liquidity significantly decreases from 1995 onwards. Our findings support the following argument: it took some time for investors to start front-running indexers after the change in the S&P announcement practice. Before 1995, the volume-liquidity pattern points at a statistically significant increase in trading activity, while spreads are unaltered. Starting from 1995 the S&P game has gained appeal, yielding positive profits and significantly worsening liquidity.

6 Conclusion

In the last two decades passive funds have gained an increasing consideration among investors as a relatively cheap tool to achieve portfolio diversification. Passive funds aim at mimicking a benchmark index. Portfolio rebalancing, as well as performance evaluation, is carried out by means of tracking error procedures. Index replacements stand as a clear rebalancing opportunity for passive managers. Starting from October 1989, changes in the S&P 500 composition are preannounced by Standard and Poor's usually five days beforehand. Passive funds are not affected by the announcement timing and their portfolio rebalancing occurs during the effective day. On the other hand this preannouncement policy induces non-fundamental speculators to enter the market. Non-fundamental speculators do not possess any information on the asset's fundamental value, rather they buy the included stock ahead of passive funds and sell it a few days later at possibly higher prices. We develop a dynamic model that explicitly keeps into account this preannouncement practice. We show that strategies based on non-fundamental information are profitable and determine a drop in market liquidity, as a direct consequence of the increased adverse selection costs faced by the specialist. Examining S&P 500 inclusions from October 1989 to December 1999 we find evidence consistent with our theoretical analysis.

	date 1		date 2	
	info set Φ_1^i	strategy	info set Φ_2^i	strategy
insider (I)	f, u_1	$x_1(f, u_1)$	$\Phi_1^I \cup u_2, \omega_1$	$x_2(f, u_2)$
market maker (M)	ω_1	$p_1(\omega_1)$	$\Phi_1^M \cup \omega_2$	$p_2(\omega_1, \omega_2)$
insider (I)	f, u_1	$x_1\left(f, u_1, E\left(s \mid \Phi_1^I\right)\right)$	$\Phi_1^I \cup u_2, s, v, \omega_1$	$x_2(f, u_2, s, v)$
speculator (S)	v	$y_1(v)$	$\Phi_1^S \cup s, \omega_1$	$y_2(s, v)$
market maker (M)	ω_1	$p_1(\omega_1)$	$\Phi_1^M \cup \omega_2$	$p_2(\omega_1, \omega_2)$

Table I: information structure for the trading games in sections 2 and 3. The information sets and the strategies for market participants are shown. Top panel: model in section 2; bottom panel: model in section 3.

day	panel A -- event day: AD				panel B -- event day: CD			
	N	*MAVR*	*t (MAVR)*	*AVR >1*	*N*	*MAVR*	*t (MAVR)*	*AVR >1*
-10	108	0.912	-1.349	**28.70**	2	0.833	-1.684	0.00
-9	108	1.051	0.576	**34.26**	2	0.966	-0.479	50.00
-8	108	0.984	-0.258	**37.04**	2	0.834	-0.911	50.00
-7	108	0.934	-1.464	**39.81**	2	0.906	-0.785	50.00
-6	108	1.021	0.229	**33.33**	11	1.951	2.536	72.73
-5	108	0.988	-0.191	**37.04**	18	**2.313**	2.139	61.11
-4	108	1.077	1.132	46.30	49	**3.055**	5.196	**77.55**
-3	108	1.071	0.999	41.67	75	**2.897**	6.762	**78.67**
-2	108	**1.229**	2.078	41.67	94	**2.569**	7.639	**86.17**
-1	108	1.018	0.276	**40.74**	108	**3.394**	8.215	**91.67**
0	108	**1.234**	2.990	51.85	108	**15.086**	13.259	**99.07**
+1	108	**4.471**	10.652	**96.30**	108	**3.266**	10.945	**94.44**
+2	94	**2.468**	7.504	**85.11**	108	**2.112**	8.360	**84.26**
+3	75	**2.313**	6.548	**78.67**	108	**2.039**	6.878	**75.93**
+4	49	**2.160**	5.172	**71.43**	108	**1.763**	6.953	**72.22**
+5	18	**1.846**	2.602	**77.78**	108	**1.445**	4.870	**63.89**
+6	11	1.457	1.510	63.64	108	**1.590**	5.696	**70.37**
+7	2	0.840	-2.966	0.00	108	**1.408**	4.857	**62.96**
+8	2	0.758	-0.937	50.00	108	**1.465**	4.494	**60.19**
+9	2	0.918	-0.762	50.00	108	**1.421**	4.546	**63.89**
+10	2	0.738	-8.752	0.00	108	**1.471**	3.153	55.56

Table II: Abnormal trading volume around the announcement and the effective days. Abnormal trading volume is defined for each stock in (29). In the sample AD precedes CD by at least one day. Since the number of trading days between AD and CD varies across firms, the columns labeled N reports the number of companies included in the cross-section for each day. Boldface numbers in columns labeled $MAVR$ denote mean trading volume significantly different from one (5% significance level). Boldface numbers in columns labeled $AVR > 1$ denote percentage significantly different from 0.5 (5% significance level).

day	panel A -- event day: AD				panel B -- event day: CD			
	N	*MASR*	*t* (*MASR*)	*ASR* >1	*N*	*MASR*	*t* (*MASR*)	*ASR* >1
-10	108	0.929	-1.962	**37.96**	2	0.504	-2.170	0.00
-9	108	0.961	-0.877	**38.89**	2	1.320	0.654	50.00
-8	108	0.934	-1.811	**37.96**	2	1.557	1.354	100.00
-7	108	0.965	-0.933	**39.81**	2	2.796	1.713	100.00
-6	108	0.983	-0.323	**37.96**	11	1.188	1.279	72.73
-5	108	1.010	0.214	43.52	18	1.233	1.444	55.56
-4	108	0.997	-0.058	43.52	49	**1.289**	2.611	57.14
-3	108	1.059	1.038	**40.74**	75	1.115	1.596	46.67
-2	108	**1.152**	2.437	53.70	94	1.114	1.885	47.87
-1	108	1.041	0.647	**36.11**	108	**1.190**	3.467	53.70
0	108	1.115	1.943	42.59	108	**1.685**	7.679	**81.48**
1	108	**1.355**	5.409	**65.74**	108	0.961	-0.816	**35.19**
2	94	1.063	1.065	44.68	108	0.983	-0.339	41.67
3	75	1.134	1.919	46.67	108	1.065	0.881	**37.04**
4	49	1.106	1.128	42.86	108	0.955	-0.885	**33.33**
5	18	1.035	0.280	44.44	108	0.933	-1.551	**36.11**
6	11	0.994	-0.051	45.45	108	1.034	0.645	46.30
7	2	2.827	1.795	100.00	108	0.956	-0.893	**35.19**
8	2	1.178	5.427	100.00	108	0.963	-0.851	42.59
9	2	1.617	4.727	100.00	108	0.994	-0.129	44.44
10	2	0.960	-0.508	50.00	108	0.981	-0.382	42.59

Table III: Abnormal bid-ask spread around the announcement and the effective days. Abnormal bid-ask spread is defined for each stock in (30). In the sample AD precedes CD by at least one day. Since the number of trading days between AD and CD varies across firms, the columns labeled N reports the number of companies included in the cross-section for each day. Boldface numbers in columns labeled $MASR$ denote mean bid-ask spread significantly different from one (5% significance level). Boldface numbers in columns labeled $ASR > 1$ denote percentage significantly different from 0.5 (5% significance level).

		panel A -- volume			panel B -- bid-ask spread		
sample	*N*	*DMAVR*	*t(DMAVR)*	*DAVR>0*	*DMASR*	*t(DMASR)*	*DASR>0*
all obs	108	**11.515**	10.435	**99.07**	**0.503**	4.95	**72.23**
1989-94	29	**10.102**	4.526	**100.00**	0.268	1.553	58.62
1995-97	40	**10.675**	6.269	**97.5**	**0.591**	3.507	**75.00**
1998-99	39	**13.429**	7.266	**100.00**	**0.587**	3.228	**79.48**

Table IV: empirical test for the model in section 3. Abnormal bid-ask spread is defined for each stock in (30). In the sample AD precedes CD by at least one day. The test for the equality of mean trading volume (panel A) and bid-ask spread (panel B) around inclusion is presented. Boldface numbers in columns labeled $DMAVR$ denote mean volume on CD significantly different from mean volume over AD+1 to CD-1 (5% significance level). Boldface numbers in columns labeled $DAVR > 0$ denote percentage significantly different from 0.5 (5% significance level). Headers for the bid-ask spread have a similar interpretation.

