Journal of Banking and Finance 89 (2018) 59–77

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Institutional trading and asset pricing☆

Bart Frijns[a], Thanh D. Huynh[b,*], Alireza Tourani-Rad[a], P. Joakim Westerholm[c]

[a] Department of Finance, Auckland University of Technology, Auckland, New Zealand
[b] Department of Banking and Finance, Monash Business School, Monash University, 900 Dandenong Rd, Caulfield East, Victoria 3145, Australia
[c] The University of Sydney Business School, Australia

ARTICLE INFO

Article history:
Received 21 September 2016
Accepted 31 January 2018
Available online 3 February 2018

Keywords:
Intermediary asset pricing
Institutional trading
Investor preferences
CAPM

ABSTRACT

This paper examines whether the trading activity of different investor types, institutional versus retail, can affect the relation between beta and average returns. We find that the beta-return relation is strong and positive on days with high institutional trading activity, and negative and significant on low institutional trading days. Our findings are robust and not driven by recently documented effects such as macroeconomic news and leverage constraints, among others. The evidence is consistent with the hypothesis that the preferences and characteristics of various investor types, which are revealed through their trading activity, cause the slope of the Security Market Line to change.

1. Introduction

Recent research has revisited the issue of the insignificant relation between beta and average returns. Frazzini and Pedersen (2014) suggest that the flat relation could be attributed to leverage constraints and show that a portfolio that is long low-beta stocks and short high-beta stocks is profitable. Savor and Wilson (2014) find that the relation is conditional on news announcements and document a positive relation on days with macroeconomic announcements. Hong and Sraer (2016) show that, in the presence of limits to arbitrage and strong investor disagreement, the Security Market Line (SML) could be downward sloping.

In this paper, we provide new evidence on this debate by examining the impact of the trading activity of different investor types on the slope of the SML. Specifically, we obtain full records of all trades conducted by different trader types on all stocks listed on the Nasdaq OMX Helsinki over the period 1996–2011 from Euroclear Finland Ltd. (Euroclear). From these records, we construct a measure of institutional trading (IT) activity as the fraction of total trading volume by all institutions over total market volume, and label a particular day as a high institutional trading (High-IT) day when IT on that day exceeds its average over the past quarter. We find strong evidence that on days with high institutional trading activity, beta is positively related to average returns, while on low institutional trading days, this relation is significantly negative. These findings hold for beta-sorted, size and book-to-market ratio (BM), and industry portfolios, as well as for individual stocks. Our results are robust to using alternative measures of IT activity, as well as controlling for the effect of positive market returns. They are not driven by well-known anomalies such as the January and turn-of-month effects (Sikes, 2014), or Nokia's stock, which is by far the largest and most liquid stock in the Finnish market, and cannot be explained by the leverage-constraints hypothesis (Frazzini and Pedersen, 2014). While we confirm the presence of a macroeconomic announcement effect as in Savor and Wilson (2014), we find that our results are distinct from this effect.

Our findings highlight the importance of financial institutions in shaping the relation between beta and returns. We explore three channels that could explain the role of institutions in shaping this relation. The first channel builds on the literature related to intermediary-based asset pricing (see, e.g., He and Krishnamurthy, 2013; Adrian et al., 2014). This literature suggests that financial intermediaries are the marginal investors who price the cross-section of stock returns. Extending this hypothesis to explain the flat-beta puzzle, we could expect that when institutions trade more actively, assets should be priced more accurately, and we would expect a positive slope of the SML on days when institutions are

☆ We are grateful to the Editor, Geert Bekaert, and two anonymous referees for their extremely helpful comments and suggestions. We thank Jonathan Brogaard, Stephen Brown, Nick Bollen, Zhi Da, Ron Giammarino, Andrew Hertzberg, Raymond Kan, Andrew Karolyi, Marios Panayides, and Pavel Savor, and workshop participants at the SFS Cavalcade 2016, McGill Global Asset Management Conference 2015, FMA Annual Meeting Orlando 2015, Conference of the Multinational Finance Society 2015, World Finance Conference 2015, Auckland Finance Meeting 2014, AUT, Massey University, Monash University, QUT, and University of Technology, Sydney. Thanh Huynh gratefully acknowledges the financial support from AUT Research Grant 2014 and SIRCA/AFAANZ Developing Researcher Grant 2014. All errors are our own.

* Corresponding author.

E-mail addresses: bfrijns@aut.ac.nz (B. Frijns), thanh.huynh@monash.edu (T.D. Huynh), tourani@aut.ac.nz (A. Tourani-Rad), joakim.westerholm@sydney.edu.au (P.J. Westerholm).

https://doi.org/10.1016/j.jbankfin.2018.01.018

very active. Thus, our findings of the positive relation between beta and average returns on High-IT days are consistent with this hypothesis.

The financial intermediary hypothesis further predicts that larger institutions should have a larger impact on the pricing of assets. Indeed, we observe that it is the trading activity of the most active institutions that can explain the switching behavior of the SML. This finding is in line with what would be expected from intermediary-based asset pricing models, and confirms the findings of Siriwardane (2016), who shows that the capital fluctuations of the largest sellers of protection in the credit default swap (CDS) market have the strongest impact on CDS spread movements.

The second channel for the effect of trading activity on the SML is based on Hong and Sraer (2016), who theoretically show that investor disagreement can affect the slope of the SML. Extending this framework, we argue that in a market with different types of investors, some with low disagreement and others with high disagreement among each other about high-beta stocks, the slope of the SML will change depending on who is more active in the market. Individual investors differ in their ability to interpret noisy signals, while institutions are more likely to have low levels of disagreement among each other. Thus, when high-disagreement investors (individuals) are more active, the SML could be downward sloping, whereas when low-disagreement investors (institutions) are more active, the SML could be upward sloping. Indeed, the household finance literature documents that, compared to institutions, individual investors are more heterogeneous. Their trading activity depends on their characteristics and preferences such as gender, language, and culture (Grinblatt and Keloharju, 2001a), biological factors such as genetic differences (Cronqvist and Siegel, 2014), growing up during the Great Depression (Cronqvist et al., 2015), IQ scores (Grinblatt et al., 2012), sensation seeking, overconfidence (Grinblatt and Keloharju, 2009), stocks with early alphabet names (Itzkowitz et al., 2015), and attention (Barber and Odean, 2008). Institutions, however, tend to base their trades on more rational trading models and are considered sophisticated investors (Foster et al., 2011; Adrian et al., 2014; Barber and Odean, 2008; Grinblatt and Keloharju, 2001b; Kumar, 2009). Consistent with this hypothesis, when we consider news dispersion as a proxy for disagreement, we find that investor disagreement is lower when institutions trade intensively, suggesting that trading activity could reflect disagreement levels among investor types.

The third channel for the effect of trading activity on the slope of the SML can be seen in the framework of Mitton and Vorkink (2007), who show that when there are different groups of investors, Traditional versus lottery-type investors, their different preferences affect the valuation of stocks. Lottery-type investors have preferences for lottery-like stocks (low price, high beta, positive skewness, see Kumar, 2009) and overvalue these securities, whereas traditional investors value stocks based on mean-variance optimality. Time variation in the activity of different investor types could thus affect the slope of SML: when traditional investors are more active, the SML is expected to be upward sloping and vice versa. By comparing lottery measures between High- and Low-IT days, we indeed find that lottery preference is stronger on Low-IT days than High-IT days, suggesting that trading activity could reflect investor-type lottery preferences.

In summary, the main contribution that our paper makes is with regard to the literature on the risk-return tradeoff puzzle. Specifically, we document a novel finding on the contrasting slope of the SML on two types of days: High- versus Low-IT. Our study has an analogue to Savor and Wilson (2014), who find that the Capital Asset Pricing Model (CAPM) holds on U.S. Federal Open Market Committee (FOMC) announcement days. However, given that non-announcement days constitute the majority of trading days in a year, whether the risk-return tradeoff holds on non-announcement days remains unanswered in the literature. We find that the IT effect is distinct and holds even on non-announcement days.

Our study also joins the growing literature on the theory of intermediary-based asset pricing (He and Krishnamurthy, 2013; Adrian et al., 2014; Siriwardane, 2016; He et al., 2017), which acknowledges the central role of financial institutions. In particular, Siriwardane (2016) shows that the largest sellers of protection in the CDS market determine the CDS pricing. We extend this hypothesis to the flat-beta puzzle and show that the trading activity of the most active financial institutions drives the positive relation between beta and average returns.

The remainder of the paper is structured as follows. Section 2 describes the unique data set employed in our study and the construction of test assets as well as the empirical methodology. Section 3 presents the main findings of the empirical analysis and robustness tests. In Section 4, we test three potential channels to explain our results. In Section 5, we explore alternative explanations. Finally, Section 6 concludes the paper.

2. Data and empirical methodology

In this section, we discuss our data for the Finnish market, which offers daily trading records of all financial institutions. This aspect of the data gives us a comparative advantage over other empirical U.S.-based asset pricing studies that have to rely on quarterly institutional holdings data or a small subset of the market. Nevertheless, using quarterly holdings data from the U.S., we confirm that the main conclusions do not qualitatively change.

