Stock returns: Is it just Supply and Demand?

Carl Hopman*†

Preliminary draft. Please do not distribute.

ABSTRACT

This paper proposes a simple model of price pressure and empirically shows that stock returns are in large part due to supply and demand imbalances, and not exclusively due to information.

Since the seminal work of Roll (1988), the failure of news to explain a significant amount of variation in stock returns has been a major asset-pricing puzzle. Using a measure of order flow imbalance, which is readily available on the Paris Bourse, I run specifications to explain stock returns with simultaneous supply and demand imbalances, and get R^2s around 50% on a random sample of 34 stocks. This result suggests that mechanical price pressure through supply and demand imbalances explains price changes better than information.

To confirm this interpretation, I show that neither public nor private information is driving the correlation between return and order flow. I use the results of French and Roll (1986) to argue that the effect of public information on prices is too small to explain the correlation. To test a private information setup in the spirit of Kyle (1985), I then assume that the price impact of orders is due to the private information that, on average, they contain. This would have two implications, which are both rejected empirically.

First, since the potential for leakage of company-specific information is greater than that of market-wide information, one would expect company-specific orders to be more privately informed and to have a higher price impact than market-wide orders. However, I find this is not the case empirically.

Second, one would expect the stock market regression to have a lower R^2 than individual company regressions, as the R^2 would correspond to the fraction of private information. Again, this is rejected empirically as the stock market regression has an even higher R^2 of 70%.

Having seen that information doesn't really succeed in explaining the correlation between order flow and price changes, I conclude that this correlation is probably due to mechanical price pressure, which explains a big fraction of stock price changes.

I finally review intriguing empirical facts that are well explained by a price pressure framework.

*I would like to thank Cécile Boyer, Xavier Gabaix, Denis Gromb, S.P. Kothari, Jonathan Lewellen, Andrew Lo, Angelo Ranaldo, Dimitri Vayanos and Jiang Wang for their comments and suggestions. I am especially grateful to Xavier Gabaix for stimulating discussions and continuous encouragement for this project. I would also like to thank Mathias Auguy, Patrick Hazart and Bernard Perrot from the Paris Bourse for their help in making the data available.

†chopman@mit.edu, Massachusetts Institute of Technology.

1 Introduction

Since at least Scholes (1972), the prevailing paradigm has been that price changes are only due to new information. The fact that new information indeed has an impact on price has been confirmed by numerous event studies.

The vision of the world held before this information-efficient paradigm, by Keynes and Lintner for example, was that an excess demand for shares would drive prices up and that an excess supply of shares would drive them down. This is often called the price pressure theory. Today's financial press probably subscribes to this explanation: it often stresses the excess of supply or demand to justify price changes.

There exists a microstructure explanation of the price pressure which is unrelated to information: a buy order will push the price upwards for inventory reasons with a market maker, mechanically with a limit order market. However, the current paradigm would suggest that this effect is only temporary: arbitrageurs will soon provide the necessary liquidity to bring back the price back to its previous[1] level. But it is also possible that these arbitrageurs don't bring it completely back, because the benefit is small and the risk is high (the price will indeed one day come back to its efficient value but you may need to wait a long time and face huge price changes in between). In sum, noisy trades are not faced by infinite liquidity, and therefore have a long-term impact on the price.

Roll (1988) tries to test whether the prevailing paradigm is supported by the data, that is, whether news can explain stock price movements. He finds that news do not explain the stock returns very well. The $\bar{R}^2$ on principal components is under 30%. It changes only by 2 percentage points when distinguishing between days with company specific news and without. So the company specific variance hardly changes whether there is company specific news or not (as a fraction of total variance, it goes from 80% with all dates to 78% when there is no company specific news). Where is this high company specific variance coming from when there is no news?

In contrast, my results indicate that the price pressure hypothesis is very consistent with the data. I avoid the traditional pitfall in defining a supply and demand imbalance for stocks: the equality between the volume bought and the volume sold. I construct the order flow imbalance using submitted limit orders, which are not necessarily executed and can therefore be unbalanced. I find that when the return is regressed on the order flow imbalance, the $\bar{R}^2$ is[2] around 50%. It appears that supply and demand do move stock prices, when they are not balanced, as we were told for widgets in Economics 101.

To confirm this interpretation, I show that neither public nor private information is driving the correlation between return and order flow. I use the results of French and Roll (1986) to argue that the effect of public information on prices is too small to explain the correlation.

To test a private information setup in the spirit of Kyle (1985), I then assume that the price impact of orders is due to the private information that, on average, they contain. This would have two implications, which are both rejected empirically.

[1] Modified for the information which has arrived in between.

[2] The reported $\bar{R}^2$ are corrected for the number of degrees of freedom.

First, since the potential for leakage of company-specific information is greater than that of market-wide information, one would expect company-specific orders to be more privately informed and to have a higher price impact than market-wide orders. However, I find this is not the case empirically.

Second, one would expect the stock market regression to have a lower R^2 than individual company regressions, as the R^2 would correspond to the fraction of private information. Again, this is rejected empirically as the stock market regression has an even higher R^2 of 70%.

Having seen that information doesn't really succeed in explaining the correlation between order flow and price changes, I conclude that this correlation is probably due to mechanical price pressure, which explains a big fraction of stock price changes.

Several previous papers distinguish buyer initiated and seller initiated transactions to explain price changes. Although I use a direct measure of the orders submitted on the Paris Bourse, this transaction-based measure is related to the order flow since an excess of potential buy orders is likely to generate aggressive buy orders which result in a buyer initiated transaction. Another distinction with previous work is that I have a sceptical view relative to the "all information" paradigm. Hasbrouck (1991) uses NYSE transaction data and concludes that the impact of a trade is a positive, increasing and concave function of its size, with R^2 of 10%. However, he takes as given that only information can have a long term (20 transactions) impact on the price, and thus interprets his finding as due to private information. Evans and Lyons (2000) also use signed transactions data, this time on the Foreign Exchange dealer market, and find R^2 similar to mine, showing that the imbalance explains the Forex returns very well. Compared to mine, their data only include realized transactions, without their volume or intraday information on a sample of 89 trading days. They interpret their finding in a private information framework. However, they also call private information the volume traders are willing to buy or sell, which I do not view as being so different from a liquidity-based interpretation. In some sense one can argue that even noise traders have some private information: they know before anyone else the order they're going to submit. However, calling this private information obscures the fact that these orders are placed without any superior knowledge about the security's cash flows or the applicable discount rate[3].

Other papers look at the impact of imbalances on stock prices by restricting themselves to large trades. This literature started with Scholes (1972). He finds that the impact of a trade does not increase with the block size, and concludes by rejecting the price pressure hypothesis. Yet in a later study of large trades, Holthausen, Leftwich and Mayers (1990) use transaction data, which yield more precise estimates of the impact, and find that the impact does increase with the trade size (as Hasbrouck (1991) finds without restricting himself to large trades).

A direct implication of the impact of trades and orders on asset prices is that an increase in trading activity will increase volatility. Many papers document a consistent link between volume and volatility, surveyed for example in Karpoff (1987).