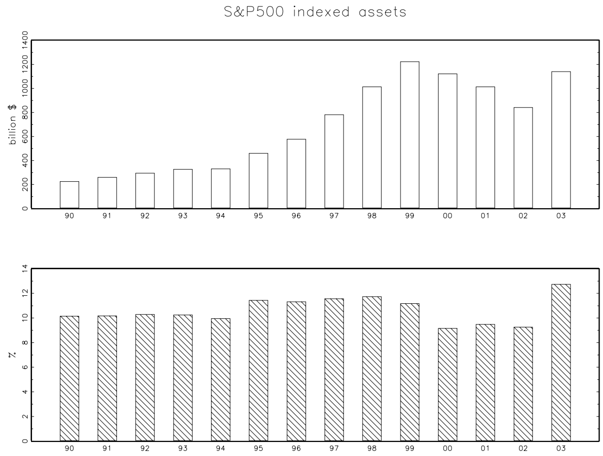

Figure 1 : assets indexed to the S&P 500. End of the year net asset value for passive funds is obtained from S&P (2003), while yearly S&P 500 capitalization is obtained from the S&P website http://www.standardandpoors.com. Top panel: S&P 500 indexed assets (billion USD); bottom panel: indexed assets NAV relative to S&P 500 capitalization (percentage).

Key to figures 2-9. Equilibrium values for several parameters and variables are shown in figures 2-9. The underlying parameters are: $\sigma^2_{f,0} = 1$, $\sigma^2_{u_1} = 1$, $\sigma^2_{u_2} + \sigma^2_{z,0} = 1$. The horizontal axis show the insider's non-fundamental information quality is $k_I = \sigma^2_{u_2}/\sigma^2_{z,0}$. The relevant parameters and variables for the model in sections 2 and 3 are computed for $k_I = 0.01, 0.5, 1, 2, 3.5, 5, 7.5, 10, 15, 20$. Further for the model in section 3 the values for the speculator's non-fundamental information quality are $k_S = 0.33, 1{,}3$ where $k_S = \sigma^2_{v,0}/\sigma^2_w$. The solid line corresponds to the model in section 2, while the dashed line is the 2 period version of Rochet and Vila (1994) (RV). For the model in section 3 values for the case $k_S = 0.33, 1$ and 3 are shown respectively with circles, squares and triangles (low, medium and high informational advantage respectively).

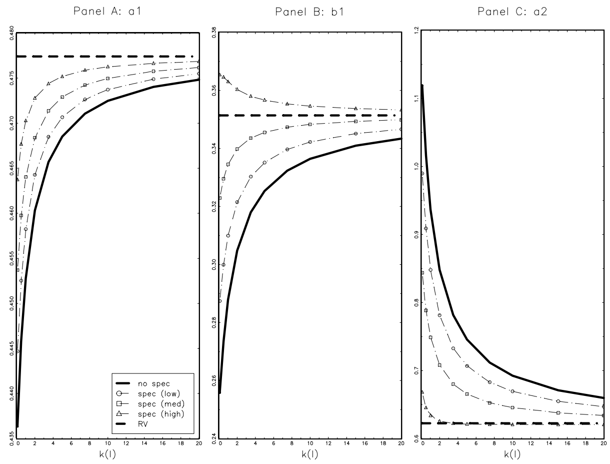

Figure 2: insider intensities. Equilibrium values for parameters a_1 (panel A), b_1 (panel B) and a_2 (panel C).

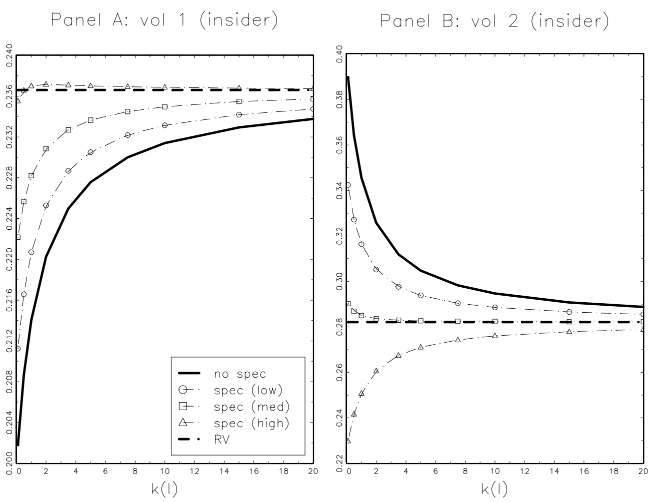

Figure 3: insider volume. Equilibrium values for V_1^I (panel A) and V_2^I (panel B).

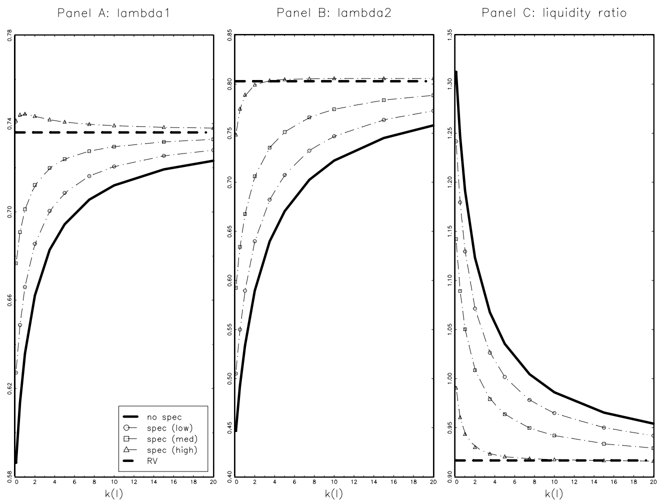

Figure 4: market liquidity. Equilibrium values for parameters λ_1 (panel A),λ_2 (panel B), and the ratio λ_1/λ_2 (panel C).

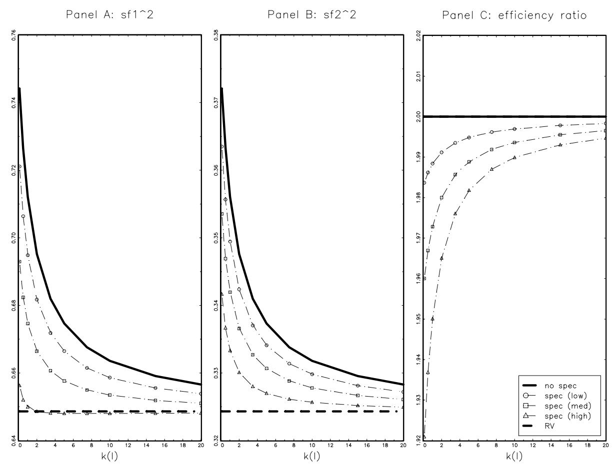

Figure 5: price informativeness. Equilibrium values for residual variances $\sigma_{f,1}^2$ (panel A) $\sigma_{f,2}^2$ (panel B) and the ratio $\sigma_{f,1}^2/\sigma_{f,2}^2$ (panel C).

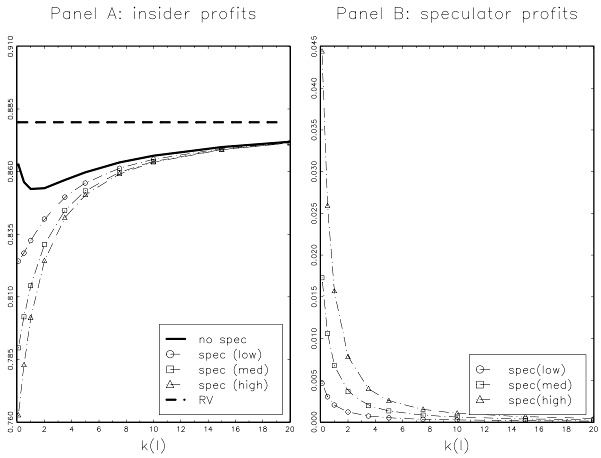

Figure 6: profits. Equilibrium values for ex-ante profits $E\left(\pi_1^I + \pi_2^I\right)$ (panel A) and $E\left(\pi_1^S + \pi_2^S\right)$ (panel B)

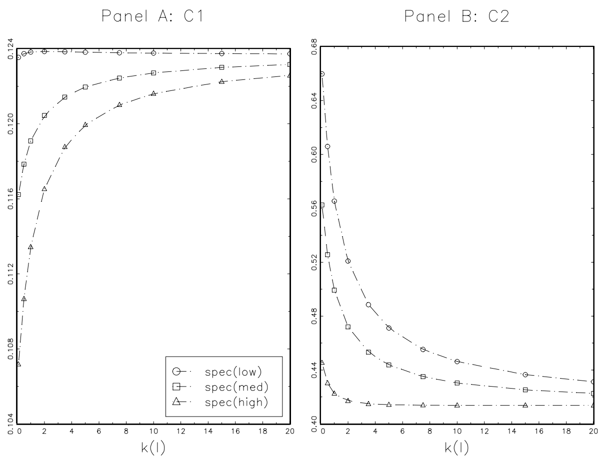

Figure 7: speculator intensities. Equilibrium values for parameters C_1 (panel A) and C_2 (panel B).