2.1. Data

We employ daily trading records of financial institutions for Finnish stocks from Euroclear (also known as Finland's Central Securities Depository). This database contains trades by all institutions (identified by a unique number of each institution aggregated at the daily level) while most U.S. data cover only large institutions (small financial firms do not file 13Fs). These features of our database allow us to compute a timely measure of the systematic impact of institutional trading. The data set has separate fields for institutional buy and sell volumes. Our data comprise 187 stocks listed on the Nasdaq OMX Helsinki exchange between 1996 and 2011 (see Grinblatt and Keloharju, 2000, Grinblatt and Keloharju, 2001b and Leung et al., 2014 for a detailed description of the Euroclear database and its classification of institutional and retail investors).

We collect daily stock prices, dividends, capitalization adjustments, and the number of shares outstanding from Compustat Global. Following Ince and Porter (2006) and Griffin et al. (2010), filters are applied on individual stock returns to eliminate data errors. Specifically, returns that exceed 100% in one day are treated as missing. If daily returns exceed 20% and reverse the following day, then returns on both days are treated as missing. Book values of equity are obtained from WorldScope. For tests that use the central bank's interest rate announcement, we collect scheduled days of monetary policy announcements from the European Central Bank (ECB) website from 1999 when the ECB was officially established. Although tests that employ ECB announcement days are limited to the period between 1999 and 2011, the daily frequency of our data gives us well over 3100 trading days, with approximately 370,000 of stock-day observations over this shorter period. All returns are in euros, and excess returns are above the five-year government bond yield obtained from Datastream. Betas are computed with respect to value-weighted market returns. Our main results do not qualitatively change when we use the MSCI index.

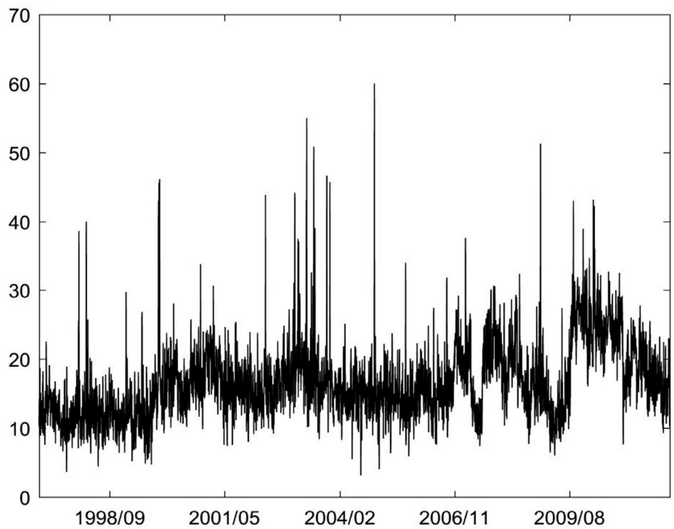


Fig. 1. Time-series plot of institutional trading. The figure plots the time series variation in institutional trading (IT) activity over time. IT is the ratio of daily institutional trading volume over the total market volume. The sample period is between 1997 and 2011.

2.2. Methodology

We first compute the aggregate institutional trading activity as $IT_t = \sum_i (buy_{i,t} + sell_{i,t})/total\ market\ volume$, where $buy_{i,\ t}$ and $sell_{i,\ t}$ are buy and sell volume of financial institutions, respectively. We define a day as a High-IT day when that day's IT_t is higher than its average over the past quarter.[1] One can think of our measure as a time series dummy variable that takes a value of one on High-IT days and zero otherwise. In some of the tests, we separately examine institutions' buying ($IT_{buy,t} = \sum_i buy_{i,t}/total\ market\ volume$) and selling activity ($IT_{sell,t} = \sum_i sell_{i,t}/total\ market\ volume$).

In Fig. 1, we provide a time series plot of the fraction of institutional trading volume over total volume. While there appears to be some persistence in the fraction of institutional trading volume, the graph also shows that there is quite some variation in this fraction. There are 1598 days out of 3567 days being classified as High-IT days. In addition, the probability that a High-IT day is followed by a High-IT day is 59.64%, while the probability that a Low-IT day is followed by a Low-IT day is 71.66%. There is limited evidence of the impact of the occurrence of macroeconomic announcements on the classification of High-IT days: of the 177 days with macroeconomic announcements, 85 are classified as High-IT days. The classification of High-IT days is also not driven by financial crises. For the 754 days that fall into the period January 2007 to December 2009 (broadly covering the Global Financial Crisis), 330 are classified as High-IT days. Thus, we can conclude that High-IT days are not confined to specific periods or events.

Our empirical tests employ the two-pass (Fama and MacBeth, 1973) procedure for the CAPM and then examine the coefficient estimates on High- and Low-IT days. First, we estimate betas using one-year rolling regressions, adjusting for the potential effect of non-synchronous trading by using Dimson's (1979)) "sum" betas. Specifically, we run the following regression:

$$R_{i,t} = \alpha_i + \beta_{i,0}R_{M,t} + \beta_{i,1}R_{M,t-1} + \beta_{i,2}\sum_{k=2}^{4} R_{M,t-k}/3 + \epsilon_{i,t}$$

where $R_{i,\ t}$ and $R_{M,\ t}$ are excess returns on asset i and the market index, respectively. The Dimson sum beta is then $\beta_0 + \beta_1 + \beta_2$.[2] The test assets are five beta-sorted portfolios, nine Fama and French size- and BM-sorted portfolios, five industry portfolios, and individual stocks.

In the second-stage regression, we run the following cross-sectional regressions:

$$R^H_{i,t+1} = \gamma^H_{0,t} + \gamma^H_{1,t}\hat{\beta}_{i,t}, \quad (1)$$

$$R^L_{i,t+1} = \gamma^L_{0,t} + \gamma^L_{1,t}\hat{\beta}_{i,t}, \quad (2)$$

where $\hat{\beta}_{i,t}$ is asset i's market beta for period t estimated in the first stage; and $R^H_{i,t+1}$ is the excess return on the test asset on high institutional trading days (High-IT or High).

As we are interested in studying the marginal effect of High-IT days on the relation between beta and average returns, we focus on the difference in the coefficient estimates, $\gamma^H_{1,t} - \gamma^L_{1,t}$. Standard errors are computed using Newey and West's (1987)) method.

Following Savor and Wilson (2014), we also estimate a pooled regression to test the difference in the implied market risk premium between High- and Low-IT days:

$$R_{i,t+1} = \gamma_0 + \gamma_1\hat{\beta}_{i,t} + \gamma_2 High_{t+1} + \gamma_3\hat{\beta}_{i,t}High_{t+1} + \epsilon_{i,t+1}, \quad (3)$$

where $High_t$ is a dummy variable that equals one for High-IT days and zero otherwise. We compute clustered standard errors by date, which adjust for the cross-sectional correlation of the residuals.[3] We expect the coefficient γ_3 to be positive, suggesting that the risk premium on High-IT days is higher than on Low-IT days.

3. Results

In this section, we present evidence on the difference in the market risk premium between High- and Low-IT days. We document these results for various test portfolios. We further confirm that our findings are robust for individual stocks, subsample periods, the January effect, the turn-of-month effect, and the use of U.S. 13F institutional holdings data.

3.1. Beta, book-to-market, size, and industry portfolios

We start our analysis by constructing various test portfolios. We form nine size- and BM-sorted portfolios in the spirit of Fama and French (1993). Stocks are sorted into three groups on the basis of their market capitalization at the end of June of each year (due to the small size of the Finnish market, sorting stocks into three bins helps to maintain a certain level of diversification in each portfolio as well as the power of our tests). Big stocks are those in the top 30% of the market capitalization and small stocks are those in the bottom 30%. Independently, stocks are also sorted into three groups on the basis of their BM ratios. We use book values for the fiscal year ending in calendar year $t-1$, while market capitalization is for the end of December in calendar year $t-1$. The nine size-BM portfolios are thus the intersections of three size and three BM portfolios. We also form five beta-sorted portfolios and five industry portfolios using Fama and French's SIC classifications.[4] Except for the beta-sorted portfolios that are rebalanced monthly, all portfolios are rebalanced annually.

[1] The results are robust to using the underlying continuous IT variable, as well as using various windows (e.g., monthly, yearly) to define High-IT days. These results are available upon request.

[2] Our results are robust to using no Dimson adjustment.

[3] According to prior studies (e.g., Fama and French, 1992; Ang et al., 2006), the CAPM makes predictions about the relation between realized beta and contemporaneous returns. For example, Ang et al. (2006, p. 1201) state that "the CAPM predicts an increasing relationship between realized average returns and realized factor loadings,... More generally, a multifactor model implies that we should observe patterns between average returns and sensitivities to different sources of risk over the same time period used to compute the average returns and the factor sensitivities."

[4] The classification is downloaded from Ken French's website http://www.mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Table 1
Summary statistics for test portfolios.

