The Impact of Information Shocks on the Dispersion of Market Betas

Jose Antunes-Neto *

Kellogg School of Management

December 5, 2023

Abstract

Keywords: News, Public Information, CAPM Beta, High Frequency

* jose.neto@kellogg.northwestern.edu

1 Introduction

The concept of beta has been a cornerstone of modern finance theory since its introduction by Sharpe (1964). Beta measures the sensitivity of a stock's returns to changes in the market as a whole, and is often used as a measure of a stock's riskiness. This measure has initially been described as a stable characteristic of a stock, representing the long term sensibility of its return relative to the market's. Later studies show that beta exhibits strong variation over time, challenging the traditional view of beta as a constant measure of risk.

More recently, Andersen, Thyrsgaard, and Todorov (2021) evidence the intraday pattern of the betas. As showed by the authors, the cross-sectional distribution of betas presents a consistent daily pattern in which betas tend to be more dispersed in the early hours of the trading day and more concentrated towards unity as the market approaches close. This work addresses that puzzle and studies whether this can be explained by the type of information shocks the market receives throughout the day. The main approach is based on the work of Campbell and Vuolteenaho (2004), which describes that betas can be decomposed into two components, one that relates to the cash flow risk of the stock and another which relates to discount rate risk. In this paper, I argue that the cash flow component of the betas is the main driver of the dispersion pattern we see in intraday.

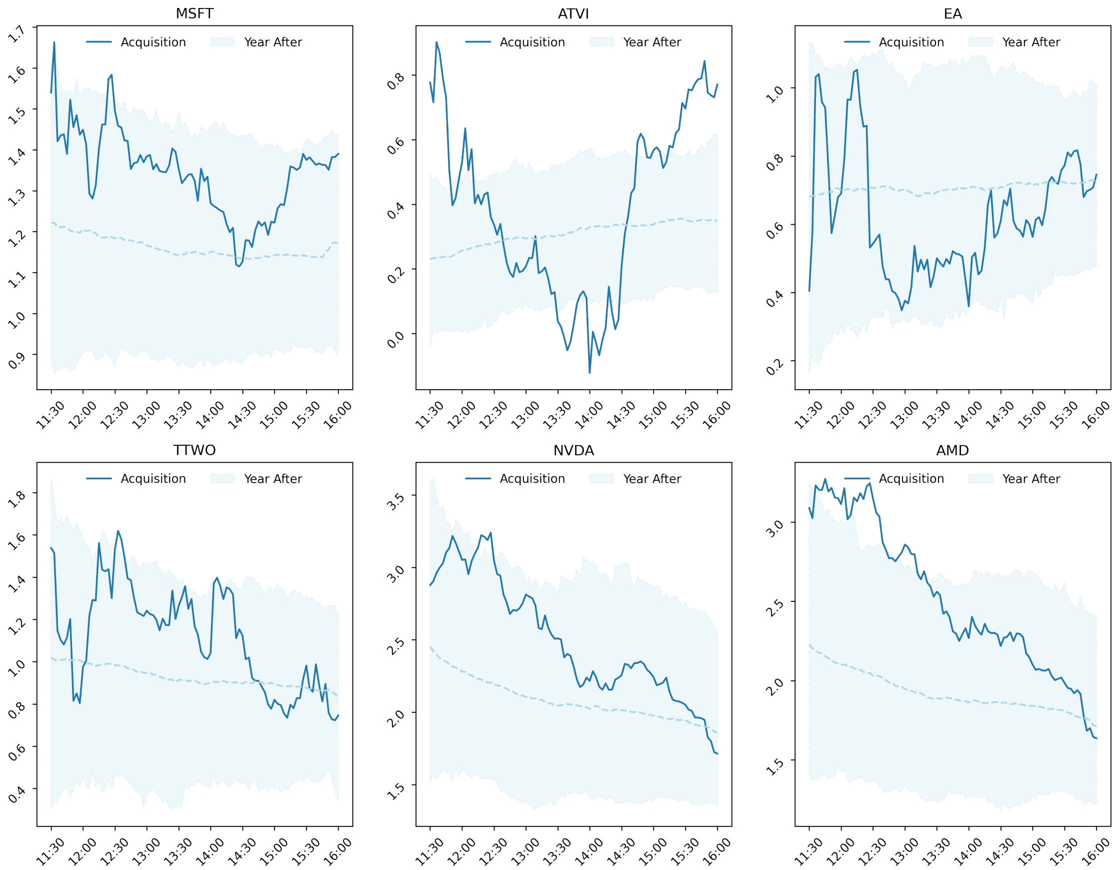
As cash flow shocks tend to be more heterogeneously spread throughout the market, they impact the risk of many stocks in different sizes and directions. Therefore these type of shocks would likely lead to a higher increase in beta dispersion if compared to discount rate shocks, which tend to be more homogeneous and affect the market more uniformly. For example consider the recent events of Microsoft (MSFT) acquiring Activision Blizzard (ATVI)[1]. The acquisition event likely displays as a positive cashflow shock for both MSFT and ATVI, therefore changing their risk perception by the market and, finally, their betas. This acquisition, however, can also be seen to affect Microsoft's competitors in the gaming industry as the merge seems to be a threat for their market share. Finally, other stakeholders in the technology industry might as well be affected due to their linkage with Microsoft. In the end, this effect would crowd out to the market, generating a more risky environment and increasing the cross-company variance of betas.