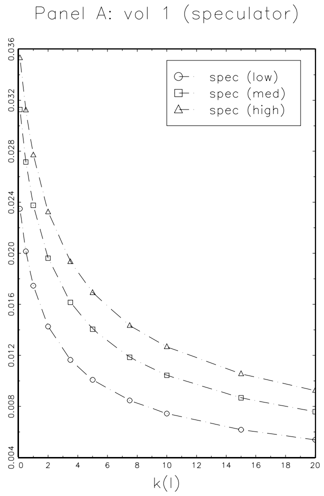

Figure 8: speculator volume. Equilibrium values for V_1^S (panel A) and V_2^S (panel B)

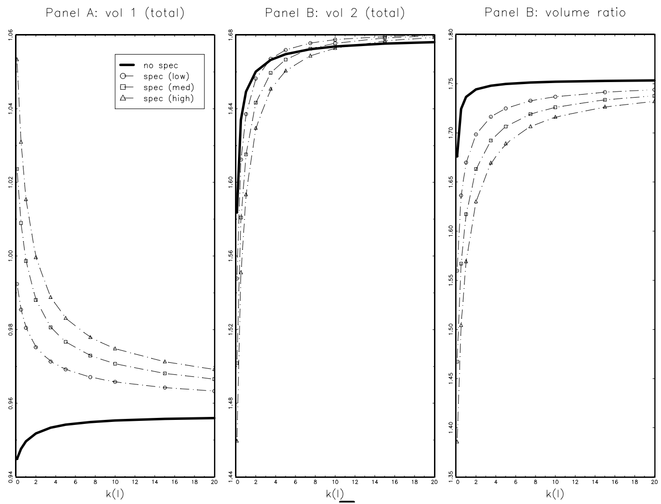

Figure 9: total volume. Equilibrium values for V_1 (panel A) and V_2 (panel B)

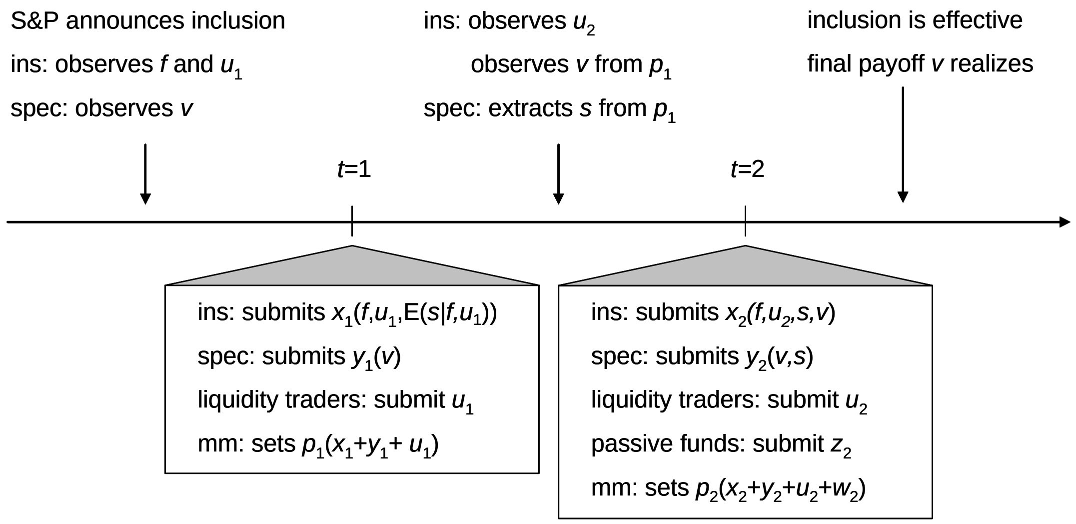

Figure 10: time line for the trading game in section 3

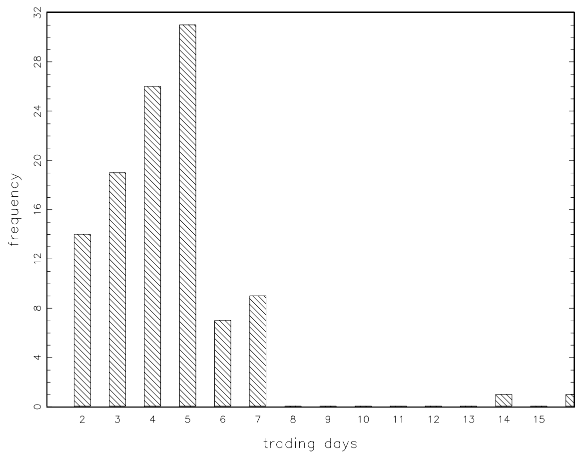

Figure 11: announcement frequencies. Frequency distribution of the number of trading days between the announcement and the effective day over the period October 1989-December 1999 for S&P 500 (108 inclusions)

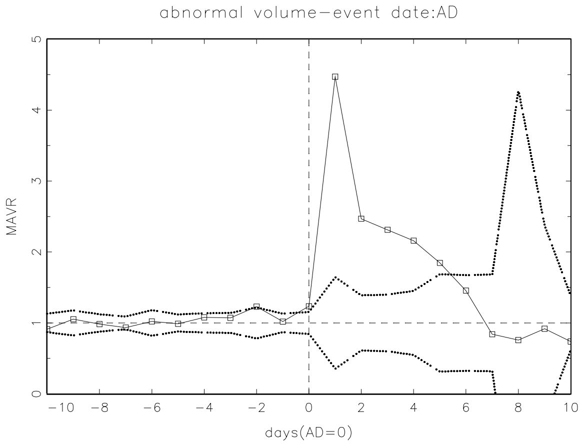

Figure 12: trading volume around the announcement day. Mean abnormal volume is defined in (29). The $MAVR$'s are displayed (bold solid line) for each trading day in the window (AD-10,AD+10) together with the 95% confidence interval for the null hypothesis $MAVR = 1$.

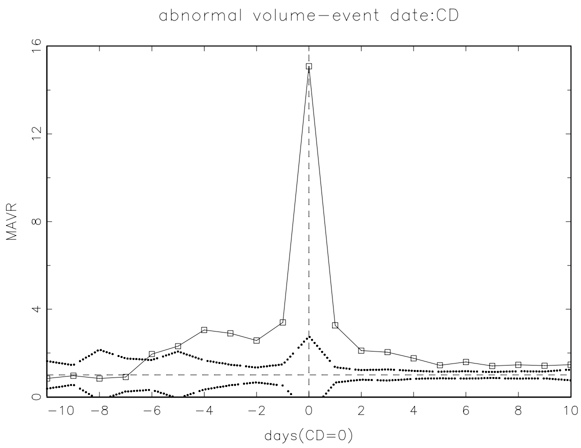

Figure 13: trading volume around the effective day. Mean abnormal volume is defined in (29). The $MAVR$'s are displayed (bold solid line) for each trading day in the window (CD-10,CD+10) together with the 95% confidence interval for the null hypothesis $MAVR = 1$.

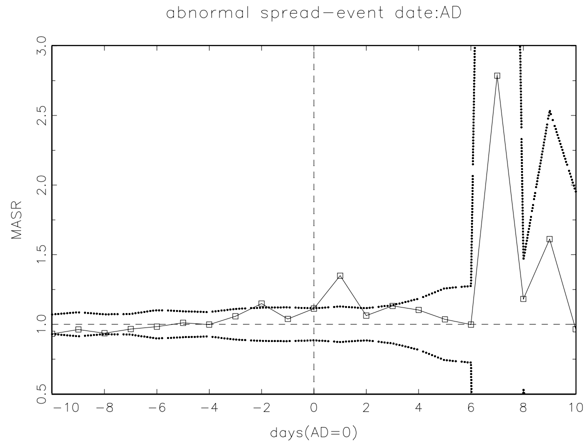

Figure 14: bid-ask spread around the announcement day. Mean bid-ask spread is defined in (30). The $MASR$'s are displayed (bold solid line) for each trading day in the window (AD-10,AD+10) together with the 95% confidence interval for the null hypothesis $MASR = 1$.

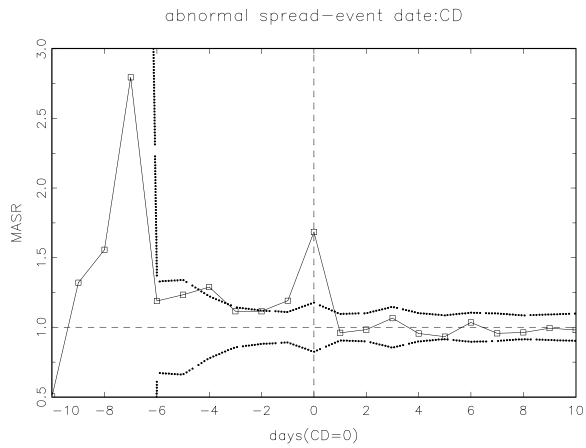

Figure 15: bid-ask spread around the effective day. Mean bid-ask spread is defined in (30). The $MASR$'s are displayed (bold solid line) for each trading day in the window (CD-10,CD+10) together with the 95% confidence interval for the null hypothesis $MASR = 1$.

