Price Discovery in the Cross Section: Leaders and Followers

Diego Amaya, M. Fabricio Perez, and Andriy Shkilko*

Wilfrid Laurier University

December, 2019

Abstract

A frictionless market should allow information to be incorporated into all assets quickly and efficiently. Market frictions however make such incorporation less than perfect. What is the extent of the issue, and what are the main frictions? To answer this question, we use a novel econometric methodology that allows us to analyse the propagation of public information across all the constituents of the S&P 500 index. We show that a sizeable portion of information is captured by only a few equities, suggesting that not all assets incorporate information in a timely manner. Our analysis suggests that stocks that are slow to adjust are those with significant liquidity demand pressures and high trading costs.

Keywords: Price adjustment, Price discovery, Systematic risk

JEL classification: G14, G23.

*We thank Canadian Derivatives Institute (CDI) for financial support.

1 Introduction

An optimally functioning financial market should allow asset managers and other market participants to quickly adjust their positions in response to the arrival of new information. In theory, such changes should happen swiftly so that all stock prices reflect public information in a timely manner. Prior research shows that the adjustments indeed occur efficiently over medium and long periods of time (e.g., Bai, Philippon, and Savov, 2016). But how quicky do prices adjust within the day? Do adjustment speeds vary in the cross-section, and if so what are the determinants of this variation? We suggest that the speeds may indeed differ across assets due to a variety of frictions faced by market participants. For instance, if the cost associated with opening and closing of a position is high, prices may adjust slowly or not at all. A similar situation could arise when liquidity demand pressures are high, causing excessive volatility. In such cases, market participants may wait until the pressures subside, otherwise facing a more volatile and uncertain environment. As such, there is a possibility that stock prices, even in today's fast markets, do not incorporate information immediately.

Understanding dislocations of common information in stock prices is important for at least three reasons. First, it provides a new cross-sectional view of how prices impound public information, departing from previous studies that seek to understand how price discovery is done within a group of securities that share a common efficient price. For instance, Harris (1989) and Chan (1992) describe the links between stock and stock index futures markets, whereas Bhattacharya (1987) and Easley, O'Hara, and Srinivas (1993) study the links between stock and option markets. Second, the ubiquity of exchange traded fund (ETF) trading has introduced large intraday demand pressures on the underlying stocks due to arbitrage activity (Ben-David, Franzzoni, and Moussawi, 2018). This new type of intraday liquidity shocks increases stock volatility and may therefore affect the efficiency of information incorporation into prices. Third, it allows us to appreciate relative gains in price efficiency, shedding light on the way markets incorporate common information across stocks (Fama, 1970).

In this paper, we investigate if the stocks exposed to common price schocks incorporate them into prices in a timely manner. Using the largest and most liquid stocks in the US market, those underlying the S&P 500 index, we document sizeable differences in reactions to new information in the cross section.

To study how market-wide information is impounded into prices, we propose a new price discovery measure that identifies the stocks that lead information incorporation in the cross section. This measure can be viewed as an extension of Hasbruck (1995) to a case in which information are the innovations in the market portfolio, rather than a common efficient stock price across multiple venues. Our interest with this shift in paradigm is to measure price discovery across stocks about a common component. We implement this measure using one-second data for all stocks that are part of the S&P 500 index during 2010, given that this group of stocks enjoys stable characteristics that allow for comparison across periods.

The information share measure builds on the common factor approach of De Jong and Schotman (2010) and Westerlund, Reese, and Narayan (2017). This methodology conveniently formulates the problem of price discovery as one in which the efficient price component is seen as a common factor, so that existing methods can be employed in their estimation. In our case, the common factor is the market portfolio return, which we estimate at a given point of time from the cross-sectional average of all stocks in the intersection of CRSP and DTAQ. The basic idea of the information share measure is to identify the proportion of information in a given stock relative to other stocks that are also exposed to common innovations in the market portfolio. Thus, this measure provides the proportion of information in a given stock relative to the total contribution of stocks in the panel.

Our empirical analyses generate four sets of main findings. First, when we look at the daily information share across stocks in the S&P 500 index, we find that a small group contains a disproportional amount of information about common innovations in the market portfolio. For instance, when we sort stocks according to their daily information share, we find that the top quintile contains about 70% of the total information in the common innovation. This amount is in striking contrasts with the total information contained in the lowest information share quintile, which contains on average less than 1%. These proportions are consistent when we estimate the daily information share from lower intraday frequencies, showing that this result is not a mechanical effect from stale quotes or non-synchronous trading (Campbell, Lo, and MacKinlay, 1997), but rather support the view that price discovery about the market portfolio is carried out by a small group of stocks.

Second, we show that there is an notable level of persistence among constituents of information share groups, particularly among the stocks that belong to the lowest information share quintile in a given day. We find that these stocks are in the same quintile the next day 45% of the time, much more often than the 29% observed for stocks in the highest quintile migrating into the same quintile over the same period. Morevoer, the lower persistence in the highest quintile suggests that there is nontrivial rotation among stocks incorporating common information, but there is still a fraction of them consistently integrating information day after day. When we compare characteristics of stocks in these two information groups, we find important differences: stocks in the highest (lowest) quintile have higher (lower) market capitalizations, lower (higher) transaction costs, and higher (lower) systematic risk.

Third, we use panel regressions to study the determinants of information share ratios. When we look at specifications that only account for stock characteristics unrelated to liquidity, we find that market capitalization and the proportion of systematic risk in the stock's return (R^2) are positively related with the information share. However, when these variables are used in conjunction with others related to liquidity, i.e., trading activity, demand pressure, and liquidity of the stock, we find that only the latter group of variables is significant. In particular, demand pressure plays a deterrent role in the proportion of information impounded in the price. We view this as evidence that the stock's ability to impound common information comes primary from microstructure factors associated with the stock's trading environment. These factors speak directly to the signal-to-noise ratio of the stock, as observed prices become more imprecise about fundamental quantities when these microstructure components grow in importance.

Fourth, we use Dimson (1979) regressions to compare the speed of adjustment of portfolio values to market returns for stocks sorted into groups according to their information share. The portfolio containing stocks with the largest information share are less sensitive to lagged information in the market return. In contrast, portfolio returns of stocks with the lowest information share are more sensitive.

The rest of the paper is organized as follows. Section 2 presents a simple model to understand price adjustments and the measure of information share. Section 3 presents the data and descriptive statistics about this variable. Section 4 empirically assess the determinants of information shares and Section 6 extends analysis to lower frequencies. Section 6 examines that speed of adjustment of portfolios based on the proposed measure. Section 7 concludes.

2 Model and Methodology

2.1 Price adjustment to Market Factor Changes

We start by describing a simple model in which stock prices incorporate common information from the market factor and lagged adjustments in prices. The objective of this model is to lay the groundwork for subsequent analyses about the incorporation of market factor innovations by a large group of stock prices.

Let $p_{i,t}$ denote the efficient log-price of stock i at time t, which we assume that evolves according to the following equation:

$$p_{i,t} = p_{i,t-1} + \alpha_i \left(\gamma_i F_{M,t} - p_{i,t-1} \right). \tag{1}$$

In this equation, $F_{M,t}$ represents the market factor (log-value) observed at time t. The parameter γ_i captures the sensitivity of the stock price to contemporaneous changes on market factor. The variable α_i measures the speed of price adjustment to permanent shocks in the market. All together, these three components govern the way stock prices incorporate common information arriving from the market factor.