To see this, we take a look at Figure 1. This figures compare the estimates of betas close to the annoucement of the acquisition of Activision Blizzard by Microsoft[2] with their 2022 average distribution. The solid line shows the average beta for a given point in time for the day of the announcement and the next day. The shaded area provides the 95% quantile range for the betas on the year of 2022. The stocks selected were Microsoft (MSFT), Activision (ATVI), Electronic Arts (EA), Take-Two Interactive (TTW), NVIDIA (NVDA) and AMD (AMD). As we see, the betas for these stocks often hit their highest levels of the year around the announcement. The downward trend is clear for these stocks, representing the learning of the market about the event and its impact in the firms future cash flows.

[1] https://www.forbes.com/sites/qai/2023/10/16/microsofts-69-billion-activision-blizzard-acquisition-finally-approve
?sh=5c759e28554b

[2] https://news.microsoft.com/2022/01/18/microsoft-to-acquire-activision-blizzard-to-bring-the-joy-and-community-of-

Figure 1: Betas around the acquisition announcement day (Jan 18, 2022)

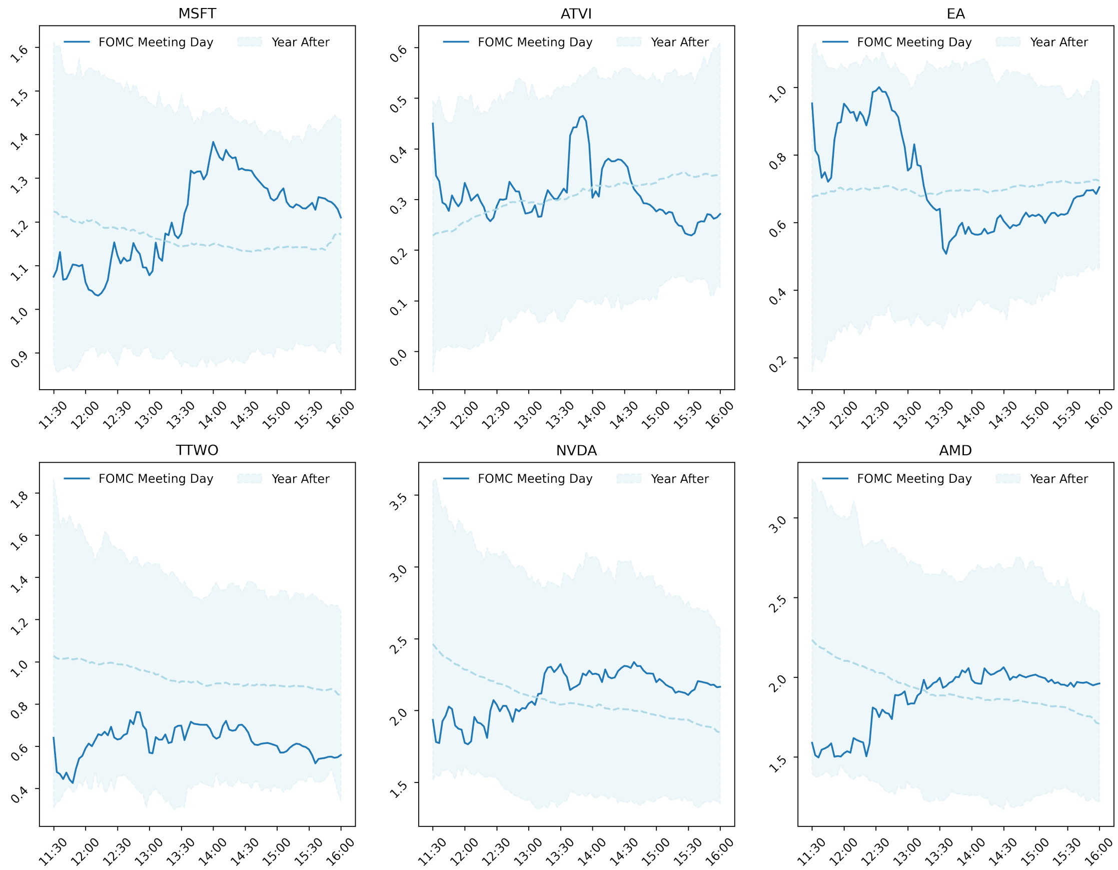


Iterestingly, seven days later, the FOMC met and announced[3] its intention to start hiking interest rates starting on their next event. This was the start of its tightening cycle[4] after the COVID-19 crisis and considered a big event for the market. What type of effect did it had for the same companies we displays in the last paragraph? Figure 2 shows the same exercise, but with betas calculated around the FOMC announcement days (January 26 and 27). The same pattern we showed on Figure 1 is not here anymore. The betas for these stocks are mostly constant throughout the day and do not show any significant changes when compared to their year average. This contributes to the idea that cash flow shocks, unlike discount rate, produce the dispersion pattern we see in the intraday betas.

[3] https://www.cnbc.com/2022/01/26/fed-decision-january-2022-.html

[4] https://www.forbes.com/advisor/investing/fed-funds-rate-history/

Figure 2: Betas around FOMC announcement (Jan 26, 2022)



The goal of this paper is to investigate this hypothesis a little further. To do this, I combine real-time news data with high-frequency stock returns to try to explain the intraday pattern of betas. News articles allow us to understand the different shocks that have affected the market each day. Distinguishing between cash flow and discount rate news, I can investigate how the dispersion of betas change depending on which type impacts the market and test whether the intraday pattern we see in the dispersion of betas is indeed driven by cash flow shocks, rather than discount rate shocks.

2 Model

Consider the loglinearization of Campbell and Shiller (1988) to decompose the returns of a stock i:

$$r^i_{t+1} - \mathbb{E}_t r_{t+1} = (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=0}^{\infty} \rho^j r_{t+1+j} = N^i_{CF,t} - N^i_{DR,t}$$

where