Appendix

Proof (Proposition 1). Given the pricing function (9), the insider chooses his second period trade x_2 in order to maximize the objective function (3)

$$E\left[x_2\left(f-p_2\left(\omega_2\right)\right)\middle|\,\Phi_2^I\right]=x_2\left(f-p_1-\lambda_2x_2-\lambda_2u_2\right)\ .$$

The first order condition gives $x_2=\dfrac{f-p_1}{2\lambda_2}-\dfrac{u_2}{2}$, such that eq. (7) obtains with $a_2=(2\lambda_2)^{-1}$. Further the second order condition is $\lambda_2>0$. Plugging eq. (7) in the objective function (3) gives

$$E\left(\pi_2^I|\Phi_2^I\right)=\frac{\left(f-p_1\right)^2}{4\lambda_2}+\frac{\lambda_2u_2^2}{4}-\frac{\left(f-p_1\right)u_2}{2}\ . \tag{31}$$

In the first trading round the insider chooses x_1 to maximize (2), i.e.

$$E\left[x_1\left(f-p_1\left(\omega_1\right)\right)+\pi_2^I\left(x_1\right)\middle|\,\Phi_1^I\right]\ .$$

By the Law of Iterated Expectations $E\left[\pi_2^I\left(x_1\right)|\Phi_1^I\right]=E\left[E\left(\pi_2^I|\Phi_2^I\right)|\Phi_1^I\right]$, where (31) gives $E\left(\pi_2^I|\Phi_2^I\right)$. Therefore when submitting his order x_1 the insider has to keep into account the impact of his trade on the price $p_1\left(x_1\right)$. Assuming that the first period price is set according to eq. (8), the first order condition is

$$\left(1-\tfrac{\lambda_1}{2\lambda_2}\right)E\left(f-p_1\left(\omega_1\right)\middle|\,\Phi_1^I\right)-\lambda_1x_1=0\ ,$$

or equivalently:

$$x_1=\frac{2\lambda_2-\lambda_1}{\lambda_1\left(4\lambda_2-\lambda_1\right)}\left(f-p_0\right)-\frac{2\lambda_2-\lambda_1}{4\lambda_2-\lambda_1}u_1\ ,$$

such that eq. (6) obtains with $a_1=\frac{2\lambda_2-\lambda_1}{\lambda_1(4\lambda_2-\lambda_1)}$ and $b_1=\lambda_1a_1$. Finally the second order condition for (2) is $\lambda_1\left(2-\frac{\lambda_1}{2\lambda_2}\right)>0$, and the inequality (10) follows since $\lambda_2>0$. We now determine equilibrium prices. Let $\sigma_{\omega,0}^2$ and $\sigma_{f\omega,0}$ denote respectively the unconditional variance of the first period aggregate order flow and the unconditional covariance between ω_1 and f. The unconditional distribution for the random variables (f,u_1) together with the first period trade (6) yields $\sigma_{\omega,0}^2=a_1^2\sigma_{f,0}^2+\left(1-b_1\right)^2\sigma_{u_1}^2$ and $\sigma_{f\omega,0}=a_1\sigma_{f,0}^2$.Therefore the efficiency condition (4) together with the Projection Theorem gives the price in eq. (8), where the regression coefficient $\lambda_1=\sigma_{f\omega,0}/\sigma_{\omega,0}^2$ is defined as

$$\lambda_1=\frac{a_1\sigma_{f,0}^2}{a_1^2\sigma_{f,0}^2+\left(1-b_1\right)^2\sigma_{u_1}^2}\ , \tag{32}$$

and the fundamental value posterior variance is $\sigma_{f,1}^2=var\left(f|\omega_1\right)=\left(1-a_1\lambda_1\right)\sigma_{f,0}^2$.

During the second trading round the market maker observes the order flow

$$\omega_2=x_2+u_2+z_2=a_2\left(f-p_1\right)+u_2/2+z_2\ .$$

It follows that $\omega_2|\omega_1\sim N\left(\bar{z}_0,\sigma_{\omega,1}^2\right)$ with $\sigma_{\omega,1}^2=var\left(\omega_2|\omega_1\right)=a_2^2\sigma_{f,1}^2+\sigma_{u_2}^2/4+\sigma_{z,0}^2$, and the second period price is given by eq. (9) with

$$\lambda_2=\frac{\sigma_{f\omega,1}}{\sigma_{\omega,1}^2}=\frac{a_2\sigma_{f,1}^2}{a_2^2\sigma_{f,1}^2+\sigma_{u_2}^2/4+\sigma_{z,0}^2}\ .$$

Substituting for $a_2 = (2\lambda_2)^{-1}$ in the latter results in a second order equation in λ_2 which admits the unique root (uniqueness follows from the insider's date 2 second order condition)

$$\lambda_2 = \frac{\sigma_{f,1}}{\left(\sigma_{u_2}^2 + 4\sigma_{z,0}^2\right)^{1/2}} \ . \tag{33}$$

The fundamental value residual variance after the second trading round is $\sigma_{f,2}^2 = var\left(f|\omega_2\right) = \left(1 - a_2\lambda_2\right)\sigma_{f,1}^2$ .In order to compute the insider's unconditional profits note that the Law of Iterated Expectations applied to (31) gives

$$E\left(\pi_2^I\right) = E\left[E\left(\pi_2^I|\Phi_2^I\right)\right] = \frac{\sigma_{f,1}^2}{4\lambda_2} + \frac{\lambda_2\sigma_{u_2}^2}{4} \ .$$

First period unconditional profits are obtained substituting the equilibrium trade (6) into (2), yielding

$$E\left(\pi_1^I\right) = E\left[E\left(\pi_1^I|\Phi_1^I\right)\right] = a_1\left(1 - a_1\lambda_1\right)\sigma_{f,0}^2 - b_1\lambda_1\left(1 - b_1\right)\sigma_{u_1}^2 \ .$$

Adding up the last two equations gives total unconditional profits as

$$E\left(\pi_1^I + \pi_2^I\right) = a_1\left(1 - a_1\lambda_1\right)\sigma_{f,0}^2 - b_1\lambda_1\left(1 - b_1\right)\sigma_{u_1}^2 + \frac{\sigma_{f,1}^2}{4\lambda_2} + \frac{\lambda_2\sigma_{u_2}^2}{4} \ . \tag{34}$$

$a_2 > 0$ follows from $a_2 = (2\lambda_2)^{-1}$ and $\lambda_2 > 0$. From the inequality (10) one has $4\lambda_1\lambda_2 - \lambda_1^2 > 0$ or equivalently $4\lambda_1\lambda_2 > \lambda_1^2 \geq 0$. Therefore $\lambda_1 > 0$ since $\lambda_2 > 0$. $a_1 > 0$ follows from $\lambda_1 > 0$ and the expression for λ_1 in eq. (32). Finally $b_1 > 0$ since $b_1 = \lambda_1 a_1$. Before deriving the expression for the expected volume, we prove the following:

Lemma 2 *Let* $X \sim N\left(\mu, \sigma^2\right)$. *Then*

$$E\left|X\right| = \sqrt{\frac{2}{\pi}}\sigma e^{-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2} + \mu\left[1 - 2\Phi\left(-\mu/\sigma\right)\right] \tag{35}$$

where $\Phi\left(\cdot\right)$ *is the cumulative distribution for the standard normal distribution.*

Proof. Let $f\left(x\right)$ be the normal probability distribution. Then:

$$\begin{aligned}
E\left|X\right| &= \int_0^{+\infty} xf\left(x\right)dx - \int_{-\infty}^0 xf\left(x\right)dx \\
&= \frac{\sigma}{\sqrt{2\pi}}\left(\int_{-\mu/\sigma}^{+\infty} ze^{-z^2/2}dz - \int_{-\infty}^{-\mu/\sigma} ze^{-z^2/2}dz\right) \\
&\quad +\mu\left(\int_{-\mu/\sigma}^{+\infty}\frac{e^{-z^2/2}}{\sqrt{2\pi}}dz - \int_{-\infty}^{-\mu/\sigma}\frac{e^{-z^2/2}}{\sqrt{2\pi}}dz\right) \\
&= \frac{\sigma}{\sqrt{2\pi}}\left[2e^{-(\mu/\sigma)^2/2}\right] + \mu\left[1 - 2\Phi\left(-\mu/\sigma\right)\right]
\end{aligned}$$

where the change in variable via $z = \left(x - \mu\right)/\sigma$ gives the second line, and the last line follows from straightforward computations. The trading volume in Admati and Pfleiderer (1988) clearly follows from eq. (35) setting $\mu = 0$. ■

The contributions of each group of traders to the total expected volume as in $(27a, 27b)$ are therefore given by

$$\begin{aligned}
V_1^I &= \sqrt{\frac{a_1^2\sigma_{f,0}^2 + b_1^2\sigma_{u_1}^2}{2\pi}} \quad ; \quad V_1^L = \frac{\sigma_{u_1}}{\sqrt{2\pi}} \quad ; \quad V_1^M = \frac{\sigma_{\omega,0}}{\sqrt{2\pi}} \\
V_2^I &= \sqrt{\frac{a_2^2\sigma_{f,1}^2 + \sigma_{u_2}^2/4}{2\pi}} \quad ; \quad V_2^L = \frac{\sigma_{u_2}}{\sqrt{2\pi}} \\
V_2^P &= \frac{\sigma_{z,0}}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\bar{z}_0}{\sigma_{z,0}}\right)^2} + \frac{\bar{z}_0}{2}\left(1 - 2\Phi\left(-\bar{z}_0/\sigma_{z,0}\right)\right) \\
V_2^M &= \frac{\sigma_{\omega,1}}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\bar{z}_0}{\sigma_{\omega,1}}\right)^2} + \frac{\bar{z}_0}{2}\left(1 - 2\Phi\left(-\bar{z}_0/\sigma_{\omega,1}\right)\right)
\end{aligned} \tag{36}$$