The model presented in Equation (1) extends Hasbrouck and Lo (1987) to one in which common information across stocks and lagged price adjustments are incorporated into prices. To illustrate the interaction between these two sources, notice first how stocks with $\alpha_i = 1$ immediately adjust to the arrival of the common innovation $F_{M,t}$. In contrast, stocks with $\alpha_i < 1$ do not incorporate new common information in a timely manner, as lagged adjustments are also impounded in current price changes. Moreover, if a stock has an $\alpha = 0$, no common information is incorporated and only idiosyncratic adjustments are reflected in current prices. The dependence of stock price changes on the adjustment speed to new information shows that for a stock to be informative about innovations in the market factor, it is necessary to not only have high systematic risk but also to promptly incorporate new information.

We assume that the dynamics of the market factor are governed by a random walk:

$$F_{M,t} = F_{M,t-1} + \eta_t, \tag{2}$$

in which η_t represents a zero-mean disturbance that is independently and identically distributed (i.i.d) with variance σ^2_η. This representation of the market factor assumes that all new information, conditional on time t, comes from the innovation η_t. This specific assumption implies that all stocks are closely-linked by this term, but it is not necessarily true that they impound this information at the same time, as it is clear from the stock price representation in Equation (1).

We link the latent efficient stock price to the observed one by assuming that these two differ by an error component that could be assoicated with microstructure effects such as transaction costs, information asymmetry, inventory costs, unpredictable shifts in demand, etc. Formally, we write the observed stock price by

$$\begin{aligned} \pi_{i,t} &= p_{i,t} + \epsilon_{i,t} \\ &= \lambda_i \left(\eta_t + F_{M,t-1}\right) + \left(1 - \alpha_i\right) p_{i,t-1} + \epsilon_{i,t}, \end{aligned} \tag{3}$$

where $\epsilon_{i,t}$ denotes an i.i.d. idiosyncratic shock of variance $\sigma^2_{\epsilon_i}$ and $\lambda_i = \alpha_i \gamma_i$ gives the loading of the observed return on the common innovation.

2.2 Stock's Information Share of the Market Factor

The above representation of observed prices is particularly useful because it illustrates how price changes could be governed by a common factor innovation η_t and an idiosyncratic component $\epsilon_{i,t}$. This structure suggests that using a cross-section of stocks can be useful to disentangle common price movements from idiosyncratic ones. The immediate result from this identification process is that cross-sectional measurements about the amount of common information absorbed by a given stock should help identify stocks that are relatively more efficient than others in the process of discovering new information about the market factor.

To formalize this idea, let $v_{i,t}$ denote the total sum of common and idiosyncratic contemporaneous innovations in the observed price, that is

$$v_{i,t} = \lambda_i \eta_t + \epsilon_{i,t}. \tag{4}$$

Let $v_t = (v_{1,t}, \ldots, v_{N,t})^\top$, $\Lambda = (\lambda_1, \ldots, \lambda_N)^\top$, and $E = (\epsilon_{1,t}, \ldots, \epsilon_{N,t})$, then we can write the total variation in observed prices as

$$\Sigma_v = \sigma_\eta^2 \Lambda \Lambda^\top + \Sigma_E, \tag{5}$$

where $\Sigma_E = diag\left(\sigma^2_{\epsilon_{1,t}}, \ldots, \sigma^2_{\epsilon_{N,t}}\right)$.

If common innovations η_t were observable, then it is possible to identify stocks that contain a larger share of information about η_t by looking at the proportion of the R^2 explained by each stock return in the following regression:

$$\eta_t = b^\top v_t + e_t, \tag{6}$$

where $b = (b_1, \ldots, b_N)^\top$ and e_t is the part of the innovation in the efficient price that is not due to shocks in observed prices. Given that the OLS estimate of vector b is $\Sigma_v^{-1}\Lambda\sigma_\eta^2$, the R^2 measure of the regression in Equation (6) is[1]

$$
\begin{aligned}
R^2 &= \frac{b^\top \Sigma_v b}{\sigma_\eta^2} \\
&= \sum_{i=1}^{N} b_i \lambda_i .
\end{aligned}
\tag{7}
$$

From the equation above, it is clear that each stock contribution to R^2 is given by $b_i\lambda_i$, which provides a measure of the stock's information share about common innovations. Following Westerlund, Reese, and Narayan (2017), the latter expression can be written in terms of parameters governing Equation (4), which leads to define the stock information share, IS, as:

$$
IS_i = b_i \lambda_i = \frac{\left(\lambda_i \frac{\sigma_\eta}{\sigma_{\epsilon_i}}\right)^2}{\sum_{j=1}^{N} \left(\lambda_j \frac{\sigma_\eta}{\sigma_{\epsilon_j}}\right)^2} \tag{8}
$$

The above equation states that the contribution of a given stock return to the total variation in the common innovation depends on two sources. The first one is the loading on the common innovation, λ_i^2, which captures the degree of association between stock i and the market factor. The second source is the ratio between the common innovation's variability and the one from the stock's idiosyncratic component. This second term measures the precision with which contemporaneous innovations are informative about market factor innovations, since low(high) relative values of σ_{ϵ_i} magnify(reduce) the contribution of λ_i.

To gain further intuition about this measure, we can use the model introduced in Equa-

[1] The derivation of the following result can be found in Appendix A of Westerlund, Reese, and Narayan (2017)

tion (3) and identify how model parameters influence the stock's information share. In particular, notice that the loading on the common innovation is given by $\lambda_i = \alpha_i \gamma_i$, so the degree of association between stock returns and the common factor depends on the speed at which the stock incorporates information and the systematic risk of the company. In addition, the amount of information is also a function of the microstructure noise present in the observed stock price, since prices that diverge more from the efficient component would have lower information shares.

2.3 Estimation of Stock's Information Share

We now address the question of how to estimate the information share of a given stock, as measured by Equation (8). Assume that a panel of log-prices is available for N companies, each of which having a number of T observations at a given observation frequency.

Since these log-prices follow the representation in Equation (3), the observed log-return for a given stock can be written as:

$$\begin{aligned} r_{i,t} &= \Delta \pi_{i,t} \\ &= \lambda_i \eta_t + (1-\alpha_i) r_{i,t-1} + (1-\alpha_i)\Delta\epsilon_{i,t-1} + \Delta\epsilon_{i,t}. \end{aligned} \tag{9}$$

This equation shows that log-returns have contemporaneous and lagged information. If we compare Equation (4) with Equation (9), we notice that the former equation only has contemporaneous innovations in returns. Thus, we need to first isolate these components from the observed return. This requirement can be achieved by working with the residual $\tilde{r}_{i,t}$ obtained from an ARMA regression on stock returns.[2] In this manner, $\tilde{r}$ only captures

[2] In our empirical implementation, we found that the AR(1) specification provided indistinguishable results from an ARMA(1,1), so we privileged the use of the former model throughout our analyses.

the contemporaneous shocks of Equation (9).

Westerlund, Reese, and Narayan (2017) estimate the common factor from the the cross-sectional average of stock returns in the panel, based on the estimator of Pesaran (2006). We use a similar estimator for our market factor based on equal- and value-weighted averages of all companies in the intersection of CRSP and DTAQ, as discussed in detail in the next section. Once this factor is obtained, the estimator of IS_i proposed in Westerlund, Reese, and Narayan (2017) consists of sample estimates of each of the individual parameters required in the computation of this quantity. Under suitable assumptions, with large N and T, the authors show that this procedure provides a consistent estimate of IS_i as defined in Equation (8).