[3]In the same spirit, one can argue that public information includes the price change itself, so that no price change can happen without being "driven" by public information.

In some cases, it is even possible to solve for the endogeneity issue[4]: whether an increased information flow generates additional volatility and trading, or whether trading activity is at the origin of the additional volatility even in the absence of news. Indeed, trading activity is in some cases determined exogenously. For example, French and Roll (1986) study days when the exchange is closed (so the trading activity is reduced) but the rest of the economy is active (so the information flow is the same). They find that the variance on these days is only 14.5% of what it is when the exchange is open. This suggests that trading creates volatility rather than the other way around.

A final branch of the literature which supports the price pressure hypothesis started with Shleifer (1986) and studies the inclusion and exclusion of companies from the S&P 500. Index funds and other funds using the S&P 500 as a benchmark need to buy the incoming stock and sell the outgoing one, so the price pressure hypothesis naturally explains the observed impact on the price. Furthermore it does not seem that new information is revealed on these days.

My first result is in explaining the stock return better than previous attempts have done. The supply and demand imbalance by itself explains around 50% of the return. The second result is in distinguishing between private information and direct price pressure. Using the lower information assymetry on the market index relative to individual companies, I can reject the private information interpretation of the order flow impact.

To summarize, I propose an answer to Roll's question: "Why is the R^2 so small?" and I exhibit the higher R^2. The explanation is that news is not the only determinant of price changes; the supply and demand disequilibrium also has a strong and direct impact on prices.

In the next section, I present a simple model of price pressure. I then describe the data from the Paris Bourse and the variables I used. In the fourth section, I present the high R^2 results. I then distinguish between private information and price pressure. I then conclude by reviewing various empirical facts which are well explained by a price pressure model and look at some implications of this work.

2 A simple Model of Price Pressure.

2.1 A mechanical impact.

I now look into more detail at the price pressure mechanism and how it can happen directly in the absence of private or public information. In the case of a market maker, Stoll (1978) has shown how inventory considerations could induce the market maker to move the price when he is faced with an order flow imbalance. However, as I have mentioned, with a market maker who regularly clears his inventory, the "order

[4] Even these cases of exogenous trading activity do not fully resolve the simultaneity problem, because privately informed traders may choose to reveal their information as a constant proportion of trading activity, as French and Roll (1986) suggest.

flow[5]" has to be balanced (since the market maker takes the opposite side of each trade and clears his inventory regularly). Although the imbalance probably exists in the feelings of participants, it is hard to define it rigorously, let alone to measure it.

Buy	Sell
\$99 (10 shares)	\$101 (10 shares)
\$98 (30 shares)	\$102 (10 shares)
\$97 (20 shares)	\$103 (40 shares)

Table 1: Example, the order book before a market buy order arrives.

I therefore turn to the case of the limit order market, where it is possible to measure the imbalance, and where I find it is highly correlated with price changes. For this type of market, it is clear that big market orders have a short term mechanical impact on the price, as we can see on the following example. For instance, let's assume that the order book is like Table 1, when a buy market order of 40 shares is submitted. It matches the book at \$101, \$102 and buys 20 shares at \$103. The new ask price is \$103. The mid-quote has gone up from \$100 to \$101. What a believer in the prevailing paradigm would argue is that this impact is only temporary, unless the order was informed. But this presupposes that some arbitrageurs will bring the price back to its "normal" value. If the incentives for the arbitrageurs to do so are not high enough (the price will indeed one day come back to its efficient value but you may need to wait a long time and face huge price changes in between), the short term impact may take some time to disappear (at least three months in Hopman (2002a)).

This price pressure framework explains naturally how market orders can move the price. As for the impact of limit orders, it is indirect: because sell limit orders provide additional liquidity on the sell side, a buy market order will have a smaller positive price impact. In my example, if someone places a limit sell order of 40 shares at \$101, the market buy order will result in a transaction price of only \$101 and a mid-quote of only \$100. So it prevents the market order to move the mid-quote up to \$101. Therefore, sell limit orders have an indirect negative impact on the price.

2.2 A simple descriptive model

Limit orders are providing liquidity, whereas market orders are demanding liquidity. However, they can both have a mechanical impact on the price as described above. To get the intuition of the implications of price pressure, I do not model the endogenous choice of liquidity demand and supply, that is market vs limit orders[6]. Instead, I assume that all orders have the same impact on the price: buy orders push the log-price by $+\lambda$ and sell orders by $-\lambda$. I also assume that the direction of the order is

[5]Defined here, with transactions instead of orders, as the volume of buyer-initiated transactions minus the volume of seller-initiated transactions.

[6]Implicitly, I assume that there are enough sell limit orders to provide liquidity for the buy market orders to avoid market breakdowns, and the other way round.

distributed randomly Buy or Sell, with an iid Bernoulli distribution, like a flipping coin[7]. If there are N_t orders between time 0 and t, the log-price change can be written:

$$p_t - p_0 = \lambda \sum_{i \leq N_t} \epsilon_i$$

where $\epsilon_i = +1$ for a buy order and -1 for a sell order.

This simple model predicts that the log price follows a random walk, thanks to the Central Limit Theorem. This result, which has been attributed to information, also ensues naturally from a price pressure model. Moreover, this price pressure model predicts that the log-price will follow a random-walk in transaction time[8] and not in physical time, as has been empirically documented by Ané and Geman (2000).

Finally, this random-walk result has important implications for the behavioral litterature. It shows that behavioral trading patterns can have an impact even if they are not systematically in the same direction: many irrational random orders will not perfectly cancel each other. Instead, this imperfect cancellation produces a random walk as I have described above. So you do not need a systematic crowd behavior to move stock prices. Random trades will do just as well[9].

This model is very simplistic. Among other things[10], it predicts that prices deviate infinitely from fundamentals[11]. To be more realistic, we need to assume that some rational "arbitrageurs" are ready to short the market when it is grossly overvalued and

[7]Again, this is a simplification, because the order flow is autocorrelated. However, the part of the order flow which is predictable has no impact on the price, as a simple aribitrage argument would predict. So I omit it here.

[8]In this simple model, the distinction between market orders (which produce a transaction) and limit orders (which do not) is blurred. Empirically however, the intensity of limit and market order submission are very correlated, so that the result of Ané Geman (2000) would probably extend with order time instead of transaction time.

[9]In fact, I find empirically in Hopman (2002a) that the fraction of the order flow which is predictable has no impact on the price, suggesting that an easily arbitrageable crowd behavior has no impact.