■

Remark (parameters in RV). The case of complete knowledge about date 2 noise trades can be obtained from Proposition 1 substituting (7) with the following

$$x_2 = a_2 (f - p_0) - (u_2 + z)/2 \; .$$

The second period expected profits in (31) become

$$E\left(\pi_2^I | \Phi_2^I\right) = \frac{(f - p_1)^2}{4\lambda_2} + \frac{\lambda_2 (u_2 + z)^2}{4} - \frac{(f - p_1)(u_2 + z)}{2} \; ,$$

while date 2 price sensitivity in (33) is $\lambda_2 = \sigma_{f,1}\left(\sigma_{u_2}^2 + \sigma_{z,0}^2\right)^{-1/2}$. Finally unconditional profits in (34) become

$$E\left(\pi_1^I + \pi_2^I\right) = a_1 (1 - a_1\lambda_1)\sigma_{f,0}^2 - b_1\lambda_1 (1 - b_1)\sigma_{u_1}^2 + \frac{\sigma_{f,1}^2}{4\lambda_2} + \frac{\lambda_2\left(\sigma_{u_2}^2 + \sigma_{z,0}^2\right)}{4} \; .$$

Other parameters and variables are defined like in Proposition 1. From the above formulas it emerges that parameter values for RV can be obtained setting $z = 0$ and $\sigma_{z,0}^2 = 0$ in Proposition 1.

Proof (Proposition 2). The proof is organized in three steps.

Step 1: date 2 trades

Given the speculator's trade (19) and the pricing function (21), the insider chooses his second period trade x_2 in order to maximize the objective function (3), i.e.

$$E\left[x_2 (f - p_2(\omega_2)) | \Phi_2^I\right] = x_2 \left[f - p_1 - \lambda_2 x_2 - C_2\lambda_2 (s - p_1) - 2\lambda_2 (v - \bar{v}_1)/3 - \lambda_2 u_2\right] \; . \tag{37}$$

The optimality conditions for (37) are:

$$\begin{aligned}
x_2 &= \frac{f - p_1}{2\lambda_2} - \frac{u_2}{2} - \frac{C_2}{2}(s - p_1) - (v - \bar{v}_1)/3 \\
\lambda_2 &> 0
\end{aligned}$$

such that eq. (17) obtains with $a_2 = (2\lambda_2)^{-1}$. Similarly, given the insider's order (17) and the pricing function (21), the speculator chooses y_2 in order to maximize the objective function (14):

$$E\left[y_2\left(f - p_2\left(\omega_2\right)\right)\middle|\, \Phi_2^S\right] = y_2\left[-\lambda_2 y_2 - 2\lambda_2\left(v - \bar{v}_1\right)/3 + \left(1 - \left(a_2 - C_2/2\right)\lambda_2\right)\left(s - p_1\right)\right] \ , \qquad (38)$$

yielding the optimality conditions:

$$\begin{aligned} y_2 &= -\frac{v - \bar{v}_1}{3} + \frac{1 - \left(a_2 - C_2/2\right)\lambda_2}{2\lambda_2}\left(s - p_1\right) \\ \lambda_2 &> 0 \end{aligned}$$

Equation (19) then obtains with $C_2 = \dfrac{1 - \left(a_2 - C_2/2\right)\lambda_2}{2\lambda_2}$. Solving for the coefficient C_2 yields $C_2 = 2a_2/3$ and the time 2 trades (17) and (19) become:

$$x_2 = \frac{f - p_1}{2\lambda_2} - \frac{u_2}{2} - \frac{s - p_1}{6\lambda_2} - \frac{v - \bar{v}_1}{3} \quad \text{and} \quad y_2 = \frac{s - p_1}{3\lambda_2} - \frac{v - \bar{v}_1}{3}$$

Note that date 2 trading intensities depend on λ_2 only. Plugging the above expressions for x_2 and y_2 in the objective functions (37) and (38) gives the conditional profits as

$$\begin{aligned} E\left(\pi_2^I | \Phi_2^I\right) &= \frac{\left(f - p_1\right)^2}{4\lambda_2} + \frac{\lambda_2 u_2^2}{4} + \frac{\left(s - p_1\right)^2}{36\lambda_2} + \frac{\lambda_2\left(v - \bar{v}_1\right)^2}{9} \\ &- \frac{\left(f - p_1\right)u_2}{2} - \frac{\left(f - p_1\right)\left(s - p_1\right)}{6\lambda_2} - \frac{\left(f - p_1\right)\left(v - \bar{v}_1\right)}{3} \\ &+ \frac{\lambda_2 u_2\left(v - \bar{v}_1\right)}{3} + \frac{u_2\left(s - p_1\right)}{6} + \frac{\left(s - p_1\right)\left(v - \bar{v}_1\right)}{9} \\ E\left(\pi_2^S | \Phi_2^S\right) &= \frac{\left(s - p_1\right)^2}{9\lambda_2} + \frac{\lambda_2\left(v - \bar{v}_1\right)^2}{9} - \frac{2\left(s - p_1\right)\left(v - \bar{v}_1\right)}{9} \end{aligned} \qquad (39)$$

Recall that the insider knows the final liquidation value, i.e. $f \in \Phi_2^I$. Using the decomposition $(s - p_1) = (f - p_1) + (s - f)$ and $(s - p_1)^2 = (f - p_1)^2 + (s - f)^2 + 2(f - p_1)(s - f)$ the insider's expected profits can be equivalently written as

$$\begin{aligned} E\left(\pi_2^I | \Phi_2^I\right) &= \frac{\left(f - p_1\right)^2}{9\lambda_2} + \frac{\lambda_2 u_2^2}{4} + \frac{\left(s - f\right)^2}{36\lambda_2} + \frac{\lambda_2\left(v - \bar{v}_1\right)^2}{9} \\ &- \frac{\left(f - p_1\right)u_2}{3} + \frac{\left(s - f\right)u_2}{6} - \frac{2\left(f - p_1\right)\left(v - \bar{v}_1\right)}{9} \\ &+ \frac{\lambda_2 u_2\left(v - \bar{v}_1\right)}{3} - \frac{\left(f - p_1\right)\left(s - f\right)}{9\lambda_2} + \frac{\left(s - f\right)\left(v - \bar{v}_1\right)}{9} \ . \end{aligned} \qquad (40)$$

Step 2: date 1 trades

During the first trading round the insider chooses x_1 to maximize (2), or equivalently:

$$E\left[x_1\left(f - p_1\left(\omega_1\right)\right)\middle|\, \Phi_1^I\right] + E\left[E\left(\pi_2^I | \Phi_2^I\right) | \Phi_1^I\right] \ , \qquad (41)$$

where $E\left(\pi_2^I | \Phi_2^I\right)$ is as in eq. (40). Therefore when submitting his order x_1 the insider has to keep into account the impact of his trade on the price $p_1(x_1)$ *and* the speculator's forecast of the final liquidation value $s(x_1)$. Assume that under the insider's conjecture the first period price follows (20) and the speculator updates his beliefs on f according to eq. (12). The expression for x_1 depends on the insider's estimate –conditional on Φ_1^I– of the speculator's forecast, $E\left[E\left(f|\,\Phi_2^S\right)|\Phi_1^I\right]$, and the signal s depends on the speculator's conjecture of the form of x_1. The following Lemma gives the insider's estimate of the speculator's forecast s:

Lemma 1 *In equilibrium* $E\left(s-f|\,\Phi_1^I\right)=\chi\left(f-p_0\right)+\psi u_1$

Proof. Assuming that Lemma 1 holds, the insider's trading strategy (15) can be rewritten as $x_1=a_1\left(v-p_0\right)-b_1u_1$, and (16) obtains setting $a_1=\alpha+\gamma\chi$ and $b_1=-\left(\beta+\gamma\psi\right)$. Given eq. (16), the speculator updates his belief on f after the first round according to $s=p_0+\phi\left(x_1+u_1\right)$ –which is eq. (22)– with regression coefficient

$$\phi=\frac{cov\left(f,x_1+u_1|\,\Phi_1^S\right)}{var\left(x_1+u_1|\,\Phi_1^S\right)}=\frac{a_1\sigma_{f,0}^2}{a_1^2\sigma_{f,0}^2+\left(1-b_1\right)^2\sigma_{u_1}^2}\ .$$

Plugging eq. (16) into the speculator's forecast (22) and taking expectations conditional on Φ_1^I gives $E\left(s-f|\,\Phi_1^I\right)=\chi\left(f-p_0\right)+\psi u_1$, where $\chi=a_1\phi-1$ and $\psi=\left(1-b_1\right)\phi$. Recall that under equation (16) the parameters a_1 and b_1 depend on the insider's forecast via χ and ψ, such that

$$a_1=\frac{\alpha-\gamma}{1-\gamma\phi}\qquad\text{and}\qquad b_1=-\frac{\beta+\gamma\phi}{1-\gamma\phi}\ .$$

For these expressions for a_1 and b_1, and the speculator's update (22) the conjecture in Lemma 1 is verified. ■

Letting $\kappa_I=2\lambda_1+\phi-2\lambda_2\mu$, under conjectures (20) and (22) the optimality conditions[23] for the objective function (41) are:

$$E\left(f-p_1\left(\omega_1\right)|\,\Phi_1^I\right)\left(1-\frac{\kappa_I}{9\lambda_2}\right)-\lambda_1x_1+E\left(v-\bar{v}_1\left(\omega_1\right)|\,\Phi_1^I\right)\frac{\kappa_I}{9}+E\left(s-f|\,\Phi_1^I\right)\frac{\kappa_I}{18\lambda_2}=0$$

$$2\lambda_1-\frac{\kappa_I^2}{18\lambda_2}>0$$

The insider's first period trade is linear, i.e.

$$x_1=d^{-1}e\left(f-p_0\right)-\left(1-d^{-1}\lambda_1\right)u_1\ ,$$

yielding eq. (16) with a_1, b_1, and d as given in the main text, and $e=1-\kappa_I\left(6\lambda_2\right)^{-1}$.