$$N^i_{CF,t} := (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} \qquad N^i_{DR,t} := (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=0}^{\infty} \rho^j r_{t+1+j}$$

we have that return innovations can be generated either by future cash flow shocks (e.g., dividends) or by market discount rates. This generates the distinction between the Cash Flow and the Discount Rate betas, which add up to the total beta:

$$\beta^i = \beta^i_{CF} + \beta^i_{DR}$$

Therefore, the current beta at any given time is the sum of its discount rate and the cash flow variants:

$$\beta^i_{CF} := \frac{\mathbb{C}\left(N^i_{CF,t}, r^M_t - \mathbb{E}\left(r^M_t\right)\right)}{\mathbb{V}\left(r^M_t - \mathbb{E}\left(r^M_t\right)\right)}$$

$$\beta^i_{DR} := \frac{\mathbb{C}\left(-N^i_{DR,t}, r^M_t - \mathbb{E}\left(r^M_t\right)\right)}{\mathbb{V}\left(r^M_t - \mathbb{E}\left(r^M_t\right)\right)}$$

2.1 Dispersion of Betas

We define the cross-sectional dispersion of betas as in Andersen et al. (2021) as the cross sectional variance of betas across all stocks in the sample:

$$\mathcal{D}\left(\beta\right) := \frac{1}{N}\sum_{i=1}^{N}\left(\beta^i - 1\right)^2$$

Including the decompositions of betas, this would imply that the dispersion of betas is the sum of the dispersion of its components plus a covariance term:

$$\mathcal{D}\left(\beta\right) = \mathcal{D}\left(\beta_{CF}\right) + \mathcal{D}\left(\beta_{DR}\right) + 2\frac{1}{N}\sum_{i=1}^{N}\left(\beta^i_{CF} - \bar{\beta}_{CF}\right)\left(\beta^i_{DR} - \bar{\beta}_{DR}\right)$$

For $\beta := \left(\beta^i\right)_{i\in N}$, $\beta_{CF} := \left(\beta^i_{CF}\right)_{i\in N}$ and $\beta_{DR} := \left(\beta^i_{DR}\right)_{i\in N}$. Where, $\bar{\beta}_{CF}$ and $\bar{\beta}_{DR}$ indicate the average cash flow and discount rate betas across all stocks in the sample.

For us to be able to identify the effect of each one of these shocks, we need to be able to separate the cash flow and discount rate effects. Under the assumption that a firm's cash flow shock do not impact the market prediction of future discount rate shocks, we can identify the dispersion of the cash. Equivalently, considering that discount rate shocks do not affect firm's future cash flows, we can identify the effect of discount rate shocks on dispersion.

This work suggests that the pattern we see in the intraday is mostly driven by cash flow shocks. Cash flow shocks tend to be more dispersed among the stocks. Think for example that Microsoft releases earnings announcements on a given day. We would expect that the company own stock would react the most to this news. Moreover, companies in the same sector will also be affected - either because of the spillover effect of the earnings announcement or because of the predicitve power of Microsoft's reports for the sector as a whole. We would also expect changes in the market for Microsoft suppliers and clients, though in a smaller magnitude. Therefore, cash flow shocks spread across the market with different intensities, generating a higher dispersion in the stock returns and in the realized beta on that day.