[10]In this simple model, I have only considered one asset (the stock market). However, my empirical results show that the price pressure also works for each stock individually. Price pressure is harder to model in this case, because risk averse "arbitrageurs" can build portfolios with apparently very high Sharpe ratios, and therefore remove a big fraction of the mispricing even with a relatively small fraction of the global wealth. Indeed, an "arbitrageur" could invest in a long/short portfolio, which removes the market component of risk, and diversify the idiosyncratic risk. However, there are several difficulties in following this strategy. First, this long-short portfolio will be heavily loaded in the book-to-market factor of Fama and French (1993). So it is in fact risky. I would also need to explain where this factor in stock returns comes from. I believe it is created by the trading of these "arbitrageurs" themselves when they get or lose money: as they invest in and out of their long/short portfolio, they generate a price pressure. This moves the price of the undervalued stocks together and in the opposite direction of the overvalued stocks (which they have in short position). Second, there is a lot of uncertainty in the distribution of future returns. The true fundamental value is difficult to estimate (a high price relative to book value could signal a growth company as well as an overvalued company). And the true covariance structure with many assets is also hard to estimate, which makes diversification harder. So even without the book-to-market factor, it would be difficult to build a very high Sharpe ratio portfolio.

[11]The cumulative imbalance between supply and demand can go to infinity over time, because it is an imbalance in submitted orders, not in realized transactions, for which there would never be any imbalance between volume bought and sold.

leverage their investment in the stock market when it is instead undervalued. This will create a dividend yield effect in the time series, as reported by Fama and French (1988), as well as the long term mean-reversion reported by Poterba and Summers (1988). When prices are high relative to fundamentals, they come back down. When they are low, they come back up.

3 Data and Definitions

3.1 The data

One of the most frequent argument against price pressure is that there is "no imbalance" in supply and demand of stocks. Indeed, the volume bought is equal to the volume sold when you look at realized transactions. To get around this problem, some researchers have distinguished between buyer and seller initiated transactions. However, this distinction does not solve the equality objection in a pure market maker setup where no limit orders are allowed. Indeed, suppose that only market orders are allowed, with a market maker who clears his inventory regularly[12], then even if the econometrician knows perfectly whether the market maker was on the buy side or the sell side, the total volume sold to the market is equal to the volume bought from it after each inventory clearing. Therefore, there is never any imbalance in volume. This property limits the effectiveness of using transactions data to measure the order flow in a pure market maker setup, and probably extends to markets where limit orders are the exception rather than the norm.

This is the main reason why I use the Paris Bourse data: limit orders are the norm not the exception, and their submission is available to the econometrician. Although in transaction terms the volume bought is still equal to the volume sold, in submission terms there can be many submitted orders which are never executed. There can therefore be an imbalance between submitted Buy and Sell orders (some are later executed, some are not) which I measure directly with this data set[13].

Although market analysts may feel the imbalance on markets which rely on market-makers instead of limit orders, it is not obvious to define or to measure it. On the Paris Bourse on the contrary, limit orders are the dominant form of orders and

[12] This condition can be weakened to having a bounded inventory with the equality between buy and sell volume true in the limit.

[13] If one goes deeper, one could ask what happens to the unexecuted limit orders. They of course get cancelled, most of them automatically at the end of the day or the month. If the impact of all the cancellations is equal to that of all the submissions, then again the total net (submitted minus cancelled) volume of orders is equal on the buy and sell side with a monthly time frame (and both are equal to the transaction volume). However, submissions are usually made relatively close to the current best quote, whereas cancellations often happen automatically, after the price has moved away and the submitted order has remained. If one considers a cancellation as a negative submission, then I have found empirically that submitted orders, for a given volume, have a decreasing impact when submitted further away from the best quotes. As a first approximation, one can then argue that submissions are close enough to the best quotes to have an impact, whereas cancellations are not and can be ignored. This is the approximation I am forced to make, since I do not have the cancellation data. Having this data would allow us to have an even better estimation of the impact of supply and demand. As we'll see, the approximation already yields very good results.

the imbalance is easy to measure. Besides, the Paris Bourse data is very clean and very complete. Because the Bourse is a fully automated electronic exchange[14], it is virtually free of errors.

The dataset goes from January 1995 to October 1999 included. I only look at the continuous trading session, which, until September 19 1999, started just after 10AM and finished at 5PM. From September 20 1999, it started at 9AM and finished at 5PM[15].

The dataset includes all the transactions and all the orders[16] that were submitted on the Paris Bourse, as well as the best quotes available at any time. In comparison, the TORQ (Trades, Orders, Reports and Quotes) database for the New-York Stock Exchange (NYSE) misses about half the total volume of submitted orders (Kavajecz (1999)).

The main French Index is the CAC40, which includes the 40 biggest stocks. The 40 stocks I study are the ones that defined the CAC40 in January[17] 1995. At the end of the sample 34 of them were still quoted as independent companies, so the results of this paper are provided for these 34 stocks.

To make things more concrete, I present the univariate results on one stock, Lafarge, which seemed to have average properties in many directions. The results are then summarized for all the 34 stocks.

3.2 Variable definitions

The regressions I present calculate the (log) return using mid-quotes. But I performed robustness checks using transaction prices instead of the mid-quote and got nearly identical results.

The time horizon used in this paper is the half hour. However, similar results were obtained with horizons from 10 min to 3 months, as reported in Hopman (2002a). The trade-off is between statistical power, which increases with the number of observations, and microstructure noise, which increases at short horizons. I chose 30 mn because the "bid-ask bounce" becomes insignificant at this horizon (no negative autocorrelation of the mid-quote), suggesting that temporary short-term noise is small at this horizon.

I distinguish the different buy orders (and similarly the sell orders) according to the level of urgency chosen by the trader submitting the order. This also corresponds to the rapidity with which it is likely to be executed. The reason for this distinction

[14] Biais, Hillion and Spatt (1995), Hamon and Jacquillat (1992) and Huang and Stoll (1991) provide a detailed description of the microstructure of the Paris Bourse.

[15] There are also 2 call auctions, one just before the opening, the other at 5:05PM which was created on June 2 1998, but I remove all the order flow data they generated, because it is harder to define and measure order imbalance in these auctions.

[16] The dataset distinguishes between visible and hidden orders. The latter are not visible by any trader or broker. However, the impact of both were quite similar, and I do not distinguish between them in this paper.

[17] This sample allows me to check for survivorship bias. However, since the results were similar for the 6 stocks that disappeared, and to be able to implement principal component analysis, I report results only for the 34 surviving stocks.

is that I expect more urgent orders (of similar size) to have a bigger impact on the price[18]. The basic distinction is between:

1. Market orders, which are executed immediately.

2. Spread orders, which are submitted between the best bid and ask (and thus change either the bid or the ask).

3. Book orders, which are submitted inside the order book[19].

To be more precise, a buy order:

- is called "market" if executed immediately, ie all the market orders and limit orders such that $P \geq ask$.

- is called "spread" if placed within the spread: $ask > P > bid$

- I calculate the log-mid-point $p_0 = \frac{ln(bid)+ln(ask)}{2}$

- is called "book" if $\ln(bid) \geq \ln(P) > p_0 - 0.005$

To have consistent and symmetric definitions, I use the natural logarithm. However one can consider the mid-quote as being roughly the arithmetic average between the bid and the ask price. One can also think of "book" orders as being above the mid-quote minus 0.5% (for buy orders). I use the best quotes available when the order is submitted, not outdated ones from the beginning of a time interval.