The speculator chooses his first period trade y_1 maximizing (13), or equivalently

$$E\left[y_1\left(f-p_1\left(\omega_1\right)\right)|\,\Phi_1^S\right]+E\left[E\left(\pi_2^S|\Phi_1^S\right)\right]\ ,$$

where $E\left(\pi_2^S|\Phi_1^S\right)$ is as in eq. (39). Plugging the expression for p_1 [see eq. (20)], the speculator's update [see eq. (22)], and the conjecture for x_1 [see eq. (16)] into the latter results in the following optimality conditions:[24]

$$E\left(s-p_1\left(\omega_1\right)|\,\Phi_1^S\right)\left(1-\frac{2\kappa_S}{9\lambda_2}\right)-\lambda_1y_1+E\left(v-\bar{v}_1\left(\omega_1\right)|\,\Phi_1^S\right)\frac{2\kappa_S}{9}=0$$

$$2\lambda_1 - \frac{2\kappa_S^2}{9\lambda_2} > 0$$

where $\kappa_S = \lambda_1 - \lambda_2\mu$. The speculator's first period trade is linear in $(v - \bar{v}_0)$, i.e. $y_1 = D^{-1}\frac{2\kappa_S}{9}(v - \bar{v}_0)$, and eq. (18) obtains with C_1 and D as in the main text.

Step 3: Prices

We now turn to determine the equilibrium prices. Let $\mathbf{r}$ denote the 3×1 random vector containing the final liquidation value, the passive trades known by the speculator, and the speculator's signal about the final payoff, i.e. $\mathbf{r} \equiv (f, v, s)^\top$. For the $i-$th component of the vector $\mathbf{r}$ we denote the unconditional variance by $\sigma_{r_i,0}^2 = var(r_i)$; similarly for $i \neq j$ the unconditional covariance is $\sigma_{r_i r_j,0} = cov(r_i, r_j)$. Further the unconditional variance of the first period aggregate order flow is denoted by $\sigma_{\omega,0}^2 = var(\omega_1)$ and the covariance between ω_1 and $r_i \in \mathbf{r}$ is $\sigma_{r_i\omega,0} = cov(r_i, \omega_1)$. The unconditional distribution for the random variables (f, u_1, v) together with the first period trades (16) and (18) and the speculator's signal s [see eq. (22)] yields

$$\begin{aligned}
\sigma_{s,0}^2 &= \phi^2\left(a_1^2\sigma_{f,0}^2 + (1-b_1)^2\sigma_{u_1}^2\right) \\
\sigma_{\omega,0}^2 &= a_1^2\sigma_{f,0}^2 + (1-b_1)^2\sigma_{u_1}^2 + C_1^2\sigma_{v,0}^2 \\
\sigma_{fs,0} &= \phi a_1\sigma_{f,0}^2 \\
\sigma_{f\omega,0} &= a_1\sigma_{f,0}^2 \\
\sigma_{v\omega,0} &= C_1\sigma_{v,0}^2 \\
\sigma_{s\omega,0} &= \phi\left(a_1^2\sigma_{f,0}^2 + (1-b_1)^2\sigma_{u_1}^2\right)
\end{aligned} \tag{42}$$

Note that plugging the expression for ϕ [see eq. (24)] in the above variances and covariances gives $\sigma_{fs,0} = \sigma_{s,0}^2$ and $\sigma_{s\omega,0} = \sigma_{f\omega,0}$. Therefore the market maker's prior joint distribution for $(\mathbf{r}^\top, \omega_1)$ becomes

$$\begin{bmatrix} \mathbf{r} \\ \omega_1 \end{bmatrix} \sim N\left(\begin{bmatrix} E_{\mathbf{r},0} \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{\mathbf{r},0} & \Sigma_{\mathbf{r}\omega,0} \\ \Sigma_{\mathbf{r}\omega,0}^\top & \sigma_{\omega,0}^2 \end{bmatrix}\right),$$

where $E_{\mathbf{r},0} = (p_0, \bar{v}_0, p_0)^\top$, and

$$\begin{aligned}
\Sigma_{\mathbf{r},0} &\equiv E\left[(\mathbf{r} - E_{\mathbf{r},0})(\mathbf{r} - E_{\mathbf{r},0})^\top\right] = \begin{bmatrix} \sigma_{f,0}^2 & 0 & \sigma_{s,0}^2 \\ 0 & \sigma_{v,0}^2 & 0 \\ \sigma_{s,0}^2 & 0 & \sigma_{s,0}^2 \end{bmatrix} \\
\Sigma_{\mathbf{r}\omega,0} &\equiv E\left[(\mathbf{r} - E_{\mathbf{r},0})\omega_1\right] = \begin{bmatrix} \sigma_{f\omega,0} & \sigma_{v\omega,0} & \sigma_{f\omega,0} \end{bmatrix}^\top.
\end{aligned}$$

After observing the first period aggregate order flow the market maker updates his distribution for the random vector $\mathbf{r}$. The Projection Theorem together with eqs. (4) and (11) give[25] $p_1 = p_0 + \lambda_1\omega_1$ and $\bar{v}_1 = \bar{v}_0 + \mu\omega_1$, with regression coefficients $\lambda_1 = \sigma_{f\omega,0}/\sigma_{\omega,0}^2$ and $\mu = \sigma_{v\omega,0}/\sigma_{\omega,0}^2$ respectively. Using the variance-covariance matrix in (42) one has:

$$\begin{aligned}
\lambda_1 &= \frac{a_1\sigma_{f,0}^2}{a_1^2\sigma_{f,0}^2 + (1-b_1)^2\sigma_{u_1}^2 + C_1^2\sigma_{v,0}^2} \\
\mu &= \frac{C_1\sigma_{v,0}^2}{a_1^2\sigma_{f,0}^2 + (1-b_1)^2\sigma_{u_1}^2 + C_1^2\sigma_{v,0}^2}.
\end{aligned}$$

We now determine $\Sigma_{\mathbf{r},1}$ (the posterior variance for the vector $\mathbf{r}$). From $\sigma_{fs,0} = \sigma^2_{s,0}$ and $\sigma_{s\omega,0} = \sigma_{f\omega,0}$ it follows that $\sigma_{fs,1} = \sigma^2_{s,1}$ and $\sigma_{vs,1} = \sigma_{fv,1}$. Furthermore using the unconditional variance-covariance matrix in (42) together with the expression for ϕ, μ and λ_1 in the main text yields:

$$\begin{aligned} \sigma^2_{f,1} &= (1 - a_1\lambda_1)\,\sigma^2_{f,0} \\ \sigma^2_{v,1} &= (1 - C_1\mu)\,\sigma^2_{v,0} \\ \sigma^2_{s,1} &= a_1\,(\phi - \lambda_1)\,\sigma^2_{f,0} \\ \sigma_{fv,1} &= -a_1\mu\sigma^2_{v,0} \end{aligned} \tag{43}$$

Note that the conditional variance between the fundamental value and the passive trades v can alternatively be written as $\sigma_{fv,1} = -C_1\lambda_1\sigma^2_{f,0}$.