The same doesn't happen with discount rate shocks. As they tend to affect the whole market in a similar way, the dispersion of the realized beta is lower. For example, a change in the interest rate will affect all stocks in the same way or at least on a similar side, and therefore will tend to have way less spread out effects.

By classifying news releases in either cash flow or discount rate types, I can identify the effect that each type of news has on the dispersion of the cross-section of betas. If the initial dispersion of betas is indeed explained by the timing of the news release, then we should see a higher dispersion of the cross-section of betas in days in which the market is widely affected by cash flow shocks. This would imply that the initial dispersion of betas is indeed explained by the timing of the news release.

3 Data

3.1 Stock Returns

Stock returns are obtained from NYSE's Trades and Quotes (TAQ) database. I extract the stock prices at the second frequency for all the constituents of the S&P 500 index since 2000-01-01 (947 companies). I observe the market from 2004-01-01 to 2020-12-31.

3.2 News Data

To better understand when does news shocks happen and their overall impact, I use RavenPack Analytics dataset on real-time news. RavenPack Analytics (RPA) provide a large database on news articles posted in top business newsletters since 2000. This database is divided into RPA Equities and RPA Global Markets, the former contains information on company-specific news releases such as earnings announcements, press releases, changes to the CEO, mergers, etc. The latter contains information on macroeconomic news, such as GDP, inflation, FOMC announcements, geopolitical events and others. Their sources include the Dow Jones Newswires, Alliance News, Benzinga Pro, The Fly, FX Street News and more. The dataset is constructed at the milisecond frequency and provide both quantitative and qualitative measures to help us better understand the effect of that news on the market.

From unstructured publications the company extract information about the entities mentioned in the text, the sentiment of the article, the topic, the novelty of the information, and others. Using this information, I can identify whether news released in a given day are informative about the companies' cash flow or about the market discount rate and how (if at all) they might affect the stock returns (consequently affecting their betas).

The dataset provides, for every news article, information on

- *timestamp_utc*: the exact milisecond the article was received by the RPA servers (UTC time);
- *rp_story_id*: a unique identifier of that story (article);
- *rp_entity_id*: the identifier of the entity (company, organization, nation, etc.) mentioned in the story;
- *relevance*: a score indicating how important that entity is for that story;
- *event_sentiment_score* (ESS): a score representing the overall sentiment for that given entity in the story;
- *event_similarity_key*: which identifies whether a similar story has been published in the past year;
- *topic, group, type, subtype*: a taxonomic classification of the story;

- *fact_level*: whether the article is about a *fact*, *forecast* or *opinion*;

- *CSS* (Composite Sentiment Score): a score representing the overall sentiment of that story from "emotionally charged words";

- *NIP* (News Impact Projection): a score representing the short-term impact of that story to the market;

From the taxonomy we can classify whether the news is related to the company's cash flow or discount rate and using the sentiment score (CSS) we can capture the magnitude of the impact of that news to the market. Using this measures, I construct the *cash flow* and *discount rate indexes* to be used in further analysis.

For cash flow news I consider the entire dataset of RPA Equities, meaning every news articles that mention a given company. Discount rate news, on the other hand, are considered as all the articles in the RPA Global Macro database that are classified as being about *Country* of *Organization*.

4 Methodology

4.1 Estimation

I estimate the betas by following Andersen, Riva, Thyrsgaard, and Todorov (2023) which pools the observations across trading days. The estimation considers discrete observations of stock returns r_t^i for $i = 1, \ldots, N$ and market returns r_t^M. Returns are sampled at equidistant time $0, \frac{1}{n}, \frac{2}{n}, \ldots, T$ and I denote the sampling frequency by $\Delta_n \equiv 1/n$. The common notation will define t as a specific date and $k\Delta_n$ as a specific time on the day so that the time index is represented as $t + k\Delta_n$ for $t = 1, \ldots, T$ and $k = 0, \ldots, n$. I estimate the covariances needed to compute the betas using local windows as in Andersen et al. (2021) defined as sets $\mathcal{I}_\kappa$ of length K where

$$\mathcal{I}_\kappa = \{\lfloor \kappa n \rfloor - K + 1, \lfloor \kappa n \rfloor - K + 2, \ldots, \lfloor \kappa n \rfloor\}$$

The betas are then astimated as a ratio of realized covariances on those intervals:

$$\widehat{V}_{t,\kappa} := \frac{n}{K} \sum_{k \in \mathcal{I}_\kappa} \left(r^M_{t+k\Delta_n}\right)^2 \tag{1}$$

$$\widehat{C}^i_{t,\kappa} := \frac{n}{K} \sum_{k \in \mathcal{I}_\kappa} r^i_{t+k\Delta_n} r^M_{t+k\Delta_n} \tag{2}$$

and we define

$$\widehat{\beta}^i_{t,\kappa} := \frac{\widehat{C}^i_{t,\kappa}}{\widehat{V}_{t,\kappa}} \tag{3}$$

I then calculate the dispersion of betas as explained in Section 2.