Besides, I use in this paper the SQRT method of aggregation. Indeed, we have found in Gabaix and Hopman (2002) and Hopman (2002a) that the impact of an order as a function of its volume is a concave function of this volume, well described by the square root function. When aggregating, I want to transform each order in something close to its own price impact, so as to obtain the "total price impact" when adding[20] up.

- I call v_i the volume in share turnover[21] of each order.

- The aggregate square root of volume is SQRT$= \sum_i (v_i)^{0.5}$.

The three order flow variables I use are thus:

1. Market$= \sum_{i \in \text{market buy}} (v_i)^{0.5} - \sum_{i \in \text{market sell}} (v_i)^{0.5}$

2. Spread$= \sum_{i \in \text{spread buy}} (v_i)^{0.5} - \sum_{i \in \text{spread sell}} (v_i)^{0.5}$

3. Book$= \sum_{i \in \text{book buy}} (v_i)^{0.5} - \sum_{i \in \text{book sell}} (v_i)^{0.5}$

	Lafarge
Sample size	16878
Volatility	0.46%
Number of orders	
market buy	17.88
market sell	18.35
spread buy	7.20
spread sell	6.90
book buy	12.25
book sell	11.75
Average volume of one order $\times 10^6$	
market buy	7.3
market sell	7.0
spread buy	9.6
spread sell	9.7
book buy	11.5
book sell	12.4

Table 2: Summary statistics for Lafarge over 30 min.

The results are reported for the Lafarge stock over 30 min.

	Mean	Std. Dev.
Sample size	16878	0
Volatility	0.45%	0.058%
Number of orders		
market buy	17.24	9.01
market sell	20.17	13.58
spread buy	5.48	2.16
spread sell	5.22	2.00
book buy	11.92	6.47
book sell	11.29	6.15
Average volume of one order $\times 10^6$		
market buy	9.38	5.89
market sell	8.49	5.33
spread buy	12.3	8.90
spread sell	12.1	7.73
book buy	16.3	16.7
book sell	15.6	10.8

Table 3: Summary statistics over 30 min.

The results reported are the average of the results on the 34 stocks, and the cross-section standard deviation.

3.3 Summary statistics

4 High R^2

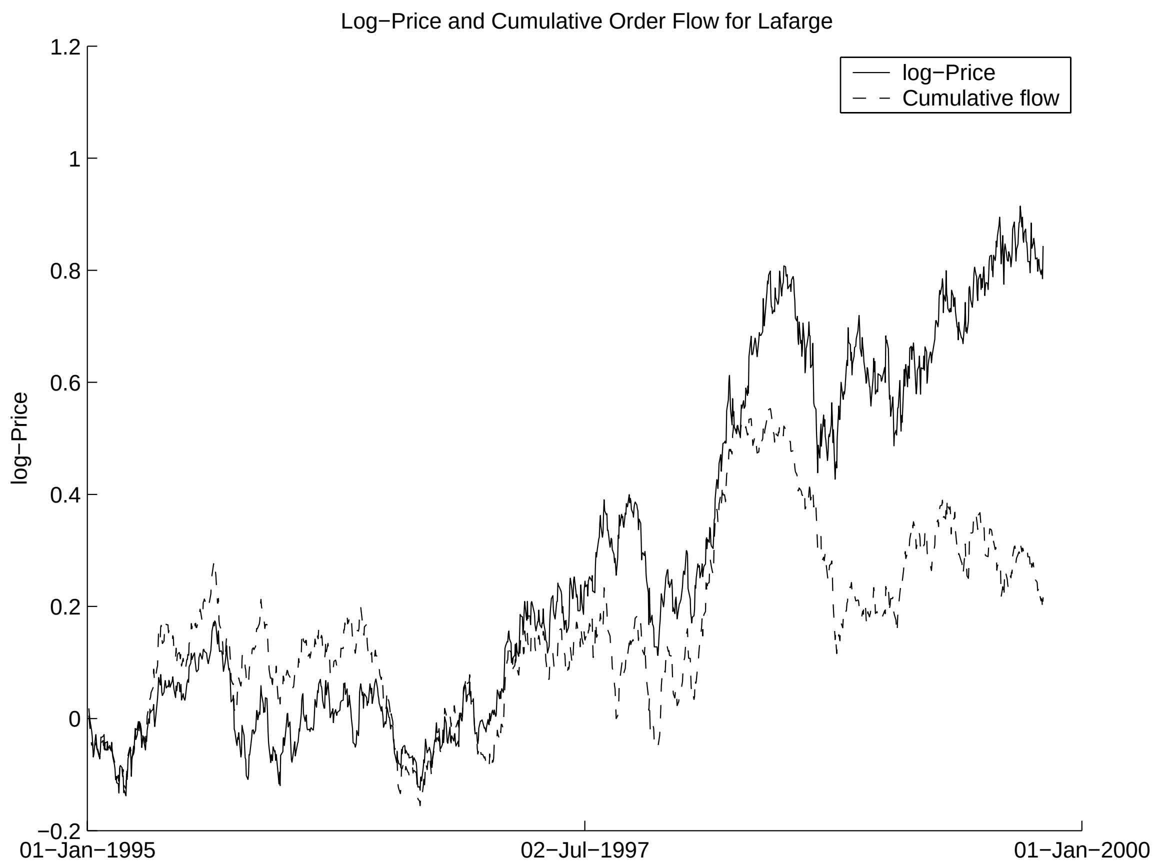

Figure 1: Cumulative return and cumulative order flow imbalance for Lafarge. The continuous line is the cumulative return of Lafarge (using daily closing prices). The dashed line is the cumulative order flow imbalance. The order flow indicator is the Sqrt of orders. To take into account the various impacts of the three different urgencies, I used the three coefficients from daily regressions when adding together the different buy and sell orders.

The first finding using order flow data is that stock return and order flow imbalance are highly correlated. I produce different results related to this finding.

[18] This is predicted by the mechanical impact, which is direct for market orders but indirect and less likely for spread and book orders. It can also be predicted by a private information setup.

[19] I discard orders too far away from the best quotes, as I have found that their impact is negligeable

[20] The log return is additive.

[21] The number of shares divided by the number of shares outstanding

4.1 A visual impression of the order flow and price

In Figure 1, the continuous line represents[22] the cumulative log return of Lafarge, using daily closing prices. It is thus the graph of (log) prices. The dashed line represents the cumulative order flow imbalance, that is, the sum of daily imbalances from date 0 to date t. The order flow indicator is the SQRT of orders. To take into account the different impacts of market, spread and book orders, I used the coefficients of a daily regression when adding the three together.

The similarity of the two lines is striking. The ups and downs of the price level are also present in the cumulative order flow imbalance. This is true not only for the daily changes, but also for longer horizon patterns, of weeks and even months.