3 Data Sources and Sample Overview

3.1 Intraday Stock Return Sample

We focus on the set of stocks that compose the S&P500 index for the entire year of 2010, since these stocks are some of the most liquid securities in the market and together constitute one of the most important benchmarks in financial markets. For our study, cross-sectional diversities in this sample provide an appropriate setting to study the differential speed with which stocks react to the arrival of common information.

We obtain the S&P 500 index constituents from the Center of Research in Security Prices (CRSP) and keep those that are always part of the index throughout 2010 and for which there is daily data in the DTAQ database. This selection criterion leaves us with 479 stocks.

For the list of eligible companies, we take the NBBO file in DTAQ and join at a millisecond level with quotes from NASDAQ that are not in this file. The resulting sample is used to construct one-second log-returns from the last recorded NBBO in the observation interval. To isolate opening or closing effects, observations between 9:35 a.m and 3:55 p.m. are the only ones consider. We remove from our sample days in which the market was not operating full day and flash-crash days (May 5, 6 and 7).

To construct the market factor at the intraday level, we first select all companies in CRSP daily file that meet five criteria: (1) the price per share must be higher than \$5; (2) market capitalization is higher than \$100 million; (3) the stock must be listed in NYSE, AMEX, or NASDAQ; (4) it must be a common stock; (5) the ticker is DTAQ master file. We use NBBO mid-quotes for each stock satisfing the previous criteria and compute an equally- and value-weighted factor by cross-sectionally averaging log-returns at a one-second frequency.

3.2 Information Share of S&P 500 Stocks

Daily estimates of the Information Share are computed for each company following the methodology presented in Section 2.3. These values are then averaged across the sample period. Table 1 presents statistics of this variable for the cross-section of stocks in the S&P 500 index. We observe that the average stock has an information share of 0.21%, which corresponds to a value close to 1/500. This value represents a situation where information about the market factor is equally impounded in the sample of stocks studied. From the skewness and percentiles in the table, it is clear that the unconditional distribution of IS is rather symmetric, showing that there are stocks with lower (larger) information shares than what would otherwise the typical level for a stock. Panel B in the table shows that this conclusion is also obtained when the market factor is computed using value weights.

We now look at the relative information share of stocks across the sample period. To this end, every day we sort stocks into 5 groups according to their IS for a given day. The group

Table 1: **Information Share Characteristics.**
This table presents summary statistics about the information share that companies in the S&P 500 have with respect to the market portfolio. Panel A shows summary statistics across stocks using as market portfolio an equally-weighted average of common stocks in the intersection of CRSP and DTAQ. The variable *IS* represents the average information share (in percentage) as defined in Equation (8) for a given stock across the sample period. The variable *IS Group* refers to the average quintile in which a stock is classified according to its daily IS (the lowest information group is 0 and the highest one is 4). The variable *# Days - Lowest IS* provides the number of days that a given stock belongs to the lowest IS quintile. The variable *# Days - Highest IS* provides the number of days in the highest IS quintile. Panel B shows results based on a value-weighted market factor. The sample period is 2010 with 248 trading days.

Panel A: Equal-Weighted Market Index

	Nobs	Median	Mean	StdDev	Skew	P10	P90
IS	479	0.20	0.21	0.11	0.56	0.07	0.36
IS Group	479	2.15	2.00	0.65	-0.94	1.06	2.71
# Days - Lowest IS	479	34.00	48.99	47.14	2.08	10	110
# Days - Highest IS	479	49.00	49.25	27.17	0.20	12	84

Panel B: Value-Weighted Market Index

	Nobs	Median	Mean	StdDev	Skew	P10	P90
IS	479	0.20	0.21	0.12	0.92	0.06	0.37
IS Group	479	2.13	2.00	0.68	-0.86	1.02	2.74
# Days - Lowest IS	479	32.00	48.99	48.83	2.05	9	111
# Days - Highest IS	479	49.00	49.25	28.79	0.34	11	88

with label 0 represents stocks in the lowest information group, while the label 4 is given to the group with the highest information share. We present in Table 1 statistics about this variable as well as statistics about the number of days that stocks stay in the most (least) informed groups. Although stocks spend a similar number of days in both groups, we observe a larger proportion of stocks in the lowest information group, as evidenced by the positive skewness and larger standard deviation.

To appreciate the level of information across groups, for each day in the sample, we sum within each group the information share attributed to each stock. Table 2 shows the median value for the total information share across the five groups. We observe a large dispersion in the amount of total information share present in each group. Whereas the lowest group captures less than 1% of the total variation in the market factor for a given day, the group with the largest IS captures about 70%. From Figure 1, it is clear that these differences are persistent across time, illustrating that a small group of stocks carries most of the information about the market factor during a given day. Nonetheless, the fact that the average stock has an IS of .21% shows that it is not necessary true that the same stocks are being attributed with the largest source of variation in the market factor component.

Table 2 also reports characteristics of stocks in each of the information groups. Each day, we compute variables such as market capitalization, CAPM-beta obtained from regressing the previous 250 daily returns on CRSP's daily market factor, and the volatility as measured by the standard deviation of the previous 250 daily returns. In addition, we report measures of trading activity on the stock such as total number of trades and dollar volume, and measures of stock's trading costs like the average quoted and effective percentage spread. These last four variables are obtained from WRDS. From the table, the group with the highest information share are large, have high systematic risk, experience a relative low level