Therefore $\mathbf{r}|\omega_1 \sim N\left(E_{\mathbf{r},1}, \Sigma_{\mathbf{r},1}\right)$, where

$$E_{\mathbf{r},1} = E\left(\mathbf{r}|\omega_1\right) = \left(p_1, \bar{v}_1, p_1\right)^\top \ ,$$

and

$$\Sigma_{\mathbf{r},1} = E\left[\left(\mathbf{r} - E_{\mathbf{r},1}\right)\left(\mathbf{r} - E_{\mathbf{r},1}\right)^\top \middle| \omega_1\right] = \begin{bmatrix} \sigma^2_{f,1} & \sigma_{fv,1} & \sigma^2_{s,1} \\ \sigma_{fv,1} & \sigma^2_{v,1} & \sigma_{fv,1} \\ \sigma^2_{s,1} & \sigma_{fv,1} & \sigma^2_{s,1} \end{bmatrix}$$

In the second trading round the market maker sets prices according to (5). Using the expression for date 2 trades (17) and (19) one has:

$$\omega_2 = a_2\left(f - p_1\right) + u_2/2 + a_2\left(s - p_1\right)/3 + v/3 + w + 2\bar{v}_1/3 \ .$$

It follows that $\omega_2|\omega_1 \sim N\left(\bar{z}_1, \sigma^2_{\omega,1}\right)$ with $\bar{z}_1 = \bar{v}_1 + \bar{w}$, and

$$\sigma^2_{\omega,1} = var\left(\omega_2|\omega_1\right) = a_2^2\sigma^2_{f,1} + \sigma^2_{u_2}/4 + 7a_2^2\sigma^2_{s,1}/9 + \sigma^2_{v,1}/9 + \sigma^2_w + 8a_2\sigma_{fv,1}/9 \ .$$

Furthermore the conditional covariance between the fundamental value and the date 2 order flow is

$$\sigma_{f\omega,1} = cov\left(f, \omega_2|\omega_1\right) = a_2\sigma^2_{f,1} + a_2\sigma^2_{s,1}/3 + \sigma_{fv,1}/3 \ .$$

Letting $\lambda_2 = \sigma_{f\omega,1}/\sigma^2_{\omega,1}$, date 2 prices follow from the Projection Theorem:

$$p_2 = p_1 + \lambda_2\left(\omega_2 - \bar{z}_1\right)$$

which is eq. (21). Note that both $\sigma_{f\omega,1}$ and $\sigma^2_{\omega,1}$ depend on the insider's trading aggressiveness a_2, which in turn depends on λ_2 only. Since $a_2 = \left(2\lambda_2\right)^{-1}$, the regression coefficient λ_2 solves the quadratic equation

$$\Sigma\lambda_2^2 + 4\lambda_2\sigma_{fv,1} + \left(\sigma^2_{s,1} - 9\sigma^2_{f,1}\right) = 0 \tag{44}$$

where Σ is defined in the main text. Lemma 2 addresses the existence of real roots for eq. (44) as well as the uniqueness for λ_2.

Lemma 2 *In equilibrium λ_2 is given by:*

$$\lambda_2 = \Sigma^{-1}\left(-2\sigma_{fv,1} + \left[4\left(\sigma_{fv,1}\right)^2 + \Sigma\left(9\sigma^2_{f,1} - \sigma^2_{s,1}\right)\right]^{1/2}\right)$$

Proof. To prove Lemma 2 we proceed in two steps. First we show that both the solutions for λ_2 in eq. (44) are real; then we use the second order conditions to pin down the positive root and get λ_2 as in the main text. From eq. (44) a sufficient condition for λ_2 to belong to the real line is $\sigma^2_{f,1} > \sigma^2_{s,1}$. Making use of the conditional variances in (43) one has $\sigma^2_{f,1} - \sigma^2_{s,1} = (1 - a_1\phi)\,\sigma^2_{f,0}$. Substituting for the expression for ϕ in the latter gives:

$$\sigma^2_{f,1} - \sigma^2_{s,1} = \frac{(1-b_1)^2\,\sigma^2_{u_1}}{a_1^2\sigma^2_{f,0} + (1-b_1)^2\,\sigma^2_{u_1}}\sigma^2_{f,0} > 0$$

such that the roots for (44) are real. Recall that in equilibrium $\lambda_2 > 0$. Thus, regardless of the sign of $\sigma_{fv,1}$, the negative root in (44) can be discarded and λ_2 in the main text obtains. ■

Eventually the liquidation value residual variance after the second trading round becomes:

$$\sigma^2_{f,2} = \sigma^2_{f,1} - \lambda_2\sigma_{f\omega,1} = \left(3\sigma^2_{f,1} - \sigma^2_{s,1} - 2\lambda_2\sigma_{fv,1}\right)/6\ .$$

Therefore an equilibrium for the trading game is described by solutions for (a_1, b_1, C_1, μ, ϕ, λ_1, λ_2) in Proposition 2 subject to the nonlinear constraint (26).

As for the expected trading volume, note that date 1 orders have mean zero. It follows that the expressions for V_1^I, V_1^L and V_1^M are given by the first line in (36) [keeping into account that parameters a_1 and b_1 are as in Proposition 2 and $\sigma^2_{\omega,0}$ is given in eq. (42)]. Similarly $E_1(u_2) = 0$, such that V_2^L is like in the second line of (36). The other contributions to the trading volume are:

$$\begin{aligned}
V_1^S &= \frac{C_1\sigma_{v,0}}{\sqrt{2\pi}} \\
V_2^I &= \frac{\sigma_{x,1}}{\sqrt{2\pi}} \quad ; \quad V_2^S = \sqrt{\frac{C_2^2\sigma^2_{s,1} + \sigma^2_{v,1}/9 - 2C_2\sigma_{fv,1}/3}{2\pi}} \\
V_2^P &= \frac{\sigma_{z,1}}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\bar{z}_1}{\sigma_{z,1}}\right)^2} + \frac{\bar{z}_1}{2}\left(1 - 2\Phi\left(-\bar{z}_1/\sigma_{z,1}\right)\right) \\
V_2^M &= \frac{\sigma_{\omega,1}}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\bar{z}_1}{\sigma_{\omega,1}}\right)^2} + \frac{\bar{z}_1}{2}\left(1 - 2\Phi\left(-\bar{z}_1/\sigma_{\omega,1}\right)\right)
\end{aligned}$$

where $\sigma_{x,1} = \left(a_2^2\sigma^2_{f,1} + \sigma^2_{u_2}/4 - 5a_2^2\sigma^2_{s,1}/9 + \sigma^2_{v,1}/9 - 4a_2\sigma_{fv,1}/9\right)^{1/2}$ and $\sigma_{z,1} = \left(\sigma^2_{w,0} + \sigma^2_{v,1}\right)$.

Proof. (Corollary 1) From the date 2 second order condition $\lambda_2 > 0$. Therefore from the date 1 second order condition $\lambda_1 > 0$, since $\lambda_1 > \kappa_s^2/18\lambda_2 \geq 0$. Now suppose that $C_1 \leq 0$. From the expression for C_1 this implies that $\kappa_s \leq 0$ and from the expression for μ this implies that $\mu \leq 0$. Since κ_S is defined by $\kappa_S = \lambda_1 - \lambda_2\mu$ one has $\kappa_S \leq 0$ if and only if $\lambda_1 \leq \lambda_2\mu$. However $\lambda_2 > 0$ and $\mu \leq 0$, implying $\lambda_1 \leq 0$ which cannot occur in equilibrium. Finally note that the insider's second order condition can be written as $2\lambda_1 - \frac{2\kappa_S^2}{9\lambda_2} - \frac{\phi^2+4\kappa_S\phi}{18\lambda_2} > 0$. Since $\kappa_S > 0$, then $2\lambda_1 - \frac{2\kappa_S^2}{9\lambda_2} - \frac{\phi^2+4\kappa_S\phi}{18\lambda_2} < 2\lambda_1 - \frac{\phi^2+4\kappa_S\phi}{18\lambda_2}$ and (26) is sufficient for $2\lambda_1 - \frac{2\kappa_S^2}{9\lambda_2} > 0$. ■

References

Admati, Anat.R., and Paul C. Pfleiderer, 1988, A Theory of Intraday Patterns: Volume and Price Variability, *Review of Financial Studies*, 1, 3-40

Beneish, Messod D., and Robert E. Whaley, 1996, An Anatomy of the "S&P Game": the Effects of Changing the Rules, *Journal of Finance*, 51, 1909-30

Beneish, Messod D., and J.C. Gardner, 1995, Information Costs and Liquidity Effects from Changes in the Dow Jones Industrial Average List, *Journal of Financial and Quantitative Analysis*, 30, 135-57

Blume, Marshall E., and Roger M. Edelen, 2004, S&P 500 Indexers, Delegation Costs and Liquidity Mechanisms, Working Paper 04-03, The Wharton School

Bos, Roger J., 2000, *Event Study: Quantifying the Effect of Being Added to an S&P Index*, Standard and Poor's report

Chen, Honghui, Gregory Noronha, and Vijay Singal, 2004, The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation, *Journal of Finance*, 59, 1901-1929

Cusick, Philip A., 2002, Price Effects of Addition and Deletion from the Standard and Poor's 500 Index, *Financial Markets, Institutions & Instruments*, 11, 349-383

Deininger Claus, Kaserer, Christoph, and Stephanie Roos, 2000, Stock Price and Volume Effects Associated with Index Replacements in Germany, Working Paper, University of Wuerzburg

Edmister, Robert O., A. Steven Graham, and Wendy L. Pirie, 1996, Trading Cost Expectations: Evidence from S&P 500 Index Replacement Stock Announcements, *Journal of Economics and Finance*, 20, 75-85

Erwin, Gayle R., and James M. Miller, 1998, The Liquidity Effects Associated with Addition of a Stock to the S&P 500 Index: evidence from bid/ask spreads, *Financial Review*, 33, 131-146

Foster, F. Douglas. and S. Viswanathan, 1994, Strategic Trading with Asymmetrically Informed Traders and Long-Lived Information, *Journal of Financial and Quantitative Analysis*, 29, 499-518

Foucault, Thierry, and Laurence Lescourret, 2003, Information Sharing, Liquidity and Transaction Costs in Floor-Based Trading Systems, *Finance*, 24, 45-78