4.2 The basic regression.

	$\lambda_{Market} \times 10^3$	$\lambda_{Spread} \times 10^3$	$\lambda_{Book} \times 10^3$	$\bar{R}^2$
estimate	82	57	19	52.4%
	95% Confidence Interval			
Lower band	77	52	17	51%
Higher band	87	62	22	54.2%
Width	9.8	10	4.8	3.2%

Table 4: The return regressed on simultaneous order flow (SQRT) for Lafarge over 30 min. I regress the 30 min log return on the simultaneous order flow imbalance, distinguishing between different urgencies, and aggregating using the square root function (SQRT= $\sum_i (v_i)^{0.5}$).

$$r_t = \alpha + \lambda_{market}\text{market}_t + \lambda_{spread}\text{spread}_t + \lambda_{book}\text{book}_t + \eta_t$$

I report the λ coefficients and the $\bar{R}^2$ corrected for the degrees of freedom. I also report the 95% confidence interval obtained from the quantiles of block bootstrap replications. The results are reported for the Lafarge stock.

In Table 4, I regress the 30 min log-return on the simultaneous order flow, distinguishing the 3 urgency levels.

$$r_t = \alpha + \lambda_{market}\text{market}_t + \lambda_{spread}\text{spread}_t + \lambda_{book}\text{book}_t + \eta_t \tag{1}$$

We find a reasonably high R^2 of 52%. I also report the block bootstraps estimates of the 95% confidence interval, using the replication quantiles. I use block bootstrapping to take into account heteroskedasticity as well as temporal[23] dependence. Indeed, we know that stocks' price volatility changes over time and that this volatility is autocorrelated. By using block bootstrap we are sure to take into account both the volatility change and its time series properties. Because the normalized regression

[22]Evans and Lyons (2000) produce a similar graph for the foreign exchange market.

[23]The block size I used is one week.

coefficients are pivotal, bootstrap also provides a second order correction for the confidence interval. This can be useful since we know that high frequency returns are non normal and fat tailed.

	$\lambda_{Market} \times 10^3$	$\lambda_{Spread} \times 10^3$	$\lambda_{Book} \times 10^3$	$\bar{R}^2$
estimate: avrg.	82	83	24	48.2%
estimate: std. dev.	25	40	11	7.7%
	95% Confidence Interval			
width: avrg.	16	20	7.7	6.1%
width: std. dev.	14	15	4.1	4.7%

Table 5: The return regressed on simultaneous order flow (SQRT) over 30 min: average results for 34 stocks.
I regress the 30 min log return on the simultaneous order flow imbalance, distinguishing between different urgencies, and aggregating using the square root function (SQRT= $\sum_i (v_i)^{0.5}$).

$$r_t = \alpha + \lambda_{market}\text{market}_t + \lambda_{spread}\text{spread}_t + \lambda_{book}\text{book}_t + \eta_t$$

I report the λ coefficients and the $\bar{R}^2$ corrected for the degrees of freedom. I also report the 95% confidence interval obtained from the quantiles of block bootstrap replications. The results reported are the average of the results on the 34 stocks, and the cross-section standard deviation.

In Table 5, I want to report the same results as in Table 4 for all the 34 stocks. For sake of brevity, I summarize the results and report the average and cross-section standard deviation of the different estimates, as well as the average and standard deviation of the width of the 95% confidence intervals. Again, we notice the high R^2 and the significance of the results.

One can find similar and more detailed results in Hopman (2002a). There, I also show that these results are true for very different time horizons, all the way from 10 min up to 3 months. Besides, it turns out that the impact of each order is best captured by a concave function of its volume, well described by the square root function. In Gabaix and Hopman (2002), we provide non parametric estimates of the impact function that confirm this type of concavity.

4.3 Interpretation of this correlation

We have seen that the log-return and the order flow imbalance are highly correlated, with R^2 around 50%. This result is important because in contrast, it has proved hard to find similar R^2 when explaining stock returns with information. The main work on this topic is from Roll (1988). His results suggest that information cannot explain more than 30% of the price changes, using daily observations. He first builds principal components and industry factors. He interprets the R^2 obtained from regressing a company's return on these factors as due to market-wide or industry-wide information. This is already a quite generous assumption in favor of the information interpretation. Indeed, it assumes that all comovements are driven by information. In fact, as observed in Hopman (2002b), the factor structure of the return is the same as that observed in the order flow. So price pressure can explain the comovements instead of information. His assumption also means that only company specific movements are allowed to be interpreted as not due to information. It is therefore all the more important to look at his result on company specific movements and company specific news. His method is to distinguish days with company news from days without. The idea is that on days without company specific news, the market wide factors should be the sole determinants and the R^2 should be close to 100%. However, the difference in R^2 between days without news and all days together is less than 2 percentage points (and both R^2 are below 30%). It seems that a lot of the idiosyncratic variance exists without any idiosyncratic news. So for the idiosyncratic part of stock returns, where no assumption is made that relates them to information, the data does not suggest that they do. It has been a puzzle since this paper that news do not seem to have a high explanatory power for stock returns. The low R^2 of information compared to order flow imbalance are a first indication that there may be some direct price pressure.

However, I propose to show more rigorously that this is indeed price pressure. First, it is interesting to understand other possible explanations, based on information, of the correlation between price changes and order flow imbalance. The first potential explanation is Public Information: traders analyze public news, and submit orders due to this news. To get the good sign for the correlation, you also need to assume that more buy orders than sell orders are placed when the news is good, and inversely when it is bad. The second potential explanation is Private Information: in a Kyle (1985) model, a market maker[24] will adjust prices to any order, whether it is informed or not, because he cannot distinguish between them. This model predicts the correct sign of correlation.

To argue that public information is not driving the results, I rely on the previous work of French and Roll (1986), who show that public information explains only a small part of stock returns. They study stocks on days when the New York Stock Exchange is closed but the rest of the economy is active, and find that their variance[25] is on average only 14.5% of what it is when the exchange is open. Since the flow of public information is at least as big during these exchange holidays (which include presidential elections), their results suggest that only 14.5% of a stock's daily variance can be attributed to public information. If public information explains so little of the price changes, it is unlikely to explain both the price changes and the order flow. The public information story does not seem to be driving the results.

Since French and Roll (1986) deal with public information[26], I propose to show

[24] Although there are no official market maker on the Paris Bourse, it is reasonable to assume that some rational "liquidity providers" play a similar role.

[25] They estimate a two day variance, from the closing price before the holiday to the closing price on the day after the holiday, so that public news have a full trading day to affect prices. The two day variance is 1.145 a normal day's variance.

[26] In fact their argument also deals with a particular type of "Private Information": faster analysis of Public Information, as long as it is analyzed during the next day. So in the next section I do not consider this kind of "Private Information".

in the next section that private information is also unlikely to be the explanation for the correlation between order flow and stock returns.