4.2 Indexes

Using the RPA dataset, I assign as cash flow news all articles related to companies in the database[5] and as discount rate news all articles in the Macro database which are classified as being

[5]This require having *entity_type* = 'COMP'

about a country or organization[6]. Discount rate news only consider news about the United States and Canada. To gain understanding on the amount of information the market receives in a given day, I construct the *cash flow* and *discount rate* indexes by summing the absolute value of the CSS of all news released in that day. The time series of these indexes is displayed in Figure 3 (on a 252-days moving average). Due to the nature of the data, I decided to work with the indexes from 2004 to 2020, in which the indexes are more stable and do not seem to exhibit any year fixed effects.

Figure 3: Cash Flow and Discount Rate Indexes

5 Results

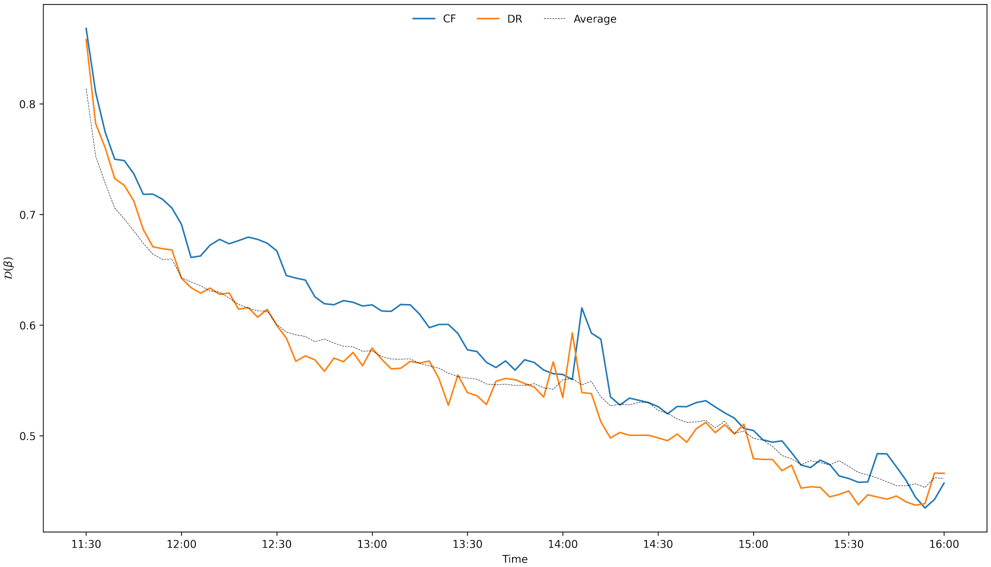
I start by looking at days in which the market is bighly affected by CF or DR shocks and compare them with days with low impact. I define high-intensity CF days as dates in which the cash flow index is on the top decile and low-intensity CF days in which the index is on the bottom decile. Similarly, I define high-intensity DR days as dates in which the discount rate index is on the top decile and low-intensity DR days in which the index is on the bottom decile. I then calculate the average dispersion of betas under these conditions:

$$\mathbb{E}\left[\mathcal{D}(\beta_{t,h} \mid \mathrm{CF}_t \geq Q_{0.9}(CF_t))\right] vs \mathbb{E}\left[\mathcal{D}(\beta_{t,h} \mid \mathrm{DR}_t \leq Q_{0.1}(DR_t))\right]$$

This comparison is seen on Figure 4. Comparing the two, we see that the dispersion of betas conditional on high CF dates is higher than the average, as the dispersion of betas conditional on high DR dates does not differs much from the unconditional mean over the sample.

[6]That removes news with *entity_type* as product ('PROD'), people ('PEOP') and commodities ('CMDT')

Figure 4: Dispersion of Betas on High CF vs High DR dates



This suggests that cash flow shocks do tend to increase the dispersion compared to the average. On the other hand, discount rate shocks do not change much compared to the average.