Year	Mean firms	Mean ME ('000,000)	Mean fraction of total IT volume	Mean fraction of sell volume	Number of accounts
Panel A: summary statistics for the Finnish market					
1995–1999	79	389.868	0.196	0.123	563
2000–2003	138	1767.843	0.169	0.104	643
2004–2007	141	1335.035	0.232	0.128	722
2008–2011	140	1095.511	0.379	0.251	656

	Estimate				*t*-stat			
Panel B: size-BM portfolios								
	Low	M	High	HL	Low	M	High	HL
		α				t_α		
Small	0.00081	0.00099	0.00042	−0.00039	3.46	1.26	1.26	−1.10
M	0.00037	0.00030	0.00002	−0.00034	1.80	1.79	0.12	−1.33
Big	0.00032	0.00037	0.00007	−0.00025	1.40	1.85	0.40	−0.81
SB	0.00049	0.00062	0.00035		1.63	2.01	0.94	
		β				t_β		
Small	0.26526	0.21463	0.31030	0.04504	23.31	16.89	19.06	2.64
M	0.30751	0.27687	0.38387	0.07636	30.98	33.77	39.42	6.09
Big	0.44942	0.45382	1.11708	0.66765	40.63	47.00	131.11	44.90
SB	−0.18416	−0.23918	−0.80678		−12.57	−15.91	−44.49	
		Adj. R^2						
Small	0.1170	0.0650	0.0814					
M	0.1897	0.2176	0.2748					
Big	0.2870	0.3501	0.8074					

Panel C: beta portfolios												
	Low	2	3	4	High	HL	Low	2	3	4	High	HL
alpha	0.00059	0.00030	0.00061	0.00033	−0.00006	−0.00065	2.91	1.79	3.16	1.65	−0.38	−2.50
Beta	0.34432	0.28650	0.34773	0.43455	1.10040	0.75608	35.77	36.24	37.91	45.56	143.83	61.21
Adj. R^2	0.2506	0.2555	0.2730	0.3516	0.8439	0.4946						

Panel D: industry portfolios										
	G1	G2	G3	G4	G5	G1	G2	G3	G4	G5
Alpha	0.00022	−0.00010	0.00017	0.00014	−0.00014	1.29	−0.54	0.86	0.66	−0.56
Beta	1.02426	0.56672	0.58071	0.67398	0.67951	123.00	59.62	61.37	65.26	55.25
Adj. R^2	0.7812	0.4562	0.4706	0.5012	0.4186					

Note: This table reports summary statistics and estimates from time series regressions of daily portfolio returns on excess returns on the market. In panel A, we present several summary statistics of our sample over different subperiods, including the number of companies in the sample, the average size of the companies, the fraction of volume due to institutional trades, the fraction of sell volume, and the number of trading accounts. Panels B to D present regression results for various test portfolios, where the left-hand side panel presents coefficient estimates, while the right-hand side panel reports the corresponding t-statistics. t-statistics are computed using Newey-West standard errors with five lags.

Panel A of Table 1 reports summary statistics for the Finnish market. The number of firms listed on the Helsinki exchange increases from an average of 79 firms per year with an average market capitalization of 389.9 million euros at the beginning of the sample to 140 firms per year with an average market capitalization of 1095.5 million euros in the final years. The average fraction of institutional trading volume (buy plus sell volume over the total market volume) increases from 19.6% in the first five years to 37.9% in the last four years of the sample. There is also a sharp increase in sell volume by institutions after 2008. Between 2008 and 2011, institutions' sell fraction (over the total market volume) is 25.1%, which is almost double the sell fraction of earlier years. The number of trading accounts held by institutions increases from 563 in the first subperiod to 722 in the third subperiod, but then decreases to 656 for the last subperiod (2008–2011). Due to the substantial rise in sell volume in the last subsample (which aligns with the Global Financial Crisis and European Debt Crisis), we ensure in one of the robustness tests that our results are not specific to this subperiod.

Panel B reports CAPM regression results for the nine value-weighted size-BM sorted portfolios (we report alphas and betas along with their respective *t*-statistics, as well as the adjusted R^2 of the regressions). While we observe some evidence of a size effect (alpha is significant for medium BM stocks and on the borderline of being significant for low BM stocks), there is no evidence of a value effect. Overall, there is little evidence for the size and value effects being present in the Finnish stock market. Panel C presents CAPM alphas and betas for the five beta-sorted portfolios. The beta puzzle is apparent in the Finnish market where the high-beta portfolio significantly underperforms the low-beta portfolio. Panel D reports CAPM alphas and betas for the five industry portfolios that are formed using industry classifications. None of the portfolios generate significant alpha over the sample period that we examine.

Fig. 2 summarizes our main finding on the relation between institutional trading and the slope of the SML. The figure plots average excess returns on nine size-BM portfolios as well as the five beta-sorted and five industry portfolios against their estimated betas.[5] The upper graph shows the relation between beta and average returns on High-IT days while the lower graph plots the relation on Low-IT days. The graph for High-IT days shows a strong, positive relation between beta and average returns.

The slope of the relation suggests that an increase in beta by 0.1 is associated with an increase in expected return of approximately 15 basis points (bps) per day (not reported in the graph) with an associated *t*-statistic of 6.35, which is significant at the 1% level. In stark contrast, on Low-IT days the relation is significantly negative, where an increase in beta by 0.1 is associated with a significant reduction in average daily excess returns of approximately 13 bps

[5] All graphs plot the full-sample portfolio beta against average returns. Beta of each test portfolio in the graph is estimated from the full-sample regression of excess portfolio returns on excess market returns. This approach is similar to Savor and Wilson (2014).

Table 2
Daily excess returns on days of high and low institutional trading.

	Fama–MacBeth regression			Pooled regression				
Type of day	Intercept	Beta	Avg. R^2	Intercept	Beta	High	Beta*High	R^2
Panel A: five beta-sorted portfolios								
High	0.000086	0.000833	0.18	0.000429	−0.000851	0.000127	**0.002515**	0.002
	(0.54)	(3.14)		(1.20)	(−1.28)	(0.24)	**(2.49)**	
Low	0.000358	−0.000915	0.23					
	(1.91)	(−2.62)						
High – Low	−0.000272	**0.001748**						
	(−1.15)	**(4.01)**						
Panel B: five beta-sorted and five industry portfolios								
High	0.000075	0.000825	0.15	0.000520	−0.001392	−0.000035	**0.003108**	0.003
	(0.46)	(3.12)		(1.88)	(−3.80)	(−0.09)	**(5.75)**	
Low	0.00027	−0.000862	0.18					
	(0.94)	(−2.04)						
High – Low	−0.00019	**0.001687**						
	(−0.14)	**(3.57)**						
Panel C: five beta-sorted, nine size and BM portfolios, and five industry portfolios								
High	0.000169	0.000740	0.11	0.000394	−0.001286	0.000238	**0.002792**	0.003
	(1.21)	(3.01)		(2.03)	(−4.70)	(0.83)	**(6.89)**	
Low	0.000216	−0.000839	0.14					
	(0.96)	(−2.36)						
High – Low	−0.00005	**0.001579**						
	(−0.20)	**(3.69)**						

Note: This table reports estimates from Fama–MacBeth regressions of daily excess returns on betas for various test portfolios. Estimates are computed for days with high institutional trading (High-IT days or High) and other days (Low-IT days or Low). Day t has High-IT volume (scaled by total market volume) when the IT volume at t is greater than the average IT volume over the past quarter. The difference in the coefficient between High- and Low-IT days is reported in the last row of each panel. There are 1598 days with High-IT volume. The right-hand side panel reports estimates from the pooled regression of excess returns on betas, High-IT day dummy, and interaction between beta and High-IT (Beta*High). Panel A shows results for five beta-sorted (value-weighted) portfolios. Panel B presents results for five beta-sorted and five value-weighted industry portfolios. Panel C reports results for nine value-weighted size-BM, five beta-sorted, and five industry portfolios. The sample period is between 1997 and 2011 (we lost one year to form portfolios). t-statistics, which are computed using Newey–West standard errors with five lags, are reported in parentheses. For pooled regressions, standard errors are corrected for clustering by trading day.