5 Private Information versus Price Pressure

As we have seen above, the most realistic information-based interpretation of the correlation between stock return and order flow imbalance is a model of private information, developed by Kyle (1985). In this model, there is a rational risk-neutral informed trader, a rational risk neutral uninformed market maker, and some noise traders. The main result for our concern is that the market maker will move the price when he receives an order flow imbalance:

$$\Delta P = \lambda \left(v_{buy} - v_{sell} \right) \tag{2}$$

In the simplest setting (single auction),

$$\lambda = 1/2 \frac{\sigma_{info}}{\sigma_{noise}}$$

where σ_{noise} measures the volume of noise trading, and σ_{info} measures the information asymmetry between the informed trader and the uninformed market-maker. This result is also very intuitive: with more noise trading, the order flow is less likely to be informed and the market maker adjusts price less. With more information asymmetry, the order flow is likely to contain more information, and he adjusts more.

If the observed correlation between order flow imbalance[27] and stock prices is explained by a private information model in the spirit of Kyle, we expect different levels of information asymmetry and noise trading to have a strong impact on the parameter λ. This is what I test in Section 5.1.

One way to have very different information assymetries is to distinguish between company-specific returns and market-wide returns. Whereas there is a lot of potential for leakage at the company level (the CEO, key employees, managers, their family and friends, inquisitive analysts or fund managers etc.), it is difficult to find much potential for leakage at the market level. It therefore seems likely that the information asymmetry between "privately informed" traders and market-makers will be higher and that the proportion of noise traders will be smaller, on each company individually than on the market as a whole. So the price impact λ of orders should be larger as well. We test this hypothesis in Section 5.1.

5.1 Similar price impact for companies and the market factor

We have seen that under a private information model in the spirit of Kyle (1985), we expect the price impact of orders to be higher for company-specific orders than

[27] In my empirical results, the imbalance is measured in order submission, and is not the volume imbalance that faces the market-maker (or liquidity-provider) as in Kyle (1985). However it is reasonable to assume that the two are closely related, as the liquidity-provider is trying to trade with investors who want to trade but find no counterpart. This relationship is also the reason why the Kyle model and private information is a potential explanation for the observed impact of order flow.

for market-wide orders. With a mechanical price pressure, we do not expect any difference.

To test this hypothesis, I first need to measure the market-wide order flow and the company-specific (idiosyncratic) order flow for the 34 stocks.

I start by aggregating and normalizing the three types of orders for each company, by regressing each stock's return on its market, spread and book order flow imbalance:

$$\begin{aligned} r_{it} &= \underbrace{\alpha_i + \lambda_{i,market}\text{market}_{it} + \lambda_{i,spread}\text{spread}_{it} + \lambda_{i,book}\text{book}_{it}} + \eta_{it} \\ r_{it} &= \qquad\qquad\qquad\qquad f_{it} \qquad\qquad\qquad\qquad + \eta_{it} \end{aligned}$$

I call this aggregate f_{it} company i's order flow imbalance. One reason for the aggregation is to simplify the rest of this section, by having only one order flow variable per stock. This regression is also necessary to normalize the λ_i coefficients to 1 for each stock. This will then allow me to make comparisons between modified λ_i^{idio}s for different stocks and with the market λ_m.

I then define the market return as the equally weighted return for the 34 stocks. Similarly, I define the market order flow[28] as the equally weighted order flow.

$$r_{mt} = \frac{1}{N}\sum_{i=1}^{N} rit$$

$$f_{mt} = \frac{1}{N}\sum_{i=1}^{N} fit$$

I then define the idiosyncratic return for stock i as the residual of stock i's return after regressing on the market return.

$$r_{it} = \theta_i + \beta_i r_{mt} + r_{it}^{idio}$$

Similarly, the idiosyncratic order flow is the residual of stock i's order flow after regressing on the market flow:

$$f_{it} = \vartheta_i + b_i f_{mt} + f_{it}^{idio}$$

I then regress the return on the order flow, for the market as a whole, and for the idiosyncratic part of each stock.

$$\begin{aligned} r_{mt} &= k_m 1 + \lambda_m f_{mt} + \xi_{mt} \qquad (3) \\ r_{it}^{idio} &= k_i 1 + \lambda_i^{idio} f_{it}^{idio} + \xi_{it} \qquad (4) \end{aligned}$$

Results for regression 3 and 4 can be found in Table 6 and 7.

I again use block bootstrapping to test the equality for each i:

$$\lambda_i^{idio} = \lambda_m \qquad (5)$$

[28]I also used other definitions of market return and market order flow. For instance, I extracted the principal component of the return, and used the resulting eigenvector for both the return and order flow. This gave very similar results, as well as the principal component from the order flow did.

	λ_m	$\bar{R}^2_m$
estimate	1.02	69.7%
	95% Confidence Interval	
Lower band	0.97	68%
Higher band	1.06	71.6%

Table 6: The market return regressed on the market order flow.
I regress the 30 min equally weighted market return on the equally weighted order flow imbalance.

$$r_{mt} = k_m 1 + \lambda_m f_{mt} + \xi_{mt}$$

I report the λ_m coefficient and the $\bar{R}^2$ corrected for the degrees of freedom. I also report the 95% confidence interval obtained from the quantiles of block bootstrap replications. The aggregation of orders is done using the square root function: SQRT= $\sum_i (v_i)^{0.5}$. The order flow of each stock is also normalized so that for each, $\lambda_i = 1$, before distinguishing idiosyncratic and market components.

	λ_i^{idio}	$\bar{R}^2_{i,idio}$
estimate	1.03	45%
	95% Confidence Interval	
Lower band	0.97	43.1%
Higher band	1.08	46.8%

Table 7: The idiosyncratic return regressed on the idiosyncratic order flow for Lafarge.
I regress the 30 min idiosyncratic return (the residual after regressing on the equally weighted market return) on the idiosyncratic order flow imbalance (the residual after regressing on the equally weighted order flow imbalance).

$$r_{it}^{idio} = k_i 1 + \lambda_i^{idio} f_{it}^{idio} + \xi_{it}$$

I report the λ_i^{idio} coefficient and the $\bar{R}^2$ corrected for the degrees of freedom. I also report the 95% confidence interval obtained from the quantiles of block bootstrap replications. The aggregation of orders is done using the square root function: SQRT= $\sum_i (v_i)^{0.5}$. The order flow is also normalized so that for each stock $\lambda_i = 1$, before distinguishing idiosyncratic and market components. The results are reported for the Lafarge stock.

	95% Confidence Interval
	$\lambda_i^{idio} - \lambda_m$
Lower band	-0.04
Higher band	0.07

Table 8: Test of the larger impact of the idiosyncratic orders for Lafarge.
Using block bootstrap, I test if $\lambda_i^{idio} = \lambda_m$. Under a private information model, we would expect $\lambda_i^{idio} > \lambda_m$. The 5% level test does not reject the equality. The aggregation of orders is done using the square root function: SQRT= $\sum_i (v_i)^{0.5}$. The order flow of each stock is also normalized so that for each, $\lambda_i = 1$, before distinguishing idiosyncratic and market components. The test is reported for Lafarge.