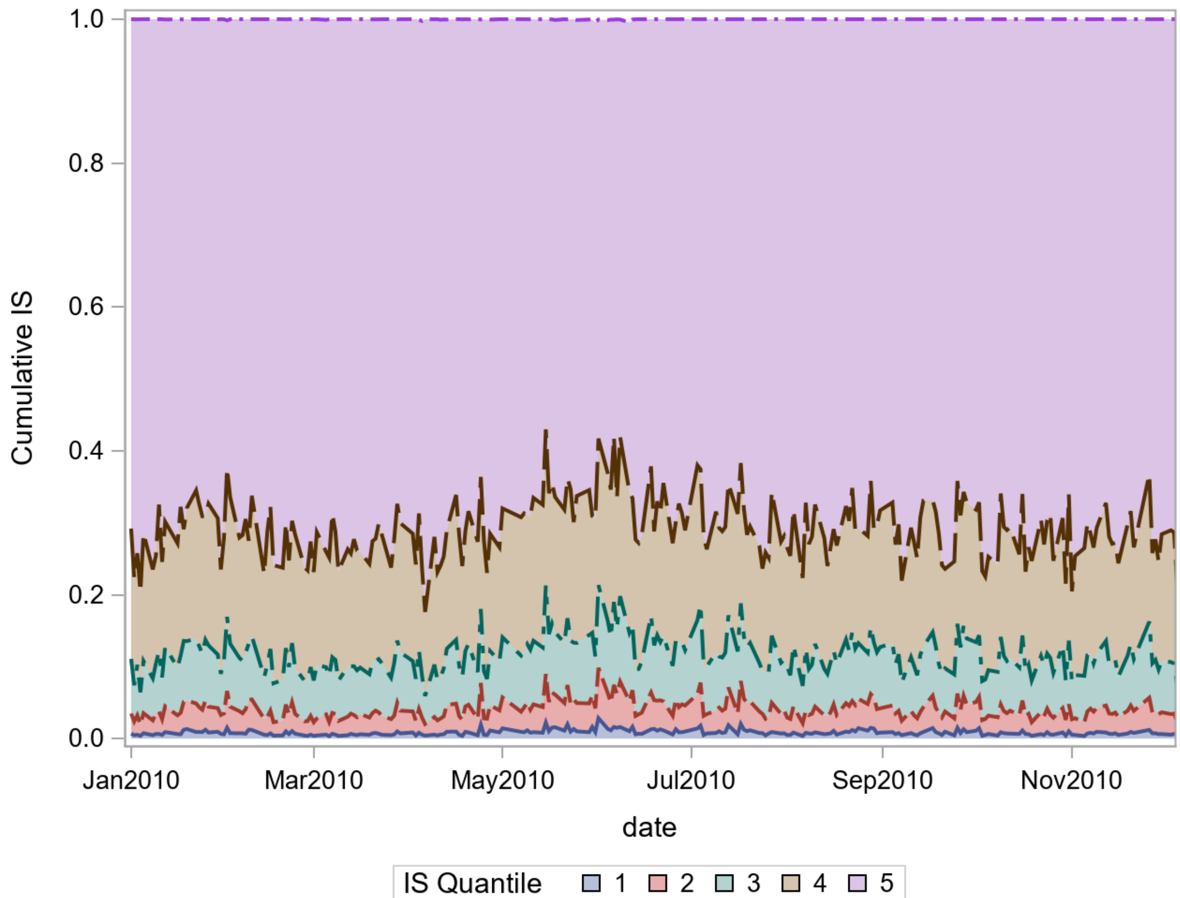

Figure 1: **Total Information Share by Quintiles.** The figure shows the total information share of IS quintile across the sample. For each day, the proportion of information in each quintile is cumulatively added and depicted with a different color. The IS quintile 1 represents the group of stocks with the lowest information share (bottom colored region in the figure), while the IS quintile 5 represents the one with the highest information share (top colored region in the figure). The sample period is 2010.

of trading activity and trading costs.

3.3 Information Persistence

Next, we examine the persistence of stocks' information share by looking at the empirical distribution of stock transitions between IS quintiles across days. Table 3 presents 5×5 transition matrices that indicate subsequent information share levels (columns) as a function of initial ones (rows). Panel A (B, D) displays 1-day (3-day, 5-day) transition probabilities from overlapping time periods.

For a 1-day period, we observe that 45% of stocks in the lowest information share quintile remain in this quintile one day after (row one, column 1 in Panel A). This number slightly decreases to 42.2% and 42.3% when looking at the same matrix entry in the 3- and 5-day transition probabilities, suggesting a high degree of persistence. Stocks with low information share are likely to exhibit low information shares ex post at 1-,3-, and 5-day horizons. This persistence contrasts with that observed in the quintile of stocks with the highest information share. When we look at the last row and column of transition probabilities, we observe lower percentage of stocks staying in this group. For instance, at a 1-day level, a stock in the highest information share quintile remains in this category 29.8% of the time. We also observe from the transition probabilities that firms with less (more) information shares are less likely to transition into a group with a higher (lower) information share. This asymmetric result shows that a large group of stocks are not likely to incorporate information about the market factor for a number of consecutive days, illustrating that the dynamic process of incorporating information about this common factor is relegated to few stocks.

Table 2: **Information Share Group Characteristics.**
This table presents characteristics for stock groups formed by the level of information share. Every day, we sort stocks into five groups with cut-off points determined by quintiles of the IS distribution. Within each group, we compute the total sum of IS -in percentage- (*Total IS*) and obtain median values of the stock's market value (*Size*), CAPM beta (*Beta*), volatility (*Total IS*), total number of trades (*Number Trades*), total volume (*Volume*), quoted percentage spread (*Quoted Spread*), and effective percentage spread (*Effective Spread*). Variables *Size* and *Volume* are measured in millions; spreads are presented in basis points. This table presents median values of each variable across the 248 trading days of 2010. Panel A presents results for the case an equal-weighted market index is used in the construction of IS. Panel B presents results when the value-weighted market index is employed.

Panel A: Equal-Weighted Market Index

	Lowest	Quintile 2	Quintile 3	Quintile 4	Highest	All
Total IS	0.73	3.00	7.36	17.48	70.37	7.36
Size	7.91	9.08	9.99	10.67	11.86	9.69
Beta	0.94	1.02	1.07	1.10	1.15	1.06
Volatility	0.32	0.31	0.31	0.31	0.31	0.31
Number Trades	19,198	16,037	15,664	15,467	16,487	16,503
Volume	4.43	3.18	2.97	2.90	2.96	3.19
Quoted Spread	5.54	4.21	3.95	3.83	3.65	4.04
Effective Spread	5.20	3.75	3.52	3.44	3.31	3.68

Panel B: Value-Weighted Market Index

	Lowest	Quintile 2	Quintile 3	Quintile 4	Highest	All
Total IS	0.66	2.78	6.98	16.84	71.39	6.98
Size	7.49	8.92	9.97	10.96	12.81	9.78
Beta	0.95	1.02	1.06	1.10	1.14	1.06
Volatility	0.32	0.31	0.31	0.31	0.31	0.31
Number Trades	18,874	15,621	15,686	15,640	16,966	16,549
Volume	4.37	3.11	2.97	2.92	3.05	3.19
Quoted Spread	5.74	4.24	3.95	3.80	3.57	4.04
Effective Spread	5.36	3.77	3.56	3.44	3.28	3.67

Table 3: **Transition Probabilities.**
This table reports transition probabilities of IS quintiles across days. Every day, we group observations into quintiles based on their information share. The result is a 5×5 matrix for which we compute the empirical distribution of transitions between a given day and a specific subsequent period. In Panels A, B, and C, we examine 1-, 3-, and 5-day transition probabilities, respectively. The sample period is 248 trading days in 2010.

	Lowest	Quintile 2	Quintile 3	Quintile 4	Highest
Panel A: 1-Day Transition Probabilities					
Lowest Information Share	0.450	0.205	0.151	0.118	0.076
Quintile 2	0.200	0.225	0.210	0.197	0.168
Quintile 3	0.147	0.209	0.216	0.216	0.213
Quintile 4	0.117	0.197	0.215	0.229	0.242
Highest Information Share	0.081	0.167	0.210	0.244	0.298
Panel B: 3-Day Transition Probabilities					
Lowest Information Share	0.422	0.209	0.155	0.128	0.086
Quintile 2	0.206	0.223	0.205	0.191	0.175
Quintile 3	0.155	0.204	0.214	0.213	0.213
Quintile 4	0.123	0.193	0.215	0.228	0.240
Highest Information Share	0.087	0.175	0.213	0.242	0.283
Panel C: 5-Day Transition Probabilities					
Lowest Information Share	0.423	0.204	0.153	0.128	0.091
Quintile 2	0.206	0.222	0.203	0.193	0.176
Quintile 3	0.150	0.207	0.214	0.215	0.215
Quintile 4	0.124	0.195	0.218	0.225	0.238
Highest Information Share	0.091	0.175	0.214	0.243	0.277

4 Determinants of Information Shares

In this section, we empirically assess which stock characteristics relate with the level of information about the market factor contained in stock prices. To select these variables, we look at the two main components of IS in Equation (8) and highlight which characteristics could play an important role in this variable empirically.

We start with the term λ_i. This term captures the stock's systematic risk and the speed at which it incorporates this information. From this perspective and given the parsimony of the model, the first set of variables to be tested correspond to parameters of the CAPM model. The first one we include is beta, which provides a standardized measure of co-movement between the stock and the market portfolio. The second one is the return's standard deviation (volatility), capturing the total risk embedded in the price coming from systematic and idiosyncratic shocks. The third one is the R^2 resulting from the regression. This variable helps to disentangle the proportion of systematic risk present in the total variance. We compute these variables for each day in the sample by using the previous 250 daily excess returns of a given stock and regressing them over the daily market factor obtained from CRSP. Given cross-sectional differences in company's size and the marked interest in the largest companies of S&P 500 index, we also include size in this set of variables.