However, to be able to assess the pure effect of news, I need to control for the correlation between the cash flow and discount rate indexes and obtain the expected dispersion under days in which only one of the effects is present. To do so, I estimate a nonparametric model of the dispersion using Local Linear Kernel Regressions. For each time of the trading day h, I estimate the conditional average of the dispersion given my cash flow and discount rate indexes:

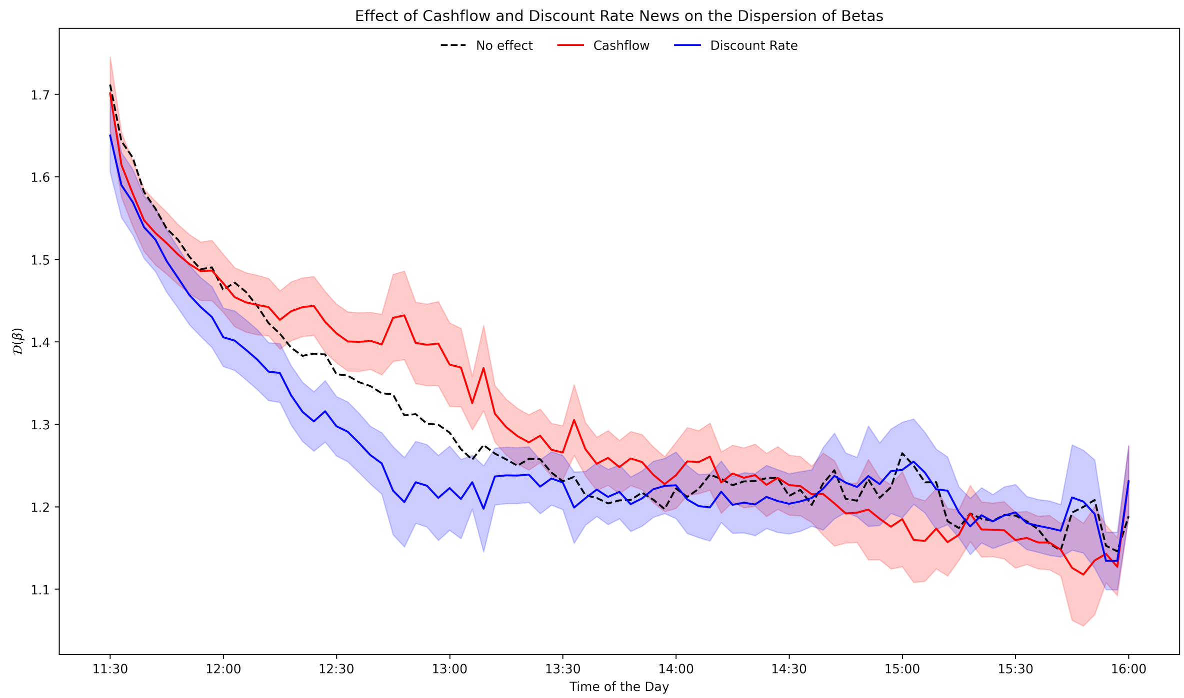
$$\mathbb{E}\left[\mathcal{D}\left(\beta_{t,h}\right) \mid \mathrm{CF}_t = x_1, \mathrm{DR}_t = x_2\right] = f_h(\mathrm{CF}_t, \mathrm{DR}_t)$$

The LLN estimator is given by:

$$f_h(x_1, x_2) = b_0(x_1, x_2) + b_1(x_1, x_2)(CF_t - x_1) + b_2(x_1, x_2)(DR_t - x_2)$$

Using these regressions, I fit the model in two different scenarios. The first scenarios, identifying cash flow days (labeled CF), takes $CF_t = Q_{0.9}(CF_t)$ and $DR_t = Q_{0.1}(DR_t)$. The second scenario, which identifies discount rate days (labeled as DR), takes $DR_t = Q_{0.9}(DR_t)$ and $CF_t = Q_{0.1}(CF_t)$. The results are shown in Figure 5 alongside estimated standard errors. As we can see, the results point to a higher dispersion of betas under days in which the market is highly affected by cash flow news. Although both of the series tend to have large dispersion values on the first hour (carried from the probably noisy initial dispersion on market open), the DR series tends converges quickly and stays flat throughout the day after 12:30. On the other hand, CF shows a slower rate of conversion and persistently high values of dispersion throughout the whole day. This indicates that cash flow shocks indeed have a higher impact on the dispersion of betas than discount rate shocks.

Figure 5: Estimation of Expected Dispersion on CF vs DR days



6 Conclusion

The results of this paper points to the hypothesis that cash flow shocks are the main driver of the intraday pattern that we observed in the dispersion of betas following Andersen et al. (2021). Combining real-time news data with high-frequency stock returns, I am able to identify for days in which the market is highly affected by cash flow (discount rate) shocks and then observe separately the effect of each one in the cross-sectional distribution of betas. The results have demonstrated that, under days in which the market is highly affected by cash flow news, dispersion tends to be higher throughout most part of the trading day and it has a clearer downward trend. On the other hand, during days with great discount rate news, betas tend to be more concentrated if we exclude the first hour of trading. Moreover, dispersion under DR days tends to be almost flat and not show any significant trend. This relates to the results initially observed in Andersen et al. (2021).