	$\lambda_i^{idio} - \lambda_m$			
	95% Confidence Interval, 4 highest lower band			
Lower band	-0.016	-0.019	-0.025	-0.033
Higher band	0.113	0.90	0.72	0.68
	95% Confidence Interval, 4 lowest higher band			
Lower band	-0.152	-0.152	-0.131	-0.135
Higher band	-0.022	-0.016	-0.012	-0.05

Table 9: Test of the larger impact of the idiosyncratic orders.
Using block bootstrap, I test if $\lambda_i^{idio} = \lambda_m$. Under a private information model, we would expect $\lambda_i^{idio} > \lambda_m$. The 5% level test rarely rejects the equality, and never in this direction (all the lower bands are negative). The aggregation of orders is done using the square root function: SQRT= $\sum_i (v_i)^{0.5}$. The order flow of each stock is also normalized so that for each, $\lambda_i = 1$, before distinguishing idiosyncratic and market components.

Under the private information interpretation, we would expect to find $\lambda_i^{idio} > \lambda_m$ for all i, or at least for most of them. However, the equality 5 is rarely rejected at the 5% level, as can be seen in Table 8 and 9. And it is never rejected in this direction, but in the other one. This lack of rejection is not due to lack of power. To get as much power as possible, I use intraday data and short, 30 mn interval, which yields 16878 observations. The confidence intervals for the λ estimates are quite narrow. Economically, the two estimates are very close [29].

I interpret these results as a rejection of the private information interpretation of the price impact of orders, and therefore a support for the direct price pressure hypothesis.

5.2 The difference in the R^2.

	95% Confidence Interval
	$\bar{R}^2_{i,idio} - \bar{R}^2_m$
Lower band	-27.1%
Higher band	-22.6%

Table 10: Test for the larger R^2 of the idiosyncratic orders on Lafarge.
Using block bootstrap, I test if $\bar{R}^2_{i,idio} = \bar{R}^2_m$. Under a private information model, we would expect $\bar{R}^2_{i,idio} > \bar{R}^2_m$. The 5% level test rejects the equality, but in the opposite direction. The aggregation of orders is done using the square root function: SQRT$= \sum_i (v_i)^{0.5}$. The order flow of each stock is also normalized so that for each, $\lambda_i = 1$ before distinguishing idiosyncratic and market components. The test is reported for Lafarge.

In this section, I report an additional result which seems to contradict the private information interpretation. Within the strict Kyle model, there is no public information and all information arrives through orders. If the order flow was perfectly measured, the R^2 when regressing the return on the order flow like in equation 1 would be 100%. But this is a very simplified model which we can extend to include public information. If this public information is incorporated directly into the price, the R^2 on the order flow will not be 100%. Instead the R^2 will correspond to the fraction of volatility due to private information and the rest will be due to public information.

Again, I use the distinction between company specific returns which should be driven more by private information and have a larger R^2 and market-wide returns which should be driven less by private information and have a smaller R^2. The result

[29] These results are for the SQRT method of aggregation of orders. However, since the volume seems better justified for theoretical reasons, like the Kyle model in equation 2, I repeated this whole test with the volume aggregation. In this case, the equality was strongly rejected for all stocks, but in the opposite direction, with $\lambda_m = 2.2$ and for all i, $\lambda_i^{idio} < 1$. This is even less consistent with the private information interpretation.

	$\bar{R}^2_{i,idio} - \bar{R}^2_m$	
	95% Confidence Interval	
	average	highest higher band
Lower band	-32.3%	-22.7%
Higher band	-25.1%	-16.5%

Table 11: Test of the larger R^2 of the idiosyncratic orders.
Using block bootstrap, I test if $\bar{R}^2_{i,idio} = \bar{R}^2_m$. Under a private information model, we would expect $\bar{R}^2_{i,idio} > \bar{R}^2_m$. The 5% level test always rejects the equality, but always in the opposite direction (the higher band is always negative). The aggregation of orders is done using the square root function: SQRT= $\sum_i (v_i)^{0.5}$. The order flow of each stock is also normalized so that for each, $\lambda_i = 1$, before distinguishing idiosyncratic and market components.

I find empirically, in tables 10 and 11, is exactly the opposite: for all 34 stocks, the idiosyncratic R^2 is smaller than the market R^2 and this difference is economically and statistically highly significant, using block-bootstrapping.

Again, this seems to reject private information as the explanation why the order flow and stock returns are correlated.

6 Conclusion

We have seen that a simple model of price pressure can generate a good approximation to the observed price process (a random walk in transaction time). I have then reported results which indicate that stock returns and order flow imbalance are highly correlated, which is suggestive of price pressure. But there are other potential explanations for this correlation, based on public or private information. I first suggest that the previous research of French and Roll (1986) isn't really compatible with the public information explanation. I then propose two implications of the private information interpretation.

- Company specific orders, where there is more information asymmetry and potential for leakage, should have a bigger impact on the price than market-wide orders. This is not observed empirically.

- For the same reason, the R^2 from regressing company specific return on company specific order should be larger than for the market return on market-wide orders. Empirically we observe the opposite.

I tentatively conclude that a large fraction of stock price changes is due to direct price pressure. This implies that prices, at least for the medium term (several months or maybe years), do not reflect the underlying fundamentals of the company, but the cumulative arrival of people's orders, whether these were placed rationally or not.

A simple price pressure model like the one we have proposed would also have natural implications which have been obtained empirically and are difficult to reconcile with an information model:

- If most volatility is due to the order flow imbalance and not to new information, it would explain the low explanatory power of news in Roll (1988).
- French and Roll (1986) find that the volatility is lower when the exchange is closed but the rest of the economy is active. This is well explained by a model where trading activity is the prime determinant of price volatility.
- This model also naturally explains the correlation of volatility and volume frequently reported, for example by Karpoff (1987).
- It explains the U-shaped pattern of volatility and of volume, with an exogenous trading pause at lunch time, in the spirit of Dacorogna, Müller, Nagler, Olsen, and Pictet (1993).
- It explains the periodic surge in volatility of Eurofutures, observed every 3 months, just before the expiration of the nearest Eurofuture contract, as a consequence of an increased trading activity due to roll-over trading[30] (see for example Ballocchi, Dacorogna, Hopman, Müller and Olsen (1999)).
- It is also consistent with the impact of large trades reported by Scholes (1972) and Holthausen, Leftwich and Mayers (1990).
- It can explain why the price of a stock rises when it is included in the S&P 500, as reported by Shleifer (1986): demand, especially from indexed funds or benchmarked funds, will rise significantly at that time.
- It is consistent with claims by Shiller (1990) that the daily price is more volatile than the discounted value of dividends. Indeed, noisy traders' demand will add volatility to the underlying fundamental price of a stock.
- It could explain market bubbles: Goetzmann and Massa (2000) report that most of the rise in the S&P500 can be attributed to inflows of money into indexed funds. Some analysts have also attributed the internet bubble to the activity of day traders and their interest in technology shares.
- It might also explain market crashes: I have found for example using the same dataset that there were twice as many sell orders as buy orders in October 1998, during the Long-Term Capital Management (LTCM) crisis.

It seems that many empirical findings which were explained by many different models of information can be explained by a simple supply and demand effect.

[30]This is a strategy that allows traders to always keep approximately the same time to maturity for their portfolio, by exchanging the Eurofutures they have with some that expire 3 months later, and do this every 3 months.