The second component of IS is the microstructure noise present in observed stock prices, σ_{ϵ_i}. The first variable we include is the daily dollar volume of the stock (in logarithm), which captures trading activity in the stock. The second one is the demand pressure exerted on the stock, measured by the absolute difference between buyer and seller initiated trades divided by the total number of trades. The third variable corresponds to the ratio between the total number of intermarket sweep orders employed in the day divided by the total number of trades. This variable is meant to capture information from trading activity by agents looking for faster execution. In addition, we add the price impact, which measures the permanent component of trades. Regarding liquidity costs, we include the daily percent quoted spread. These two last variables are compouted following Holden and Jacobsen (2014). All of these variables are obtained from WRDS.

Next, to analyse these effects, we estimate the equation

$$IS_{i,t} = a + \mu_i + \gamma' X_{i,t} + bIRet_t + e_{i,t}, \tag{10}$$

where IS is the information share for stock i in day t, μ_i denotes a firm fixed effect, and X corresponds to a vector with stock-level characteristics. The model also includes the percentage return on the S&P 500 index, $IRet$, to control for specific daily variability (time effects) associated with the index. We estimate the model using ordinary least squares, clustering observations by stock when calculating standard errors.

Table 4 shows results for three specifications of the above model. The first specification, Column (1), contains characteristics related to the stock's systematic component. In this case, we observe a positive and significant relation of stock's information share with size and R^2, and no significant relation with the other three variables. To assess the economic significance of these two coefficients, we multiply the variable's interquartile difference by its coefficient in the regression (see Table A in the Appendix for statistics about regression variables). This result implies that an increase in size (i.e., a change between 16.79 and 15.52) is associated with a 0.08% increase in the stock's information share. Recall from Table 1 that the standard deviation of IS is 0.11%, meaning that an increase of 0.08% represents roughly three quarters of a typical change in this variable. Regarding R^2, an increase in this value (i.e., a change between 0.57 and 0.33) implies a rise of 0.04% on IS, about half of what

Table 4: **Stock's Characteristics and Information Share.**

This table reports results from panel regressions in which the dependent variable is the information share of a stock over day t. The variable (*Size*) represents the total market capitalization computed from closing prices (in logarithm). To compute the variables α_{CAPM} and β_{CAPM}, for a given day and stock, we estimate the equation $R^e = a + \beta_{CAPM} R^{M,e} + u_t$ using the previous 250 daily observations of stock and market excess returns. The variable R^2 represents coefficient of adjustment in this regression. The variable $Volatility$ corresponds to the standard deviation of previous 250 daily returns. The variable *Volume* is the dollar volume of the stock (in logarithm). The variable *Quoted Spread* corresponds to the average of quoted percentage spread computed from all NBBO in a given day. The variable *Price Impact* is the average percentage price impacts across the day. The variable *Absolute OIB* is the absolute difference between the number of buyer minus seller initiated trades, divided by the sum of the two. The variable *ISO Orders* is the total number of intermarket seep orders divided by the total number of trades for the stock. The variable *Index Return* is the close-to-close daily return on the S&P 500 index. All coefficient estimates, expect for *Quoted Spread* and *Price Impact*, are multiplied by 100. Standard errors are computed with robust errors clustered by stock. There are a total of 119645 observations in the panel. The sample period is 2010. Statistical significance at the 1% and 5% levels is indicated by ** and *, respectively.

	(1)		(2)		(3)	
	Parameter	T-Stat	Parameter	T-Stat	Parameter	T-Stat
Constant	-0.903**	-3.09	1.654**	21.10	1.891**	4.95
Size	0.066**	3.70			-0.013	-0.63
β	-0.026	-1.31			-0.004	-0.22
R^2	0.159**	4.55			-0.025	-0.70
Volatility	0.080	0.15			0.880	1.50
Volume			-0.105**	-21.37	-0.107**	-21.12
Quoted Spread			-0.882**	-2.83	-1.000**	-2.86
Price Impact			0.469**	5.80	0.452**	5.58
Absolute OIB			-0.116**	-4.60	-0.115**	-4.63
ISO orders			0.393**	13.36	0.411**	13.85
Index Return	-0.079	-0.727	-0.291**	-2.60	-0.286*	-2.57
Firm Fixed Effects	Yes		Yes		Yes	
R2	0.11%		1.88%		1.90%	

is observed for the size variable.

Column (2) in Table 4 examines the role of variables related to the microstructure noise. The results show that all these variables are significantly related with the stock's information share. When we look at the economic significance of these variables, the largest impact is observed for volume, where a change in this variable leads to a decrease of 0.15% in the stock's IS: a change that corresponds to about 1.36 times the size of the IS standard deviation. The proportion of ISO has the second largest impact, with an increase in this variable translating into a 0.05% increase in IS. Regarding the quoted spread and the absolute order imbalance, an increase in one of these variables leads to a decrease of 0.02% and 0.01%, respectively. Finally, the lowest impact is observed for the price impact, with an increment of 0.01% when this variable is increased.

We next consider in Column (3) of Table 4 a specification including all variables. In contrast with the first specification, none of the variables related with the stock's systematic characteristics is significant. On the contrary, all variables in the second specification continue to be significant and exhibit similar coefficient values. These results indicate that the information share of stock prices at a one-second frequency is principally associated with trading activity and liquidity. The negative association with volume and absolute OIB shows that high trading activity leads to a decrease in the level of information about the market portfolio. According to Equation (3), this could be explained if the trading activity captured by these variables is associated with the arrival of idiosyncratic shocks that make observed prices less informative about the common innovation. This could also explain why a high proportion of ISOs is positively associated with the stock's information share, given that the use of this type of order is associated with traders taking advantage of short-lived information (Chakravarty et al., 2012). If this information is related with innovations in the market portfolio, then this positive relation is expected. Regarding liquidity variables, we observe a negative relation with quoted spreads, confirming the intuition that information is impounded in stocks with lower trading costs. The positive relation with the price impact highlights the permanent component of the price captured by this variable rather than the liquidity dimension of this measure. To the extent that this permanent component is related to common new information, stocks with high price impact exhibit a higher information about the common innovation in the market portfolio.

One aspect underlying the previous results is that these relations could depend on the frequency at which the Information Share is measured. For instance, at a one-second interval, if stock prices are influenced by short-lived information about the market portfolio, only those capturing this type of information would obtain larger shares of information. However, once this information has been incorporated into the price, the information share could depend on other determinants. To shed light on the role that the observation frequency plays on these determinants, we estimate the information share over a 10-minute frequency. To further contrast these results, we use a rolling window of 250 daily return and compute the information share from these data. Note that these daily returns contain intraday and overnight information.