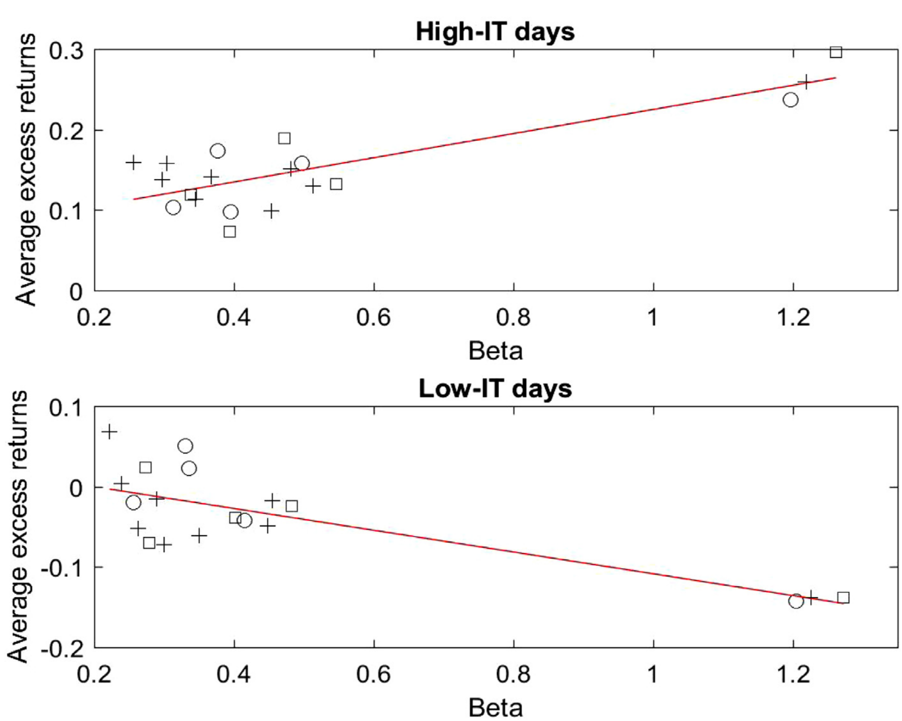


Fig. 2. Capital asset pricing model on high and low institutional trading. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ○ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol). Day t has High-IT volume (scaled by total market volume) when the IT volume at t is greater than the average IT volume over the past quarter. The first graph shows the relation between average returns and beta on High-IT days, while the second graph is constructed on days with Low-IT. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns. The implied ordinary least squares estimates of the securities market line for each type of day are also plotted. The sample period is between 1997 and 2011.

per day. This graph represents our novel findings that beta risk is important when institutions trade intensively.

Table 2 reports results for regressions (1) and (2) using the Fama–MacBeth procedure. Specifically, panel A shows the results for five beta-sorted portfolios. On High-IT days, the slope $\gamma_{1,t}^{H}$ is positive 8.3 bps per day with an associated t-statistic of 3.1, which is significant at the 1% level. This positive slope is consistent with the prediction of the CAPM and that beta risk is priced on High-IT days. In contrast, the slope $\gamma_{1,t}^{L}$ on Low-IT days is −9.2 bps (t-statistic = −2.6). The difference in the slope between High- and Low-IT days (the bottom row) is 17.5 bps per day, which is significant at the 1% level.

The right-hand side of panel A presents the results for the pooled regression (3), which tests the difference in the market risk premium between High- and Low-IT days. Consistent with the Fama–MacBeth results, the coefficient on the interaction between beta and High-IT days is positive and significant (25 bps, with a t-statistic of 2.5). Again, these results show that beta is a more important systematic risk factor on High-IT days. It is also interesting to note that the coefficient on the High-IT dummy is insignificant, suggesting that the difference in returns between High- and Low-IT days is attributable to their portfolio betas, which carry a higher risk premium on High-IT days.

As suggested by Lewellen et al. (2010), it is important to examine how the model performs when more test portfolios are used. Consequently, panel B adds five industry portfolios to the test and brings the total of test portfolios to 10. Consistently, we observe that High-IT days exhibit a strong, positive relation between beta and average returns. The coefficient on beta is positive 8.3 bps, which is significant at the 1% level. On the other hand, this coefficient is −8.6 bps on Low-IT days. The difference between the two days is 16.9 bps per day with an associated t-statistic of 3.6. Further, the pooled regression shows a positive coefficient of 31 bps (t-statistic = 5.8) on the interaction term (*beta* × *High*). These results suggest that adding five industry portfolios to the set of test assets does not change our conclusions that beta risk is more important on High-IT days. Panel C further raises the hurdle by including nine size-BM portfolios as additional test assets. Consistently, the Fama–MacBeth regression results show that the slope coefficient on beta is positive 7.4 bps on High-IT days (t-statistic = 3.0). The difference in the coefficient on beta between the High- and Low-IT samples is 15.8 bps with an associated t-statistic of 3.7, which is significant at the 1% level.[6]

3.2. Excess returns on individual stocks

In this section, we employ individual stocks as test assets and examine whether the market risk premium is higher on High-IT days than on Low-IT days. We report the results in Table 3.

Panel A shows that, on High-IT days, the relation between beta and returns is still positive 3.9 bps with an associated t-statistic of 1.8. In contrast, on Low-IT days, the slope coefficient on beta is negative 7.8 bps (t-statistic = −3.2), which is significant at the 1% level. The difference in the slope between High- and Low-IT days (last row) is positive 11.7 bps and significant at the 1% level. In panel B, the coefficient on the interaction term (*beta* × *High*) is also significantly positive. Thus, our conclusion that the market risk premium is higher on High-IT days than Low-IT days holds for individual stocks.

In panel C, we include log(size), log(BM), and past one-year returns as additional controls to the baseline model. On High-IT days, the coefficient on beta remains positive while on Low-IT days, the coefficient is significantly negative. Again, the difference in the slope on beta between two samples is positive and statistically significant. The negative coefficient on BM is consistent with Chordia et al. (2016), who examine the relation between trades and seasonality effects in the U.S. markets. Panel D shows that including size, BM, and past returns in the pooled regression does not affect the positive coefficient on the interaction term (*beta* × *High*).

A confounding factor that may affect the trade of financial institutions is liquidity. As institutions tend to hold liquid stocks that have more efficient prices, the significant difference in betas on High- and Low-IT days may be affected by a liquidity effect: Low-IT days may represent low liquidity days and therefore stock prices are less efficient. However, this effect is unlikely to drive our results. First, Boehmer and Kelly (2009) show that institutions' trades enhance market efficiency beyond the effect of liquidity provision itself. Second, the effect of liquidity would be larger for selling than for buying activity of institutions. As we show in the next section, our results are driven by the buying rather than the selling activity of institutions. Nevertheless, we control for liquidity (using turnover as a proxy) in panels E and F of Table 3. Turnover is defined as the average daily trading volume over the past year divided by the number of shares outstanding. On each day t, we rank stocks based on their turnover and use the rank (*Turn*) as an additional control in both Fama–MacBeth (panel E) and pooled regression (panel F).[7] The coefficient on *Turn* is positive but insignificant. On High-IT days, the coefficient on beta is 4 bps (t-statistic = 1.9) whereas on Low-IT days, this coefficient is −7.7 bps (t-statistic = −2.8). This leads to the average difference in the slope between High- and Low-IT days of 12 bps (t-statistic = 3.1). The pooled regression in panel F shows consistent results.

Panels G and H repeat the regressions in panels E and F, but exclude Nokia. Over our sample period, Nokia is by far the most liquid and largest firm by market capitalization (approximately 38% of the total market capitalization). As seen, our findings are not driven by Nokia as beta risk remains significant on High-IT days.

3.3. Turn-of-month and January effects

Institutions are generally known to display window-dressing behavior (Sikes, 2014). Hence, we examine whether our results are due to their trading regularity throughout the calendar year by removing days that are turn-of-month from the sample. Fig. 3 contrasts the SML between High- and Low-IT days conditioned on non-turn-of-month days. High-IT days still exhibit the positive relation between beta and average returns while Low-IT days show a negative relation. On High-IT days, an increase in beta by 0.1 is associated with a rise in average returns of 16 bps per day (t-statistic = 6.8, not reported in the graph). On Low-IT days, an increase in beta by 0.1 is associated with a decrease in average returns of 14 bps per day (t-statistic = −5.5). These results suggest that the turn-of-month behavior of institutions is not the explanation.

To complete the analysis, Fig. 4 controls for the well-known January effect (Sias and Starks, 1997) by excluding the month of January from the sample. Outside January, High-IT days still show strong support for the CAPM while Low-IT days indicate the CAPM's failure in explaining average returns. On High-IT days, an increase in beta by 0.1 is associated with a rise in average returns

[6] We confirm that our main results still hold when we employ *lagged* High-IT and control for the market returns in regression (3), when we conduct our analysis based on the residual IT, after controlling for market returns and market direction, and when we form test portfolios based on idiosyncratic volatility portfolios (as in Savor and Wilson, 2014). These additional tests are available on request.

[7] We also use the number of occurrences of zero returns over the past year as an alternate proxy for the level of liquidity (see Lesmond et al., 1999). Griffin et al. (2010) show that, in smaller and emerging markets, this measure captures the illiquidity and transaction costs better than other measures. Using this measure as a proxy for illiquidity does not alter our findings.

Table 3
Daily excess returns for individual stocks on days of high and low institutional trading.