6.1 Further Research

This findings suggest interesting questions for further research. As the pattern seems to be systematically across the market, we can ask whether the initial dispersion opens space for active traders to participate on the market, gathering short-term alphas generated by the informational shock. As these investors start to agree on the new price, risk tends to concentrate around one and the market looks less volatile. Therefore, passive investors, which follow systematic trading strategies, would be able to enter the market to rebalance their portfolios. This implies an interesting trading dynamic which can be studied in future research.

Moreover, this approach also allow us to answer other questions in the literature such as “do news cause betas to jump?”. Related work is done by Patton and Verardo (2012), which analyse the impact of earnings announcements on betas. By applying this more comprehensive dataset to the question, I can assess whether this is a feature purely of earnings shocks or if it extends to other type of information releases as well.

References

Torben G. Andersen, Martin Thyrsgaard, and Viktor Todorov. Recalcitrant betas: Intraday variation in the cross-sectional dispersion of systematic risk. *Quantitative Economics*, 12(2):647–682, 2021. ISSN 1759-7323. doi: 10.3982/QE1570. URL http://qeconomics.org/ojs/index.php/qe/article/view/QE1570.

Torben G Andersen, Raul Riva, Martin Thyrsgaard, and Viktor Todorov. Intraday cross-sectional distributions of systematic risk. *Journal of Econometrics*, 235(2):1394–1418, 2023.

Matthias Apel, André Betzer, and Bernd Scherer. Real-time transition risk. *Finance Research Letters*, 53:103600, May 2023. ISSN 15446123. doi: 10.1016/j.frl.2022.103600. URL https://linkinghub.elsevier.com/retrieve/pii/S1544612322007760.

H Kent Baker, Yang Ni, Samir Saadi, and Hui Zhu. Competitive earnings news and post-earnings announcement drift. *International Review of Financial Analysis*, 63:331–343, 2019.

Brad M. Barber and Terrance Odean. All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies*, 21 (2):785–818, April 2008. ISSN 0893-9454, 1465-7368. doi: 10.1093/rfs/hhm079. URL https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhm079.

Thomas D. Berry and Keith M. Howe. Public Information Arrival. *The Journal of Finance*, 49 (4):1331–1346, September 1994. ISSN 00221082. doi: 10.1111/j.1540-6261.1994.tb02456.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1994.tb02456.x.

Jacob Boudoukh, Ronen Feldman, Shimon Kogan, and Matthew Richardson. Information, Trading, and Volatility: Evidence from Firm-Specific News. *The Review of Financial Studies*, 32(3): 992–1033, March 2019. ISSN 0893-9454, 1465-7368. doi: 10.1093/rfs/hhy083. URL https://academic.oup.com/rfs/article/32/3/992/5061375.

Phillip A. Braun, Daniel B. Nelson, and Alain M. Sunier. Good News, Bad News, Volatility, and Betas. *The Journal of Finance*, 50(5):1575–1603, December 1995. ISSN 00221082. doi: 10.1111/j.1540-6261.1995.tb05189.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1995.tb05189.x.

Jonathan Brogaard, Terrence Hendershott, and Ryan Riordan. High-Frequency Trading and Price Discovery. *Review of Financial Studies*, 27(8):2267–2306, August 2014. ISSN 0893-9454, 1465-7368. doi: 10.1093/rfs/hhu032. URL https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhu032.

Charles W. Calomiris and Harry Mamaysky. How news and its context drive risk and returns around the world. *Journal of Financial Economics*, 133(2):299–336, August 2019. ISSN 0304405X. doi: 10.1016/j.jfineco.2018.11.009. URL https://linkinghub.elsevier.com/retrieve/pii/S0304405X18303180.

John Y Campbell and Robert J Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228, 1988.

John Y Campbell and Tuomo Vuolteenaho. Bad beta, good beta. *American Economic Review*, 94 (5):1249–1275, 2004.

Zhi Da and Mitchell Craig Warachka. Cashflow risk, systematic earnings revisions, and the cross-section of stock returns. *Journal of Financial Economics*, 94(3):448–468, 2009.

Robert F. Engle and Victor K. Ng. Measuring and Testing the Impact of News on Volatility. *The Journal of Finance*, 48(5):1749–1778, December 1993. ISSN 00221082. doi: 10.1111/j.1540-6261.1993.tb05127.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1993.tb05127.x.

Robert F Engle, Martin Klint Hansen, Ahmet K Karagozoglu, and Asger Lunde. News and Idiosyncratic Volatility: The Public Information Processing Hypothesis*. *Journal of Financial Econometrics*, 19(1):1–38, March 2021. ISSN 1479-8409, 1479-8417. doi: 10.1093/jjfinec/nbaa038. URL https://academic.oup.com/jfec/article/19/1/1/6043100.