Table 5 reports the results including all covariates used when running the third specification in Table 4. To facilitate comparisons, the first column in the table contains results for the one-second frequency. The results in this table show that the stock's information share loads on differents covariates depending on the frequency over which it is measured. For the 10 minute frequency, we find that variables such as quoted spread and price impact are no longer statistically significant, while others like volume and ISOs are still significant but with lower economic values. When we look at the daily frequency, we observe a different picture

Table 5: **Stock's Characteristics and Information Share for Different Frequencies.** This table presents the results of panel regressions in which the dependent variable is the information share of a stock computed at one of three frequencies: one second, ten minutes, and one day. The variable ($Size$) represents the total market capitalization computed from closing prices (in logarithm). To compute the variables α_{CAPM} and β_{CAPM}, for a given day and stock, we estimate the equation $R^e = a + \beta_{CAPM} R^{M,e} + u_t$ using the previous 250 daily observations of stock and market excess returns. The variable R^2 represents coefficient of adjustment in this regression. The variable $Volatility$ corresponds to the standard deviation of previous 250 daily returns. The variable *Volume* is the dollar volume of the stock (in logarithm). The variable *Quoted Spread* corresponds to the average of quoted percentage spread computed from all NBBO in a given day. The variable *Price Impact* is the average percentage price impacts across the day. The variable *Absolute OIB* is the absolute difference between the number of buyer minus seller initiated trades, divided by the sum of the two. The variable *ISO Orders* is the total number of intermarket seep orders divided by the total number of trades for the stock. The variable *Index Return* is the close-to-close daily return on the S&P 500 index. All coefficient estimates, expect for *Quoted Spread* and *Price Impact*, are multiplied by 100. Standard errors are computed with robust errors clustered by stock. There are a total of 119645 observations in the panel. The sample period is 2010. Statistical significance at the 1% and 5% levels is indicated by ** and *, respectively.

	1 sec		10 min		1 day	
	Parameter	T-Stat	Parameter	T-Stat	Parameter	T-Stat
Constant	1.891**	4.95	1.225**	5.73	3.646**	5.76
Size	-0.013	-0.63	0.005	0.41	-0.225**	-6.07
β_{CAPM}	-0.004	-0.22	0.001	0.06	0.063	1.18
Systematic Risk	-0.025	-0.70	0.003	0.11	0.31**	3.91
Volatility	0.880	1.50	0.556	1.33	-3.951**	-3.05
Volume	-0.107**	-21.12	-0.001**	-27.31	0.011	1.85
Quoted Spread	-1**	-2.86	0.076	0.45	-0.364	-0.95
Price Impact	0.452**	5.58	-0.036	-0.54	0.009	0.11
Absolute OIB	-0.115**	-4.63	-0.133**	-6.67	-0.006	-0.30
ISO orders	0.411**	13.85	0.118**	6.77	-0.11*	-2.40
Index Return	-0.286*	-2.566	-0.231**	-2.63	0.036	0.49
Fixed Firm Effects	Yes		Yes		Yes	
R^2	1.90%		1.11%		2.15%	

from the one obtained with high frequency data. From the list of variables associated with the microstructure component of the information share, only the ISO variable comes significant in the regression. Regarding variables associated with the systematic component, all but β_{CAPM} are statistically significant. These changes in statistical significativity, combined with the fact that these variables show different signs, shows that intraday information shares similar characteristics across different frequency levels, but differs from the one contained in daily returns.

5 Information Share across Frequencies

Given the previous evidence that information captured at a one-second frequency relates to some extent with that one observed at a lower intraday frequency, we study in this section the allocation of common information across stocks as a function of the sampling frequency. This characterization provides evidence about the speed with which information is incorporated in stock prices as innovations are aggregated over time. Specifically, it allows us to observe how the disproportional amount of information captured by stocks in the largest IS quintile (see Figure 1) is incorporated at lower frequencies.

We start by estimating the stock's information share for different intraday frequencies. For each frequency, we sort stocks into groups according to IS quintiles and compute the total amount of information allocated to each group. The results are reported in Panel A of Table 6. We find that the group of stocks with the highest information share always account for the largest share of information across groups. The highest proportion is observed at a one-second frequency, with about 70% of the total information impounded in the top quintile. At the 30-second frequency, this value has monotonically decreased to 66.6%, continuing to

Table 6: **Information Share for Different Sampling Frequencies.**
This table presents average values of information shares fo quintile groups formed at different frequencies. Every day and for a given frequency, we sort stocks into five groups according to their information share. Within each group, we compute the total sum of IS -in percentage- and compute the average value across the sample period for each group. In Panel A, we present results based on eight different intraday frequencies over which the value of IS is computed. In Panel B, the information share is computed based on daily returns. In this case, we use a rolling window of the previous 250 days to estimate the information share of a stock for a given day. The sample period is 248 trading days in 2010.

Panel A: Intraday Sampling Frequencies

	Lowest	Quintale 2	Quintale 3	Quintale 4	Highest
1 sec	0.83	3.22	7.52	17.51	70.92
3 sec	0.98	3.51	7.93	18.00	69.55
5 sec	1.07	3.66	8.11	18.18	68.96
10 sec	1.20	3.86	8.38	18.50	68.04
30 sec	1.42	4.19	8.78	18.93	66.65
1 min	1.53	4.35	9.02	19.15	65.91
5 min	1.57	4.54	9.24	19.32	65.27
10 min	1.39	4.50	9.48	19.63	64.94

Panel B: Daily Sampling Frequency

	Lowest	Quintale 2	Quintale 3	Quintale 4	Highest
Daily	2.45	5.76	10.52	19.71	61.17

slowly decrease to a value of 64.9% at the 10-minute frequency. This pattern is different from the one observed in other quintiles, since their total information share increases as the frequency decreases. Panel B in Table 6 shows the total information share per quintile obtained with daily data. At this frequency we still observe that almost two thirds of the total information is contained in the top quintile of stocks. This result suggests that different frequencies contain distinct information, but some of the information arriving at a higher frequency is slowly impounded into lower ones.

To shed further light into the amount of information flowing between frequencies, we look at how stocks with the highest information share move across estimation frequencies. Specifically, we look at the proportion of stocks in the highest information quintile at the one-second frequency and record the groups in which they are classified at lower observation frequencies. Figure 2 shows the flows across three frequencies. From the figure, transitions from the one-second frequency into the one-minute show that the largest proportion of stocks (54.17%) is observed for stocks migrating into the highest IS quintile (red flow in the figure). The total information share contained in these stocks is 46.62%–the average IS for a stock in this group is .85% at the one-second frequency.[3] At a one-minute frequency, the average IS for these stocks is .76%, representing 41.7% of the total information share at this frequency. Given that the amount of information in the highest quintile of the one-minute frequency is 65.91% (see Table 6), we conclude that almost two thirds of this value comes from stocks with the highest information share in the one-second frequency.