Panel A: beta only (Fama–MacBeth regressions)

Type of day	Beta	Avg. R^2
High	0.00039	0.02
	(1.80)	
Low	−0.00078	0.02
	(−3.20)	
High – Low	**0.00117**	
	(3.52)	

Panel B: beta and interaction with High-IT (pooled regressions)

	Beta	High	Beta*High	R^2
	−0.001335	0.000637	**0.002334**	0.001
	(−7.49)	(3.54)	**(8.77)**	

Panel C: size, BM, and past returns as controls (Fama–MacBeth regressions)

Type of day	Beta	Size	BM	Past one-year	R^2
High	0.00044	−0.00008	−0.00009	−0.02032	0.04
	(1.81)	(−2.59)	(−3.13)	(−1.18)	
Low	−0.00077	−0.00007	−0.00007	0.00026	0.05
	(−2.85)	(−2.29)	(−2.19)	(1.43)	
High – Low	**0.001205**	−0.00001	−0.00002	−0.02058	
	(3.21)	(−0.02)	(−0.40)	(−1.76)	

Panel D: size, BM, and past returns as controls (pooled regressions)

	Beta	Size	BM	Past one-year	High	Beta*High	R^2
	−0.000790	−0.000256	0.000072	−0.000241	0.000338	**0.001862**	0.001
	(−4.90)	(−7.28)	(5.79)	(−1.39)	(1.23)	**(4.60)**	

Panel E: size, BM, past returns, and liquidity as controls (Fama–MacBeth regressions)

Type of day	Beta	Size	BM	Past one-year	Turn	Avg. R^2
High	0.00041	−0.00008	−0.00009	−0.01801	0.00413	0.04
	(1.88)	(−2.85)	(−3.15)	(−1.05)	(0.76)	
Low	−0.00077	−0.00007	−0.00007	0.00028	−0.00001	0.06
	(−2.82)	(−2.21)	(−2.19)	(1.51)	−0.20	
High – Low	**0.00118**	−0.00001	−0.00002	−0.01829	0.00415	
	(3.14)	(−0.28)	(−0.38)	(−1.73)	(0.58)	

Panel F: size, BM, past returns, and liquidity as controls (pooled regressions)

	Beta	Size	BM	Past one-year	Turn	High	Beta*High	R^2
	−0.000851	−0.000293	0.000071	−0.000240	0.00007	0.000336	**0.001869**	0.001
	(−5.21)	(−7.55)	(5.72)	(−1.39)	(1.06)	(1.12)	**(4.62)**	

Panel G: size, BM, past returns, and liquidity as controls (Fama–MacBeth regressions) – excluding Nokia

Type of day	Beta	Size	BM	Past one-year	Turn	Avg. R^2
High	0.00039	−0.00009	−0.00010	−0.01746	0.00419	0.04
	(1.78)	(−2.96)	(−3.26)	(−1.02)	(0.77)	
Low	−0.00078	−0.00007	−0.00008	0.00028	−0.00001	0.06
	(−2.79)	(−2.24)	(−2.32)	(1.53)	(−0.08)	
High – Low	**0.001171**	−0.00002	−0.00002	−0.01773	0.00419	
	(3.10)	(−0.32)	(−0.31)	(−1.69)	(0.50)	

Panel H: size, BM, past returns, and liquidity as controls (pooled regressions) – excluding Nokia

	Beta	Size	BM	Past one-year	Turn	High	Beta*High	R^2
	−0.000843	−0.000293	0.000066	−0.000242	0.00008	0.000365	**0.001787**	0.001
	(−5.11)	(−7.51)	(4.79)	(−1.40)	(1.09)	(1.32)	**(4.34)**	

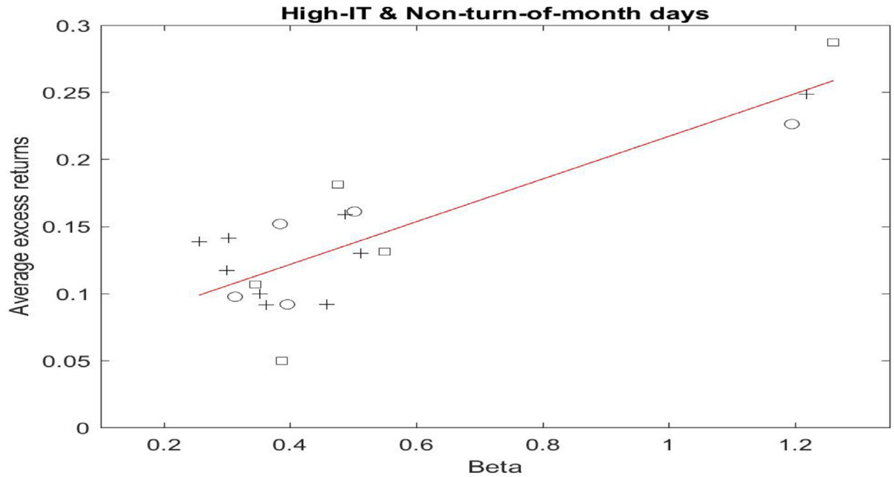
Note: This table reports estimates from Fama–MacBeth regressions of daily excess returns on betas for various test portfolios. Estimates are computed for days with high institutional trading (High-IT days or High) and other days (Low-IT days or Low). Day t has High-IT volume (scaled by total market volume) when the IT volume at t is greater than the average IT volume over the past quarter. The difference in the coefficient between High- and Low-IT days is reported in the last row of each panel. There are 1598 days with High-IT volume. Betas of individual stocks are estimated daily using rolling regressions of daily excess returns on excess market returns over the past one year, corrected for potential asynchronous trading using the Dimson method. Panels A and B show results for beta as the right-hand side variable in Fama–MacBeth regressions and pooled regressions, respectively. Panels C and D are similar to the first two panels, but include log(*size*), BM, and past one-year returns as additional controls. Panels E and F include liquidity (Turn) as an additional control. "Turn" is the rank of turnover, defined as the average daily volume over the past year divided by the number of shares outstanding. The sample period is between 1997 and 2011 (we lost one year to form portfolios). Panels G and H repeat the regressions in Panels E and F, but excluding Nokia, which is the largest firm on the Finnish market. t-statistics, which are computed using Newey–West standard errors with five lags, are reported in parentheses.

of 17 bps per day (t-statistic = 6.9). On Low-IT days, an increase in beta by 0.1 is associated with a decrease in average returns of 12 bps per day (t-statistic = −4.4). We thus conclude that regularities in stock returns throughout the calendar year cannot explain our findings.

3.4. Subsample analysis: pre- and post-2008

The summary statistics in Table 1 show that there is an increase in trading volume of institutions over the period between 2008 and 2011 that covers the Global Financial Crisis. Consequently, this subsection examines whether our results are attributable to this later period by looking at two subsamples: pre- and post-2008.

Table 4 reports results of Fama–MacBeth and pooled regressions that are estimated using samples before (panel A) and after 2008 (panel B). As before, estimates are reported separately for High- and Low-IT days and test assets are five beta-sorted, nine size-BM, and five industry-sorted portfolios. The general conclusion is that the difference in market risk premium between High- and Low-IT days remains positive and statistically significant in all subsamples. Before 2008, the slope of the SML is positive 15.4 bps on High-IT days with an associated t-statistic of 2.3. On Low-IT days, the slope



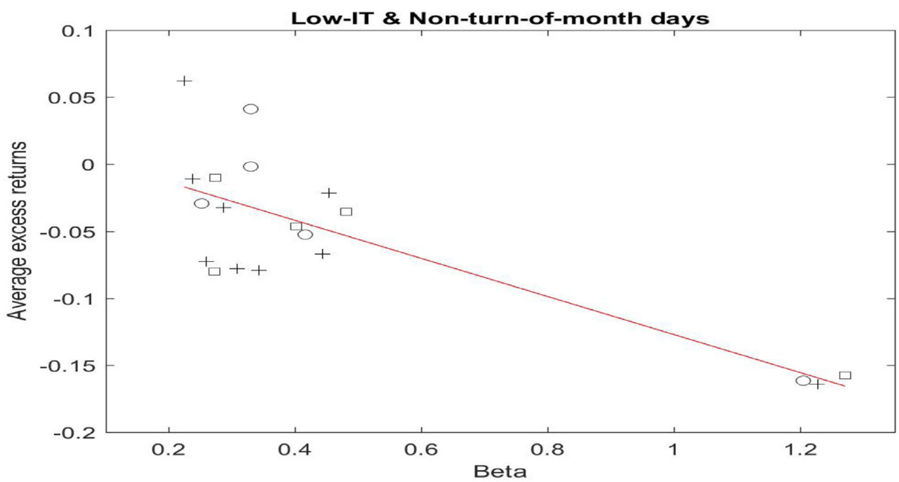


Fig. 3. Capital asset pricing model on non-turn-of-month days. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ◯ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol) conditioned on the IT effect and non-turn-of-month days. The first graph shows the relation between average returns and beta on days with High-IT but not the turn-of-month day, while the second graph presents the relation on days with Low-IT and not the turn-of-month day. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns. The implied ordinary least squares estimates of the securities market line for each type of day are also plotted. The sample period is between 1997 and 2011.

Table 4
Daily excess returns on High- and Low-IT volume: sub-sample analysis.