Andrea Frazzini and Lasse Heje Pedersen. Betting against beta. *Journal of Financial Economics*, 111(1):1–25, January 2014. ISSN 0304405X. doi: 10.1016/j.jfineco.2013.10.005. URL https://linkinghub.elsevier.com/retrieve/pii/S0304405X13002675.

Kenneth R French and Richard Roll. Stock return variances: The arrival of information and the reaction of traders. *Journal of financial economics*, 17(1):5–26, 1986.

Matthew Gentzkow, Bryan Kelly, and Matt Taddy. Text as Data. *Journal of Economic Literature*, 57(3):535–574, September 2019. ISSN 0022-0515. doi: 10.1257/jel.20181020. URL https://pubs.aeaweb.org/doi/10.1257/jel.20181020.

Peter Reinhard Hansen and Yiyao Luo. Robust estimation of realized correlation: New insight about intraday fluctuations in market betas. *arXiv preprint arXiv:2310.19992*, 2023.

Terrence Hendershott, Dmitry Livdan, and Dominik Rösch. Asset pricing: A tale of night and day. *Journal of Financial Economics*, 138(3):635–662, 2020.

David Hirshleifer and Jinfei Sheng. Macro news and micro news: complements or substitutes? *Journal of Financial Economics*, 145(3):1006–1024, 2022.

Harrison Hong and David A. Sraer. Speculative Betas: Speculative Betas. *The Journal of Finance*, 71(5):2095–2144, October 2016. ISSN 00221082. doi: 10.1111/jofi.12431. URL https://onlinelibrary.wiley.com/doi/10.1111/jofi.12431.

Jean Jacod and Philip Protter. *Discretization of processes*, volume 67. Springer Science & Business Media, 2011.

Larry HP Lang and RenéM Stulz. Contagion and competitive intra-industry effects of bankruptcy announcements: An empirical analysis. *Journal of financial economics*, 32(1):45–60, 1992.

Tim Loughran and Bill McDonald. Textual Analysis in Finance.

Mark L. Mitchell and J. Harold Mulherin. The Impact of Public Information on the Stock Market. *The Journal of Finance*, 49(3):923–950, July 1994. ISSN 00221082. doi: 10.1111/j.1540-6261.1994.tb00083.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1994.tb00083.x.

Andrew J. Patton and Michela Verardo. Does Beta Move with News? Firm-Specific Information Flows and Learning about Profitability. *Review of Financial Studies*, 25(9):2789–2839, September 2012. ISSN 0893-9454, 1465-7368. doi: 10.1093/rfs/hhs073. URL https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhs073.

Roel Popping. Knowledge Graphs and Network Text Analysis. *Social Science Information*, 42 (1):91–106, March 2003. ISSN 0539-0184, 1461-7412. doi: 10.1177/0539018403042001798. URL http://journals.sagepub.com/doi/10.1177/0539018403042001798.

Sundaresh Ramnath. Investor and analyst reactions to earnings announcements of related firms: An empirical analysis. *Journal of Accounting Research*, 40(5):1351–1376, 2002.

S Rozeff and R Kinney. CAPITAL MARKET EEASONALITY: THE CASE OF STOCK RETURNS.

William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964.

Lee A. Smales. News sentiment and the investor fear gauge. *Finance Research Letters*, 11(2): 122–130, June 2014. ISSN 15446123. doi: 10.1016/j.frl.2013.07.003. URL https://linkinghub.elsevier.com/retrieve/pii/S1544612313000354.

Paul C. Tetlock. Information Transmission in Finance. *Annual Review of Financial Economics*, 6(1):365–384, December 2014. ISSN 1941-1367, 1941-1375. doi: 10.1146/annurev-financial-110613-034449. URL https://www.annualreviews.org/doi/10.1146/annurev-financial-110613-034449.

Paul C. Tetlock. The Role of Media in Finance. In *Handbook of Media Economics*, volume 1, pages 701–721. Elsevier, 2015. ISBN 978-0-444-63691-1. doi: 10.1016/B978-0-444-63685-0.00018-8. URL https://linkinghub.elsevier.com/retrieve/pii/B9780444636850000188.

Paul C. Tetlock, Maytal Saar-Tsechansky, and Sofus Macskassy. More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance*, 63(3):1437–1467, June 2008. ISSN 00221082. doi: 10.1111/j.1540-6261.2008.01362.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2008.01362.x.

Wen Yang, Dongtong Lin, and Zelong Yi. Impacts of the mass media effect on investor sentiment. *Finance Research Letters*, 22:1–4, August 2017. ISSN 15446123. doi: 10.1016/j.frl.2017.05.001. URL https://linkinghub.elsevier.com/retrieve/pii/S1544612317302465.