Regarding the percentage of stocks that belong to the highest quintile at the ten-minute frequency, we find that about 31% of stocks migrated into the largest IS quintile: 20.15% came from the highest IS quintile of the 1-min frequency and the remaining from the other

[3]Since the number of stocks in the quintile is about 100, the total information share contained by this group of stocks is $54.15 \times .85\% = 46.62\%$

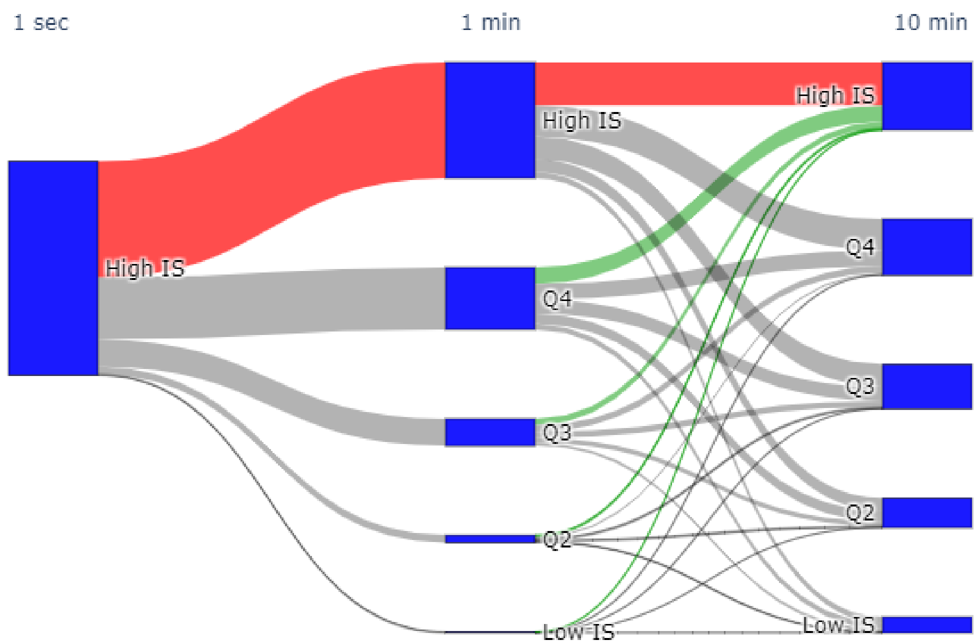

Figure 2: **Flow of Stocks from the Highest Information Quintile at the one-second frequency.** The figure plots quintile migrations of stocks in the highest information quintile at the one-second frequency. The nodes in blue represent the quintile in which stocks are grouped at a given frequency. The frequency is provided on top of the figure. A red link represents the proportion of stocks that belong to the highest quintile and transition to the the highest quintile when this group is formed at a lower frequency. A green link represents a proportion of stocks that transition into the highest quintile at the 10-minute frequency from the one-minute frequency (not including the ones transitioning from the highest quintile). Gray links represent all other transitions.

4 groups (red and green flows, respectively). At a ten-minute frequency, the average IS for these stocks is .71%, so the total information share they convey is 22%–one third of the total value in the highest quintile at this frequency (see Table 6).

In summary, we find that stocks with the highest information share at the one-second frequency contain important proportions of information shares at lower frequencies given that some of these stocks remain in the highest quintiles at lower frequencies. We do not pin down which part of the information in the highest frequency is shared with lower ones, but we show that an important part of the information at lower frequencies comes from stocks that impound it at higher frequencies.

6 Speed of Information Share Adjustment

In this section, we compare the speed of adjustment of IS quintile porfolio values to market returns. Specifically, we use Dimson (1979) regressions to determine the fraction of the information in the market portfolio returns that is present in a given quintile.

We denote by $R^n_{j,t}$ the equal-average return of stocks in the nth quintile on day j for second t. Similarly, $R^M_{j,t}$ represents the equal-average market portfolio return. For each day in our sample, we run the following regression:

$$R^n_{j,t} = a + \sum_{k=-60}^{60} b_{j,k} R^M_{j,t-k} + u_{j,t}, \quad n \in \{1, \ldots, 5\}. \tag{11}$$

This specification combines lagged ($k > 0$) and lead ($k < 0$) values of the market factor to capture the total sensitivity of the quintile portfolio's return to the portfolio market. This sensitivity is computed for day j as the sum of coefficients b_k, $\beta^{All}_j = \sum_{k=-60}^{60} b_{j,k}$.

Panel A in Table 7 reports summary statistics for the total sensitivity of each quintile portfolio. We observe a monotonic increase from quintiles with the lowest information share to the highest, which is expected given the nature of the sorting variable (the stock's information share). The quintile with the lowest information share has a total sensitivity of 0.81, much lower than the 1.13 observed for the highest quintile.

In addition to the total sensitivity, we also examine the leading, lagging, and contemporaneous sensitivity for each quintile portfolio. We define the lagging sensitivity as the sum of coefficients b_k of lagged variables, $\beta_j^{Lag} = \sum_{k=1}^{60} b_{j,k}$. The leading sensitivity is defined as $\beta_j^{Lead} = \sum_{k=-60}^{-1} b_{j,k}$ and the contemporaneous one corresponds to the coefficient b_0 in the regression. Panel B in Table 7 shows average values for each quintile. We observe that the contemporaneous beta is the sensitivity with the largest component in β^{All}. For all quintile portfolios, we find that leading sensitivities are small but positive, suggesting a modest predictability component in these portfolios. Regarding the lagging sensitivity, we find that it is negative for all quintiles except for the one with the lowest information share. This suggest that this portfolio is the one with the highest sensitivity to lagged information in the market portfolio, that is, the one adjusting more slowly to common information. In contrast, the portfolio with the highest information share has the lowest negative sensitivity to lagged information, showing that it is the fastest to impound information. These interpretations go in line with Brennan, Jegadesh, and Swaminathan (1993) who find differential of speed adjustment to common information in portfolios of firms that are followed by different number of analists, and with Chordia and Swaminathan (2002) who also document these differences for portfolios based on volume.

To visually contrast the speed of adjustment to market information of quintile portfolios, we compute the cumulative fraction of reaction to market information realized at a given second ($\sum_{k=-60}^{d} b_{j,k}/\beta_j^{All}$). This quantity provides the fraction of information impounded at

Table 7: **Dimson Betas.**
This table presents results for Dimson betas, which are the coefficients obtained from regressing one-second quintile portfolio returns on lead, contemporaneous, and lagged returns of market returns for a given day. This regression is given in Equation (11). Panel A shows summary statistics of the total sensitivity of quintile returns to market returns, as defined by $\beta^{All} = \sum_{k=-60}^{60} b_{j,k}$. Panel B presents average values of leading sensitivity $\beta^{Lead} = \sum_{k=-60}^{-1} b_{j,k}$, the lagging sensitivity $\beta^{Lag} = \sum_{k=1}^{60} b_{j,k}$, and the contemporaneous sensitivity $\beta^{Ctmp} = b_{j,0}$. The sample period is 2010.

Panel A: Summary Statistics of β^{All}

	Nobs	Median	Mean	StdDev	Skew	P10	P90
Lowest Information Share	248	0.82	0.81	0.14	-0.66	0.62	0.97
Quintile 2	248	0.98	0.98	0.16	-0.49	0.78	1.19
Quintile 3	248	1.05	1.04	0.16	-0.75	0.83	1.23
Quintile 4	248	1.09	1.08	0.17	-0.56	0.88	1.28
Highest Inforamtion Share	248	1.13	1.14	0.18	-0.26	0.92	1.37

Panel B: Average Values of Different Betas

	β^{All}	β^{Lead}	β^{Ctmp}	β^{Lag}
Lowest Information Share	0.81	0.03	0.75	0.03
Quintile 2	0.98	0.05	1.19	-0.26
Quintile 3	1.04	0.06	1.36	-0.39
Quintile 4	1.08	0.07	1.49	-0.48
Highest Inforamtion Share	1.14	0.08	1.68	-0.62

a given point in time before ($d > 0$) or after ($d < 0$) the portfolio return is observed. Figure 3 plots the cumulative fractions for the largest and lowest information quintiles. We observe that both portfolios contain no significant information that anticipates the market ($d < -1$) and that there is some level of predictability one-second ahead of the realization of the market return. The figure shows how the largest increase in information comes when the market portfolio return is observed ($d = 0$). Once this information has arrived, cumulative fractions follow different patters as lagged information for several seconds continues to be impounded in the lowest information share portfolio. On the contrary, betas for lagged market returns are negative for the portfolio with the largest information share, so the cumulative fraction of reaction decreases with the lag.