	Fama–MacBeth regression			Pooled regression				
Type of day	Intercept	Beta	Avg. R^2	Intercept	Beta	High	Beta*High	R^2
Panel A: five beta-sorted, nine size-BM portfolios, and five industry portfolios, pre-2008								
High	0.000453	0.001541	0.25	0.000542	−0.000610	0.000135	**0.002160**	0.002
	(1.22)	(2.26)		(2.50)	(−1.80)	(0.43)	**(4.61)**	
Low	0.000351	−0.000658	0.14					
	(1.23)	(−1.42)						
High – Low	0.000102	**0.00220**						
	(0.25)	**(2.82)**						
Panel B: five beta-sorted, nine size-BM portfolios, and five industry portfolios, post-2008								
High	0.00018	0.00186	0.25	−0.000306	−0.001803	0.000702	**0.003373**	0.007
	(0.317)	(2.06)		(−0.68)	(−3.10)	(1.04)	**(3.80)**	
Low	0.000159	0.001932	0.23					
	(−0.30)	(−2.57)						
High – Low	0.00026	**0.003196**						
	(0.43)	**(2.95)**						

Note: This table reports estimates from Fama–MacBeth regressions of daily excess returns on betas for various test portfolios in two subsamples: before and after 2008. Day *t* has High-IT volume (scaled by total market volume) when the IT volume at *t* is greater than the average IT volume over the past quarter. The difference in the coefficients between High- and Low-IT days is reported in the last row of each panel. The right-hand side panel reports estimates from pooled regression of excess returns on betas, High-IT day dummy, and interaction between beta and High-IT (Beta*High). Test assets are five beta-sorted, nine value-weighted size-BM and five industry portfolios. The sample period is between 1997 and 2011. Panel A uses sample before 2008 while the data in Panel B are after 2008. *t*-statistics, which are computed using Newey–West standard errors with five lags, are reported in parentheses. For pooled regressions, standard errors are clustered by trading day.

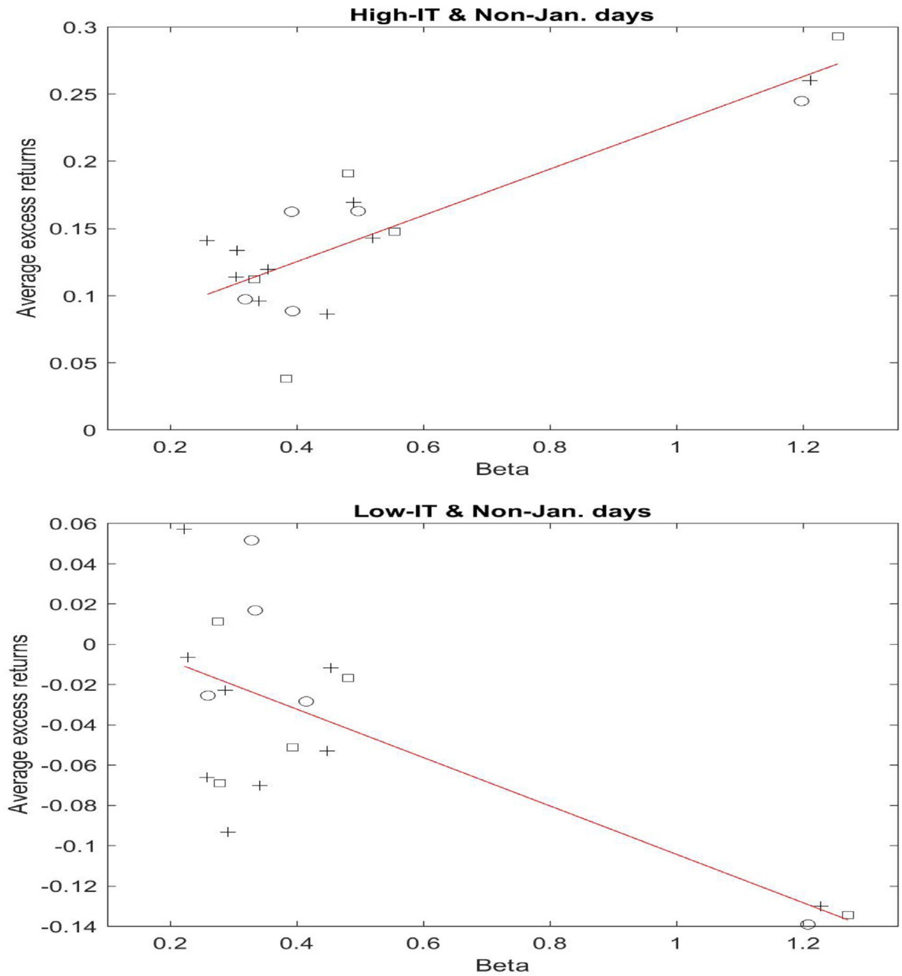


Fig. 4. Capital asset pricing model on non-January days. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ◯ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol) conditioned on the IT effect and non-January days. The first graph shows the relation between average returns and beta on days with High-IT volume but not in January, while the second graph presents the relation on days with Low-IT volume and not January days. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns. The implied ordinary least squares estimates of the securities market line for each type of day are also plotted. The sample period is between 1997 and 2011.

is −6.6 bps although this coefficient is insignificant. The difference in betas between the two samples is 22 bps with a significant t-statistic of 2.8. The pooled regression also shows that the coefficient on the interaction term is positive 21.6 bps (t-statistic = 4.6).

We see a similar picture in panel B for the period between 2008 and 2011. The relation between beta and average returns is still strong and positive on High-IT days while this relation is significantly negative on Low-IT days. The pooled regression also shows that the market risk premium is much higher on High-IT days than on Low-IT days. Consequently, the IT effect is not specific to a subsample period.

3.5. Out-of-sample test: the use of U.S. 13F holdings data

As an out-of-sample test, we employ Thomson Reuters 13F quarterly holdings data for the U.S. markets and show that the High-IT effect is not specific to the Finnish market. In each quarter, we compute the aggregate change in holdings of all 13F institutions scaled by the total trading volume between 1980 and 2014. Similar to the main analysis, we define a quarter as High-IT when the fraction of institutional volume over total market volume is higher than its average over the past year. The High-IT dummy is then a monthly time series dummy variable that takes the value of one in months of High-IT quarters. We collect monthly market data for all U.S. common equities (share codes of 10 and 11) from the Center for Research in Security Prices (CRSP). We estimate pre-ranking betas by regressing 60 months of excess returns on excess market returns (adjusting for potential non-synchronous trading using Dimson's sum beta). We then form 10 value-weighted portfolios on the basis of individual stocks' monthly betas. We repeat the Fama–MacBeth and pooled regressions and report the results in Table 5. The primary limitation of 13F holdings data is its low frequency. Since the quarterly frequency of 13F data cannot capture the timely trade of institutions, the power of our tests is significantly reduced, and inferences would have to rely on the sign and magnitude of the coefficient rather than its statistical significance.

Panel A of Table 5 reports the estimate of Fama–MacBeth regressions. On High-IT quarters, the coefficient on beta is 28 bps whereas, on Low-IT quarters, this coefficient is −35 bps. The difference in betas is 63 bps (t-statistic = 1.8). The pooled regression shows a similar picture that the coefficient on the interaction term ($beta \times High$) is positive 51 bps with an associated t-statistic of 2.4. Compared with the Finnish results in Table 2, the magnitude of these coefficients is higher, and hence economically significant. These findings suggest that the relation between beta and average returns is also pronounced on High-IT quarters in the U.S. markets.

Table 5
U.S. markets' results: excess returns on quarters of high and low institutional holdings.

Type of day	Intercept	Beta	Avg. R^2	
Panel A: Fama–MacBeth regression				
High	0.00040	0.002791	0.25	
	(0.20)	1.040		
Low	0.006491	−0.003520	0.27	
	(3.56)	−1.19		
High − Low	−0.00609	**0.00631**		
	(−2.55)	**(1.80)**		
Panel B: pooled regression				
Intercept	Beta	High	Beta*High	R^2
0.006599	−0.003290	−0.004729	**0.005102**	0.002
(4.01)	(−2.26)	(−1.98)	**(2.41)**	

Note: This table reports estimates from Fama–MacBeth regressions of daily excess returns on betas for ten beta-sorted portfolios using U.S. CRSP data. Estimates are computed for quarters with high institutional trading (High-IT or High) and other quarters (Low-IT or Low). Using data from Thomson Reuters SEC 13F holdings data, we compute aggregate change in institutional holdings scaled by CRSP's total market trading volume per quarter. Quarter *t* has Low-IT volume (scaled by total market volume) when the aggregate change in quarterly holdings at *t* is greater than the average IT volume over the past year (four quarters). The difference in the coefficient between High- and Low-IT quarters is reported in the last row of each panel. Panel B reports estimates from pooled regression of excess returns on betas, High-IT quarter dummy, and interaction between beta and High-IT (Beta*High). The sample period is between Q1-1984 and Q1-2014. *t*-statistics, which are computed using Newey–West standard errors with five lags, are reported in parentheses. For the pooled regression, standard errors are clustered by trading day.

These out-of-sample tests confirm our main conclusions and show that the High-IT effect is not specific to the Finnish market. The results mitigate any potential concerns of small markets and suggest that the institutional settings of the Finnish market do not seem to be driving our findings.

4. Investor-type characteristics

The findings thus far are robust and demonstrate the important role of institutions in shaping the relation between beta and expected returns. In this section, we explore several hypotheses that may explain the role of institutions. Existing theory suggests three potential channels based on institutions' characteristics and preferences. The first channel highlights the role of institutions in asset pricing and suggests that larger institutions play a more important role in shaping the risk premium dynamics of the assets they trade (Siriwardane, 2016). The second channel focuses on the literature on investor disagreement (Hong and Sraer, 2016), while the third channel focuses on preferences for stocks with lottery-like characteristics (Mitton and Vorkink, 2007).