This work suggests several directions for future research. The principal is probably to better understand what drives the order flow: is it known and identifiable psychological biases of human behavior? Is it also changing fads, fashionable theories and rules of thumbs that investors use to make their decisions? Or is it, at least for some, a rational and quantitative analysis of the data? Then is it possible to analyze the order flow data to distinguish between these types of orders?

It would also be very interesting to better understand the potential link between supply and demand and market crashes/bubbles, with possibly behavioral explanations like unrealistic optimism and infectious panic.

Finally, I believe that disentangling the two notions of efficiency could have important effects on the way we understand finance. On the one hand, efficiency relative to arbitrage means that you cannot make high profits without taking a significant amount of risk. On the other hand, information efficiency, which can be proved in a fully rational world, implies that price changes are only driven by news on future company dividends and modifications in rational agents' discount rate. The first efficiency concept is guaranteed by profit-seeking individuals. The second one needs much stronger assumptions, but has become a paradigm of the recent financial literature, despite some challenging evidence, including the empirical results presented in this paper.

References

[1] Ané Thierry and Hélyette Geman, 2000, *Order flow, transaction clock and normality of asset returns*, Journal of Finance.

[2] Giuseppe Ballocchi, Michel M. Dacorogna, Carl Hopman, Ulrich A. Müller and Richard B. Olsen, 1999, *The intraday multivariate structure of the Eurofutures markets*, Journal of Empirical Finance, Vol. 6, Iss. 5, 479-513.

[3] Barclay, Michael J., Robert H. Litzenberger and Jerold B. Warner, 1990, *Private Information, Trading Volume, and Stock-Return Variances*, Review of Financial Studies, vol. 3 no. 2, 233-53.

[4] Biais, B., P. Hillion, and C. Spatt, 1995, *An Empirical Analysis of the Limit Order Flow in the Paris Bourse*, Journal of Finance, Vol. 50, Issue 5.

[5] Campbell, J., A. Lo, and C. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press.

[6] Chen, Zhiwu, Werner Stanzl and Masahiro Watanabe, 2001, *Price Impact Costs and the Limit of Arbitrage*, Mimeo Yale.

[7] Dacorogna M. M., U. A. Müller, R. J. Nagler, R. B. Olsen, and O. V. Pictet, 1993, *A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market*, Journal of International Money and Finance, Vol. 12 No. 4, 413-38.

[8] Engle Robert F. and Joe Lange, 2000, *Measuring, Forecasting and Explaining Time Varying Liquidity in the Stock Market.* NBER Working Paper No. W6129. Forthcoming, Journal of Finance.

[9] Evans, Martin D.D. and Richard K. Lyons (2000), *Order Flow and Exchange Rate Dynamics*, working paper Berkeley.

[10] Fama, Eugene and Kenneth French, 1988 (a), *Dividend yields and expected stock returns*, Journal of Financial Economics, 22, 3-25.

[11] Fama, Eugene and Kenneth French, 1993, *Common risk factors in the returns on stocks and bonds*, Journal of Financial Economics, 33, 3-56.

[12] French, Kenneth, and Richard Roll, 1986, *Stock Return Variances: The Arrival of Information and the Reaction of Traders*, Journal of Financial Economics, 17, Elsevier Science Publishers.

[13] Glosten, L. and P. Milgrom, 1985, *Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders*, Journal of Financial Economics, 13, 71-100.

[14] Goetzmann, William N., and Massimo Massa ,2000, *Index Funds and Stock Market Growth*, NBER Working paper No. W7033.

[15] Granger, C., W.J., and O. Morgenstern: *Predictability of Stock Market Prices*, Lexington, Mass.: D.C. Heath and Co., 1970.

[16] Hamon, J., and B. Jacquillat, 1992, *Le marché français des actions: Études empiriques*, Presses Universitaires de France.

[17] Harris, Lawrence and Eitan Gurel, 1986, *Price and volume effects associated with the S&P 500 List: New evidence for the existence of price pressures*, Journal of Finance 41, 815-829.

[18] Harris, Milton and Artur Raviv, 1993, *Differences of Opinion Make a Horse Race*, Review of Financial Studies, vol. 6 no. 3, 473-506.

[19] Hasbrouck, Joel, 1991, *Measuring the Information Content of Stock Trades*, Journal of Finance, vol. XLVI, no. 1, 179-207.

[20] Hausman, Jerry, Andrew W. Lo and A. Craig MacKinlay 1992, *An Ordered Probit Analysis of Transaction Stock Prices*, Journal of Financial Economics, 31, 319-379.

[21] Holthausen, Robert W., Richard W. Leftwich and David Mayers, 1990, *Large-block transactions, the speed of response, and temporary and permanent stock-price effects*, Journal of Financial Economics, vol. 26, 71-95.

[22] Huang, R. and H. Stoll, 1991, *Descartes before the Horse*, Working paper, Vanderbilt University.

[23] Jones, C., G. Kaul and M. Lipson, 1994, *Transactions, volume, and volatility*, Review of Financial Studies, vol. 7,no. 4.

[24] Karpoff, J.M., 1987, *The relation between Price Changes and Trading Volume: A Survey.* Journal of Financial and Quantitative Analysis 22, 109-126.

[25] Kavajecz, Kenneth A., 1999, *A specialist's quoted depth and the limit order book*, Journal of Finance 52, 747-771.

[26] Kyle, A. S., 1985, *Continuous Auctions and Insider Trading*, Econometrica 53, 1315-1336.

[27] Lee, C.M.C. and M.J. Ready, 1991, *Inferring trade direction from intraday data*, Journal of Finance 46, 733-746.

[28] Merton, Robert, 1973, *An Intertemporal Capital Asset Pricing Model*, Econometrica 41, 867-888.

[29] Poterba, James M. and Lawrence H. Summers, 1988, *Mean reversion in stock prices ; Evidence and Implications* Journal of Financial Economics, 22, 27-60.

[30] Piccinato Barbara, Giuseppe Ballocchi, Michel M. Dacorogna, Ramazan Gencay, 1999, *Intraday Statistical Properties of Eurofutures*, Derivatives Quarterly, vol. 6, No. 2, 28-44.

[31] Roll, Richard, 1988, R^2, Journal of Finance, vol. XLIII, no. 2, 541-67.

[32] Scholes, Myron S., 1972, *The market for securities: Substitution versus price pressure and the effects of information on share price*, Journal of Business 45, 179-211.

[33] Shleifer, Andrei, 1986, *Do demand curves for stocks slope down?*, Journal of Finance 41, 579-590.

[34] Shiller Robert J., 1990, *Market Volatility*, MIT Press.

[35] Stoll H., 1978, *The Supply of Dealer Services in Securities Markets*, Journal of Finance 33, 1133-1151.

[36] Spiegel, Matthew and Avanidhar Subrahmanyam, 1992, *Informed Speculation and Hedging in a Noncompetitive Securities Market*, Review of Financial Studies, 4, 709-726.