7 Conclusion

This paper provides a measure of information share for a panel of stocks that have as common component the market portfolio. We show that the information share of stocks largely differs across stocks, pointing to inefficiencies about the way common information is impounded in a cross-section of stocks. When we look at possible determinants of this inefficiencies, our results show that stocks that experience high demand pressures are more likely to have lower shares of information about the market portfolio. Examination of adjustment speeds to information shows that stocks with the lowest share are sensitive to lagged information in the market portfolio, different from what is found for the portfolio containing the stocks with higher information share levels.

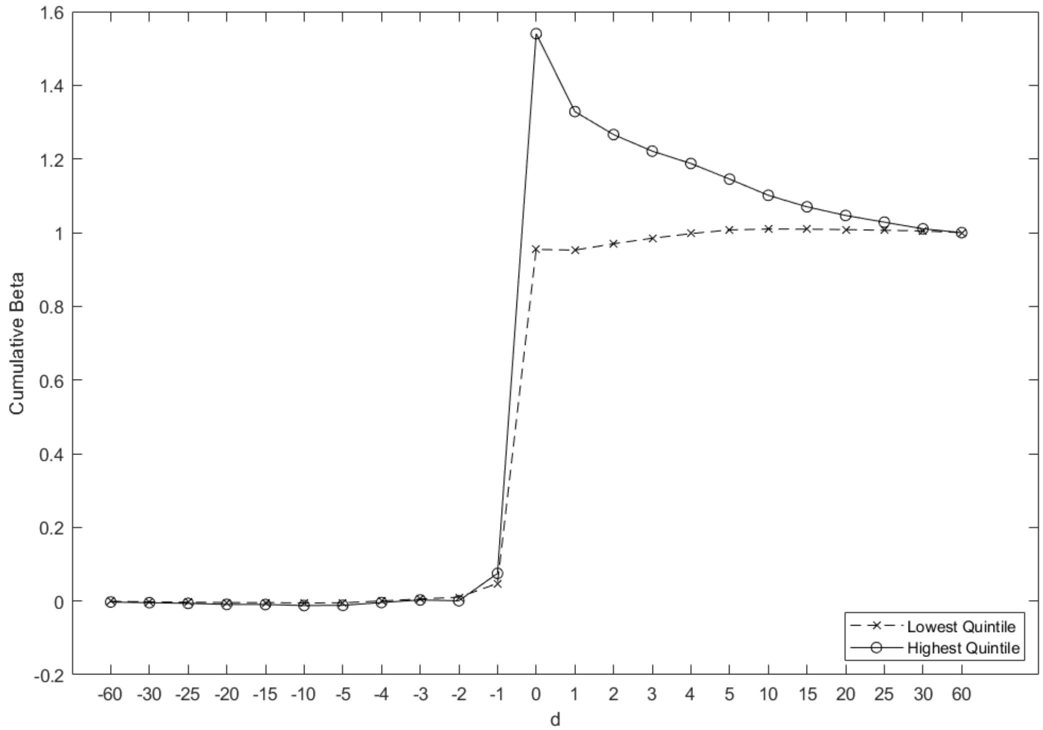

Figure 3: **Cumulative sums of coefficients in Dimson regressions.** This figure plots the average cumulative sum of coefficients in a Dimson regression for the lowest (dotted line) and the highest (continous line) quintile portfolio returns. For each day in the sample, the one-second quintile portfolio return is regressed over several lead and lag terms of the market portfolio return. From the estimated coefficents in a given day, the cumulative beta is computed for each d in $\{-60, \ldots, 60\}$ as $\sum_{k=-60}^{d} b_{j,k}/\beta_j^{All}$, where the denominator represents the sum over all lead and lag components (including the contemporaneous term). The cumulative beta presented in the figure corresponds to the average of this value over the sample period (2010).

References

Bai, J., Philippon, T. and Savov, A., 2016. Have financial markets become more informative?. Journal of Financial Economics, 122(3), pp.625-654.

Ben-David, I., Franzoni, F. and Moussawi, R., 2018. Do ETFs increase volatility?. The Journal of Finance, 73(6), pp.2471-2535.

Bhattacharya, M., 1987. Price changes of related securities: The case of call options and stocks. Journal of Financial and Quantitative Analysis, 22(1), pp.1-15.

Brennan, M.J., Jegadeesh, N. and Swaminathan, B., 1993. Investment analysis and the adjustment of stock prices to common information. The Review of Financial Studies, 6(4), pp.799-824.

Campbell, J.Y., Lo, A.W., and MacKinlay, A.C., 1997. The econometrics of financial markets. princeton University press.

Chakravarty, S., Jain, P., Upson, J. and Wood, R., 2012. Clean sweep: Informed trading through intermarket sweep orders. Journal of Financial and Quantitative Analysis, 47(2), pp.415-435.

Chan, K., 1992. A further analysis of the lead–lag relationship between the cash market and stock index futures market. The Review of Financial Studies, 5(1), pp.123-152.

Chordia, T. and Swaminathan, B., 2000. Trading volume and cross-autocorrelations in stock returns. The Journal of Finance, 55(2), pp.913-935.

De Jong, F. and Schotman, P.C., 2009. Price discovery in fragmented markets. Journal of Financial Econometrics, 8(1), pp.1-28.

Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. Journal of Financial Economics, 7(2), pp.197-226.

Easley, D., O'hara, M. and Srinivas, P.S., 1998. Option volume and stock prices: Evidence on where informed traders trade. The Journal of Finance, 53(2), pp.431-465.

Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. The journal of Finance, 25(2), pp.383-417.

Harris, L., 1989. The October 1987 S&P 500 stock-futures basis. The Journal of Finance, 44(1), pp.77-99.

Hasbrouck, J. and Ho, T.S., 1987. Order arrival, quote behavior, and the return-generating process. The Journal of Finance, 42(4), pp.1035-1048.

Hasbrouck, J., 1995. One security, many markets: Determining the contributions to price discovery. The journal of Finance, 50(4), pp.1175-1199.

Holden, C.W. and Jacobsen, S., 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. The Journal of Finance, 69(4), pp.1747-1785.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor errorstructure. Econometrica, 74, pp.967–1012.

Westerlund, J., Reese, S. and Narayan, P., 2017. A factor analytical approach to price discovery. Oxford Bulletin of Economics and Statistics, 79(3), pp.366-394.

A Appendix

Table 8: **Descriptive Statistics for Variables in Panel Regressions.**

This table presents descriptive statistics for independent variables used in panel regressions. The sample period is 247 trading days in 2010.

	Mean	Std	Min	P25	P50	P75	Max
Size	16.23	1.00	14.43	15.52	16.08	16.79	18.99
β_{CAPM}	1.12	0.48	0.31	0.75	1.05	1.42	2.64
R^2	0.44	0.16	0.00	0.33	0.46	0.57	0.82
Volatility	0.02	0.01	0.01	0.02	0.02	0.03	0.10
Volume	14.97	1.06	12.64	14.23	14.92	15.63	17.94
Quoted Spread	0.05	0.03	0.02	0.03	0.04	0.06	0.20
Price Impact	0.02	0.02	-0.02	0.01	0.01	0.02	0.10
Absolute OIB	0.06	0.05	0.00	0.02	0.05	0.08	0.24
ISO orders	0.43	0.08	0.24	0.37	0.43	0.49	0.63
Index Return	0.05	1.11	-3.24	-0.42	0.08	0.60	3.13