Regarding the second and third channels, there may be a confounding effect that it is the market-wide disagreement and lottery preferences that drive the relation between beta and average returns. In the analysis below, we rule out this possibility, but leave room for an explanation based on investor disagreement and lottery preference being investor-type characteristics. While the three potential explanations originate from seemingly different theories, they are not necessarily mutually exclusive. Overall, our findings are supportive of the bigger central theme in that financial institutions are the marginal investors, who shape the positive risk-return relation when they trade.

4.1. Investor trading activity

The first potential explanation for our findings is motivated by the financial intermediary-based asset pricing theory (He and Krishnamurthy, 2013; Adrian et al., 2014). In particular, Adrian et al. (2014) argue that between two prominent types of investors, households and financial institutions, the latter tends to fit the assumptions of modern finance theory and, if they are the marginal investors in the market, their trading activity can affect stock prices. Consequently, our finding that the risk and return relation holds when institutions dominate the market is by itself consistent with this financial intermediary hypothesis.

The financial intermediary hypothesis further posits that the heterogeneity of wealth among financial institutions is also important in driving asset prices (He and Krishnamurthy, 2013; Siriwardane, 2016). In particular, Siriwardane (2016) shows that it is the funding frictions within financial institutions that drive their trading activity and in turn, determine CDS pricing. As a result, the most active institutions (who face lower internal capital frictions) will determine CDS prices at a given point in time. Under this hypothesis, we also expect that it is chiefly the most active institutions that shape the positive relation between beta and average returns. Since we have data on each individual institution over time, we can indeed provide formal tests of this prediction.[8]

On each day *t*, we group institutions into quintiles based on their trading volume on day $t-1$, where quintile 1 contains firms with the lowest trading activity and quintile 5 comprises those with the highest trading activity. For each quintile, High-IT days are defined as follows: when the total volume in each quintile group (scaled by market volume) is greater than the average of its own quintile over the past quarter. We then estimate the beta-return relation for High- versus Low-IT days for each quintile and report the results in Table 6.

In Table 6, the Fama–MacBeth regressions show that the slope of the SML is significant only for the High- versus Low-IT variable based on the very active institutions (quintiles 4 and 5). Likewise, the pooled regression analysis confirms that it is the more active institutions that drive the relation between beta and average returns. On the whole, these results lend support to the explanation based on intermediary-based asset pricing.

4.2. Investor disagreement

Hong and Sraer (2016) contend that high-beta stocks are more sensitive to market-wide disagreement about the future of the economy. Consequently, in times of high disagreement and together with the presence of limits to arbitrage, the optimists outweigh the pessimists in their trading impact on high-beta stocks. When disagreement is large, the expected return on high beta stocks exhibits an inverted U-shape curve. When disagreement is low, the SML will be upward sloping.

Empirically, testing this hypothesis requires choosing a proxy for disagreement. A traditional measure of disagreement is the dispersion of analyst forecasts of earnings-per-share. However, in the context of our study, this measure is not feasible because of its low frequency and little variation over time. Consequently, we employ an alternative measure of aggregate news tone dispersion that is similar in spirit to Dzieliński and Hasseltoft (2017) as a proxy of investor disagreement. As Dzieliński and Hasseltoft (2017) show, this news dispersion is strongly associated with, and can even drive, analyst disagreement.

We collect news data from Thomson Reuters News Analytics (TRNA) that systematically quantifies the tone of firm-specific news for Finnish stocks between 2003 and 2011 (2267 trading days). This database has recently received considerable attention in the literature. For example, Hendershott et al. (2015) employ TRNA to ex-

[8] We thank the referee for suggesting this test. Our data set does not contain the size of institutions as all trader identities are removed in the data. Hence as a proxy for the size of institutions, we focus on their trading activity.

Table 6
Daily excess returns on days of high and low institutional trading split by volume.

	Fama–MacBeth regression			Pooled regression				
Type of day	Intercept	Beta	Avg. R^2	Intercept	Beta	High	Beta*High	R^2
Panel A: quintile 1 of institutional trading volume								
High	0.000170	−0.000019	0.088	0.000689	−0.000370	−0.000435	**0.000625**	0.0001
	(1.16)	(−0.07)		(3.88)	(−1.38)	(−1.43)	**(1.36)**	
Low	0.000198	−0.000105	0.17					
	(0.88)	(−0.28)						
High – Low	−0.000028	**0.000086**						
	(−0.12)	**(0.20)**						
Panel B: quintile 2 of institutional trading volume								
High	0.000275	−0.000199	0.10	0.000452	0.000049	0.000232	−**0.000532**	0.0001
	(1.89)	(−0.75)		(2.47)	(0.18)	(0.78)	(−**1.19**)	
Low	0.000093	0.000075	0.15					
	(0.41)	(0.22)						
High – Low	0.000181	−**0.000275**						
	(0.77)	(−**0.63**)						
Panel C: quintile 3 of institutional trading volume								
High	0.000306	0.000027	0.11	0.000454	−0.000427	0.000210	**0.000625**	0.0002
	(1.65)	(0.09)		(2.40)	(−1.49)	(0.72)	**(1.42)**	
Low	0.000062	−0.000151	0.15					
	(0.36)	(−0.49)						
High – Low	0.000244	**0.000179**						
	(1.03)	**(0.41)**						
Panel D: quintile 4 of institutional trading volume								
High	0.000161	0.000290	0.12	0.000556	−0.000732	−0.000032	**0.001270**	0.0004
	(0.92)	(0.96)		(2.87)	(−2.49)	(−0.11)	**(2.90)**	
Low	0.000206	−0.000414	0.14					
	(1.16)	(−1.41)						
High – Low	−0.000045	**0.000705**						
	(−0.19)	**(1.73)**						
Panel E: quintile 5 of institutional trading volume								
High	0.000262	0.000486	0.13	0.000072	−0.000822	0.000877	**0.001345**	0.0020
	(1.61)	(1.68)		(0.34)	(−2.59)	(3.04)	**(3.09)**	
Low	0.000106	−0.000610	0.13					
	(0.50)	(−1.89)						
High – Low	0.000156	**0.001097**						
	(0.66)	**(2.53)**						

Note: This table reports results from regressions that replicate the main results (Table 2) in five sub-samples split based on the trading volume of financial institutions. Specifically, on each day t, institutions are ranked and grouped into quintiles based on their trading volume on day $t-1$, where quintile 1 contains institutions with the lowest volume while quintile 5 consists of those with the highest volume. For each quintile, High-IT days are defined as usual: when the total volume in the quintile group (scaled by market volume) is greater than its average over the past quarter. Test assets are five beta-sorted, nine value-weighted size-BM, and five industry portfolios. Panel A reports the results for the institutions with the lowest trading activity, whereas panel E reports the results for the institutions with the highest trading activity. The panels in between report the results for the trading activity quintiles that sit in between.

amine the relation between news and stock returns or institutional trading behavior. For every news item, TRNA provides the probability that a news item has good, neutral, or negative tone (the three scores sum to one). For each news article, we compute a unified tone score as the difference between good and bad score, and compute news tone dispersion as the standard deviation of the tone score of all news articles on that day. We define a day to have high news dispersion if the day's tone dispersion is greater than the average dispersion over the past month. As TRNA is reasonably comprehensive, approximately 90% of the trading days have news and, on average, there are 52 news articles per day. The average news dispersion is 0.48 per day and 1260 days are defined as having high news dispersion (strong investor disagreement).

Fig. 5 plots the SML on weak- and strong-disagreement days. On weak disagreement days, the relation between beta and average returns is positive. An increase in beta by 0.1 is associated with a 4 bps increase in average returns (not reported in the graph). In contrast, on days with strong disagreement, the SML is downward sloping: an increase in beta by 0.1 is associated with a 14 bps lower average return per day. In short, using the new proxy for investor disagreement, we can confirm the theoretical prediction of Hong and Sraer (2016).

To formally control for the effect of market-wide disagreement, we run the following regression for the 19 test portfolios (five beta, nine size-BM, and five industry portfolios):

$$R_{i,t+1} = \gamma_0 + \gamma_1 \hat{\beta}_{i,t} + \gamma_2 High_{t+1} + \gamma_3 \hat{\beta}_{i,t} High_{t+1} + \gamma_4 Control_{t+1} + \gamma_5 \hat{\beta}_{i,t} Control_{t+1} + u_{i,t+1} \quad (4)$$

where *Control* takes a value of one if disagreement in the market is strong on a day and zero otherwise. The coefficient of interest is γ_3, which should be significantly positive if the IT effect cannot be explained by the disagreement effect.

Panel A of Table 7 shows that the coefficient on $beta \times MktDisagree$ is negative, confirming that the risk premium is lower on days when disagreement is high. The coefficient on $beta \times High$ is still positive and significant, suggesting that, even on days when disagreement is strong, the risk premium is still higher when there is high institutional trading activity. Thus, market-wide disagreement cannot explain our findings.