Harris, Lawrence, and Eitan Gurel, 1986, Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures, *Journal of Finance*, 41, 815-29

Hollander, Myles, and Douglas A. Wolfe, 1999, *Nonparametric Statistical Methods*, New York: John Wiley and Sons

Jain, Prem C., 1987, The Effect on Stock Price of Inclusion in or Exclusion from the S&P 500, *Financial Analysts Journal*, 43, 58-65

Kyle, Albert S., 1985, Continuous Auctions and Insider Trading, *Econometrica*, 53, 1315-35

Lin, Ji-Chai, Gary C. Sanger, and G. Geoffrey Booth, 1995, Trade Size and Components of the Bid-Ask Spread, *Review of Financial Studies*, 8, 1153-1183

Lynch, Anthony W., and Richard R. Mendenhall, 1997, New Evidence on Stock Price Effects Associated with Changes in the S&P 500 Index, *Journal of Business*, 70, 351-83

Madhavan, Ananth, 2003, The Russell Reconstitution Effect, *Financial Analysts Journal*, 59, 51-64

Madrigal, Vicente, 1996, Non-Fundamental Speculation, *Journal of Finance*, 51, 553-78

Polonchek, John, and Tim Krehbiel, 1994, Price and Volume Effects associated with Changes in the Dow Jones Averages, *Quarterly Review of Economics and Finance*, 34, 305-316

Pruitt, Stephen W., and K.C. John Wei, 1989, Institutional Ownership and Changes in the S&P 500, *Journal of Finance*, 44, 509-513

Rochet, Jean-Charles, and Jean-Luc Vila, 1994, Insider Trading without Normality, *Review of Economic Studies*, 61, 131-152

Shleifer, Andrei, 1986, Do Demand Curves for Stocks Slope Down?, *Journal of Finance*, 41, 579-90

Singal, Vijay, 2003, *Beyond the Random Walk: a Guide to Stock Market Anomalies and Low-Risk Investing*, Oxford: Oxford University Press

Standard and Poor's, 2003, *Annual Survey of S&P Indexed Assets*, Standard and Poor's Index Services

Wurgler, Jeffrey, and Ekaterina Zhuravskaya, 2002, Does Arbitrage Flatten Demand Curves for Stocks?, *Journal of Business*, 75, 583-608

Yu, Fan, 1999, What is the Value of Knowing Uninformed Trades?, *Economics Letters*, 64, 87-98

Notes

[1]Investigations on indices other than the S&P 500 can be found in Bos (2000) and Jain (1987) who report evidence for supplementary S&P Indices. Madhavan (2003) focuses on changes to the Russell 2000 and 3000 indices. The Dow Jones Industrial Average is considered in Beneish and Gardner (1995) and Polonchek and Krehbiel (1994), while Deininger, Kaserer and Roos (2000) study the German DAX and MDAX.

[2]We stress the fact that indexers enter the market at the second date only with the subscript 2 in the passive industry demand.

[3]As in Kyle (1985) the efficiency condition arises from price competion à-la-Bertrand in the market making sector. In equilibrium one can consider a single market maker that operates according to a zero expected profits condition.

[4]One can accomodate the insider receiving a noisy signal of the fundamental value rather than the realization f. In this case the insider's informational advantage would be captured by the signal to noise variance, rather than the fundamental value variance only. Qualitatively this does not affect our main conclusions.

[5]We use depth and liquidity as synonyms, even though they capture different aspects of market behaviour.

[6]The insider's informational advantage with respect to the fundamental value is captured by the ratio between fundamental and non-fundamental uncertainty, like in Kyle (1985) and its various extensions. Since we are focusing on the role of non-fundamental information, we refer to k_I as the insider's advantage only.

[7]Parameters at equilibrium are computed using the following values for $k_I = 0.01$, 0.5, 1, 2, 3.5, 5, 7.5, 10, 15 and 20.

[8]This is not surprising since parameters for RV can be obtained from Proposition 1 setting $\sigma^2_{z,0} = 0$ (appendix A contains further details). For $k_I = 500$ the difference between the parameters in Proposition 1 and their RV counterparts is of the order of 10^{-4}. For reasons of space we consider $k_I \leq 20$ in figures 2–6.

[9]The following expression follows from the fact that insider's orders are conditionally normal with mean zero, implying that the volume is (a multiple of) the conditional standard deviation. The total volume pattern is discussed in section 4.

[10]We note that it might be difficult to justify that both the speculator and the insider observe the *same* fraction of date 1 liquidity orders like in Madrigal (1996).

[11]A related problem is that other fund managers might implement mimicking techniques only for a fraction of their portfolios, and thus would not show up in surveys on completely passive funds like the Standard and Poor's (2003).

[12]Madrigal (1996) imposes a hierarchical information structure as well. As previously noted in his model the insider and the speculator share the knowledge of past noise trades when trading at date 2. This way the insider knows –on top of the final liquidation value– the speculator's informational advantage relative to the market maker, i.e. the difference $s - p_1$, like in our specification

[13]On the other hand the first period information sets are non-nested. This assumption can be easily modified including v into Φ_1^I. In this case, the speculator would lose his informational advantage (with respect to the insider) and competition between the two informed agents would arise at date 1. However, given the previous considerations on the difficulty in gathering information about passive funds' techniques, we do not regard this situation as particularly interesting.

[14]In equilibrium $E\left(s|\,\Phi_1^I\right) = f + \chi\left(f - p_0\right) + \psi u_1$ and $s = p_0 + \phi\left(x_1 + u_1\right)$ are mutually consistent, and as a consequence the insider's first period trade can be expressed as $x_1 = a_1\left(f - p_0\right) - b_1 u_1$. We derive the expression for coefficients χ, ψ and ϕ in appendix B (see Lemma 1).

[15]The first period aggregate order flow does not contain information about u_2. Therefore conditional on ω_1 the second period liquidity trades have mean zero. Recall that liquidity trades are independent through time and orthogonal to passive funds' orders.

[16]Recall that the speculator has long-lived non-fundamental information, as in Kyle (1985) the insider has long-lived information about the final liquidation value f. Thus it is not surprising that the speculator's behaviour with respect to v closely resembles the insider's aggressiveness on f in Kyle (1985).

[17]V_1^M decreases with k_I, but a slower rate than the increase for V_1^I; moreover V_1^L does not depend on k_I.

[18]We are grateful to Nicholas Barberis and Jeffrey Wurgler for sharing their dataset.

[19]Other authors use market adjusted trading turnover, i.e. $V_{i,t}/V_{M,t}$, where $V_{M,t}$ is the NYSE volume during day t. Market adjustment results in stronger tests by taking into account market variation. Harris and Gurel (1986) report that the qualitative results are not affected by the way one measures trading volume. Accordingly, our results are qualitatively the same when using raw trading volume $V_{i,t}$. Cusick (2002) and Lynch and Mendenhall (1997) use a logarithmic transformation of the market adjusted trading volume.

[20]The evidence reported in this and the following subsections is quite robust to the pre-event window choice. Inclusion of the announcement day in computing average volume and bid-ask spread does not qualitatively change our results.

[21]When computing the $t-$statistic we opt for the cross-sectional dispersion of $MAVR_t$ to estimate its variance. See Lynch and Mendenhall (1997) for an alternative method of computing standard errors. Details on the binomial test are in Hollander and Wolfe (1999).

[22]Using the effective relative spread $2\left|\ln\left(P_{i,t}/Q_{i,t}\right)\right|$ does not change the results presented here [for a definition of the effective relative spread, see Lin, Sanger and Booth (1995)]. For reasons of space we report findings for the relative spread only.

[23]Using eqs. (20) and (22) the terms in the insider's profits (40) can be written as functions of x_1 according to $(f-p_1)=(f-p_0)-\lambda_1\left[x_1+C_1\left(v-\bar{v}_0\right)+u_1\right]$, $(s-f)=-\left(f-p_0\right)+\phi\left(x_1+u_1\right)$ and $(v-\bar{v}_1)=(v-\bar{v}_0)\left(1-C_1\mu\right)-\mu\left(x_1+u_1\right)$.

[24]Note that (20), (22) and (16) allow to write $f-p_1=\left(1-a_1\lambda_1\right)\left(f-p_0\right)-\lambda_1\left[y_1+\left(1-b_1\right)u_1\right]$, $s-p_1=\left(\phi-\lambda_1\right)\left[a_1\left(v-p_0\right)+\left(1-b_1\right)u_1\right]-\lambda_1 y_1$ and $v-\bar{v}_1=-\mu\left[a_1\left(f-p_0\right)+\left(1-b_1\right)u_1\right]-\mu y_1+\left(u_2-\bar{u}_0\right)$. Further, since the date 1 speculator's information set does not include f one has $E\left(f-p_1\middle|\Phi_1^S\right)=E\left(s-p_1\middle|\Phi_1^S\right)=-\lambda_1 y_1$.

[25]From the definition of s one has $E\left(s|\omega_1\right)=E\left[E\left(f|x_1+u_1\right)|\omega_1\right]$. Since ω_1 is coarser than x_1+u_1, then using the Law of Iterated Expectations $E\left[E\left(f|x_1+u_1\right)|\omega_1\right]=E\left(f|\omega_1\right)=p_1$.