Our original hypothesis suggests that disagreement could be an investor-type characteristic. Compared to individual investors, institutions tend to have lower disagreement among each other because they tend to use similar trading models, employ security analysts who have similar training in business schools and are ex-

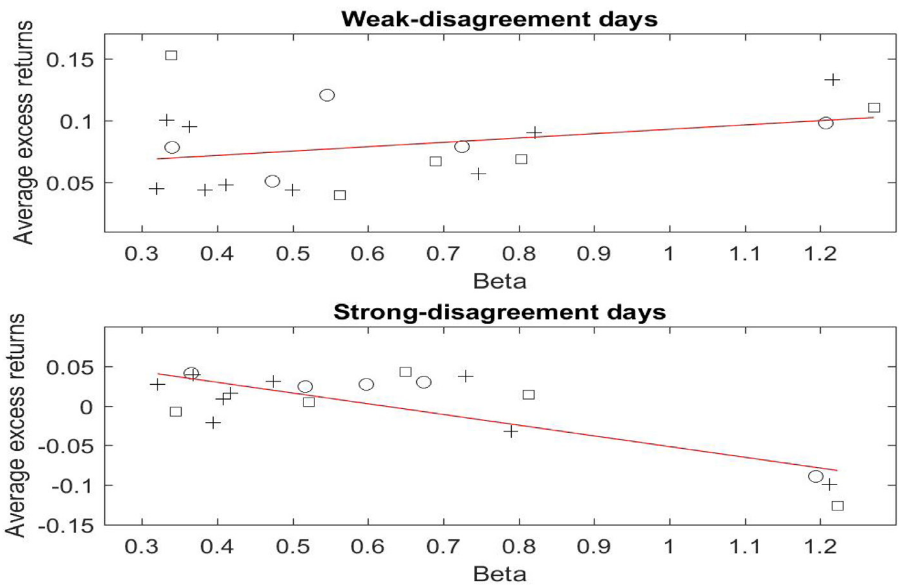


Fig. 5. Capital asset pricing model on days with high or low degree of investor disagreement. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ○ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol) on high and low-disagreement days in the market. A day t has strong investor disagreement if the cross-sectional standard deviation of news tone on day t is higher than its average over the past month. Dzieliński and Hasseltoft (2017) show that high news dispersion is associated with strong investor disagreement. The first graph shows the relation between average returns and beta on days with weak investor disagreement, while the second graph presents the relation on days with strong investor disagreement. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns. The implied ordinary least squares estimates of the securities market line for each type of day are also plotted. The sample period is daily frequency between 2003 and 2011 (the news data set starts in 2003).

ante more likely to be trading based on traditional analysis of risk and return. If so, our findings are consistent with the notion that High-IT days exhibit an upward sloping SML because there is less disagreement among institutions. When their trading activity dominates the market, the investor disagreement hypothesis predicts a positive relation between beta and returns. In fact, further evidence in panel B of Table 7 shows that investor disagreement is lower on days when institutions trade intensively. Specifically, panel B compares average disagreement between High- and Low-IT days. We observe that disagreement is lower on High-IT than on Low-IT days and the difference is significant. These results suggest that the low disagreement among institutional investors compared to household investors is an important characteristic that helps explain the difference of risk premiums between High- and Low-IT days.

4.3. Lottery preferences

The last potential explanation for our findings is the lottery preference hypothesis (Mitton and Vorkink, 2007). According to this hypothesis, individual investors are willing to sacrifice returns to hold stocks with lottery features such as high beta and high coskewness. Similar to investor disagreement, lottery preference could be either market-wide or specific to an investor type (i.e., individual investors). In this subsection, we show that market-wide lottery preference cannot explain our findings. Rather, the evidence lends support to our original hypothesis that low lottery preference is a characteristic of institutions, whose presence is more apparent on High-IT days.

We estimate a stock's coskewness (denoted as the coefficient c_i) as in Harvey and Siddique (2000) and Mitton and Vorkink (2007) by running the following quarterly regression:

$$R_{i,t} = a + b_i R_{M,t} + c_i R_{M,t}^2 + \epsilon_{i,t}, \tag{5}$$

where $R_{i,\,t}$ and $R_{M,\,t}$ are excess returns on the stock and the market, respectively. The coefficient c_i is the coskewness of the stock. The market coskewness is then the simple average of individual stocks' coskewness on each day. Day t exhibits high coskewness if the market coskewness is higher than its average over the past quarter.

Fig. 6 shows that on days of low coskewness, the SML is upward sloping while on high-coskewness days, the SML is downward sloping. These results suggest that lottery preference can affect the relation between beta and returns.

Panel A of Table 7 reports results for regression (4) in which we control for days with high market coskewness. Panel A shows that the market risk premium is indeed lower on days when market coskewness is higher. However, we also see that the coefficient on $beta \times High$ is still positive and significant, suggesting that lottery preference in the market cannot explain the IT effect.

Our hypothesis is that lottery preference could be an investor-type characteristic, particularly for individual investors, rather than institutions (Mitton and Vorkink, 2007; Kumar, 2009). If so, we would expect High-IT days to exhibit lower lottery-type characteristics. In addition to coskewness, we employ two additional measures of lottery characteristics, namely total skewness and MAX portfolio. As in Mitton and Vorkink (2007), we compute skewness for each stock using daily returns over the past quarter:

$$skewness = \frac{\frac{1}{90}\sum_{t=1}^{90}(r_t - \mu)^3}{\hat{\sigma}^3}, \tag{6}$$

where μ and σ are average returns and the standard deviation of stock returns over the past quarter, respectively. The market skewness is then the simple average across stocks on each day.[9] The advantage of this calculation is that it captures the incremental skewness over variance because these statistics are positively correlated. We also follow Bali et al. (2011) and employ the MAX portfolio; the more negative (lower) the average MAX return is, the stronger is the preference for skewness.

[9] The conclusions do not qualitatively change if we compute skewness for the market index.

Table 7
Betas, investor disagreement, and lottery preferences.

Panel A: market-wide explanations

Controlling for high market-wide disagreement days and market directions

Intercept	Beta	High	Beta*High	MktDisagree	Beta*MktDisagree	R^2
0.00094	−0.00154	0.000553	**0.00270**	−0.00021	−0.00101	0.0069
(3.32)	(−3.94)	(1.76)	**(6.18)**	(−0.66)	(−2.29)	

Controlling for high market coskewness days and market directions

Intercept	Beta	High	Beta*High	MktCoskew	Beta*MktCoskew	R^2
0.00048	−0.00074	0.00025	**0.00279**	0.00018	−0.00103	0.0034
(1.88)	(−2.07)	(0.88)	**(6.90)**	(0.61)	(−2.54)	

Panel B: institutional investors' characteristics

				News				Difference: Buy–Sell				
IT day	Buy/sell	IT Imb	Beta	Dispersion	Skewness	Coskewness	MAX	Beta	Dispersion	Skewness	Coskewness	MAX
Low	Sell	−0.0504	0.6783	0.4944	−0.0163	−0.5972	−0.0018	0.0136#	0.0001	−0.0015	0.0028	0.0000
	Buy	0.0408	0.6919	0.4945	−0.0178	−0.5945	−0.0019					
High	Sell	−0.0605	0.6453	0.4849	−0.1044	−1.1951	0.0016	0.0114#	0.0003	−0.0009	0.0196	0.0000
	Buy	0.0514	0.6567	0.4852	−0.1053	−1.1754	0.0017					
High - Low	Sell	−0.0101#	−0.0330#	−0.0095#	−0.0881#	−0.5978#	0.0035#					
	Buy	0.0106#	−0.0352#	−0.0094#	−0.0875#	−0.5810#	0.0035#					

Note: This table reports the test results of the investor disagreement hypothesis and lottery preference hypothesis. Panel A tests whether the market-wide disagreement and market-wide lottery preference can explain away the IT effect. The regression specification is Model (4), where Control is either market-wide news dispersion (a proxy for disagreement) or market-wide coskewness (a proxy for lottery preference). Test assets are five beta-sorted, nine value-weighted size-BM, and five industry portfolios. Panel B reports the statistics for IT imbalance, beta, skewness, coskewness, MAX, and news dispersion (a proxy for investor disagreement) on High- and Low-IT days. In panel B, stocks are split into two groups on the basis of their institutional trade imbalance where the first group (Buy) has positive average imbalance and the second group (Sell) has negative average imbalance. "IT imb" is the average imbalance (imb=(buy volume–sell volume)/total market volume) of institutions. "Beta" is the average beta across stocks in each buy or sell portfolios. "Skewness" is the average skewness of the market. "Coskewness" is the average coefficient on c_i in regression (5). "MAX" is the average return on the MAX portfolio that is constructed in a similar way to Bali et al. (2011). Specifically, stocks are ranked and split into five groups on the basis of their average five maximum daily returns over the past month. Portfolio 1 contains stocks with the lowest maximum daily returns while portfolio 5 consists of stocks with the highest maximum daily returns. The value-weighted average raw return difference between portfolio 5 and 1 is the MAX portfolio. Bali et al. (2011) show that this MAX portfolio earns a negative average return because investors have a strong preference for skewness stocks. "News Dispersion" is the average of the cross-sectional standard deviation of news tone. *High – Low* in the bottom panel reports the difference in beta and news dispersion between High- and Low-IT days. The last five columns report the difference in the characteristic between Buy and Sell. # denote the difference that is statistically significant at either the 5% or 1% level. Standard errors are computed using the Newey–West method with five lags.

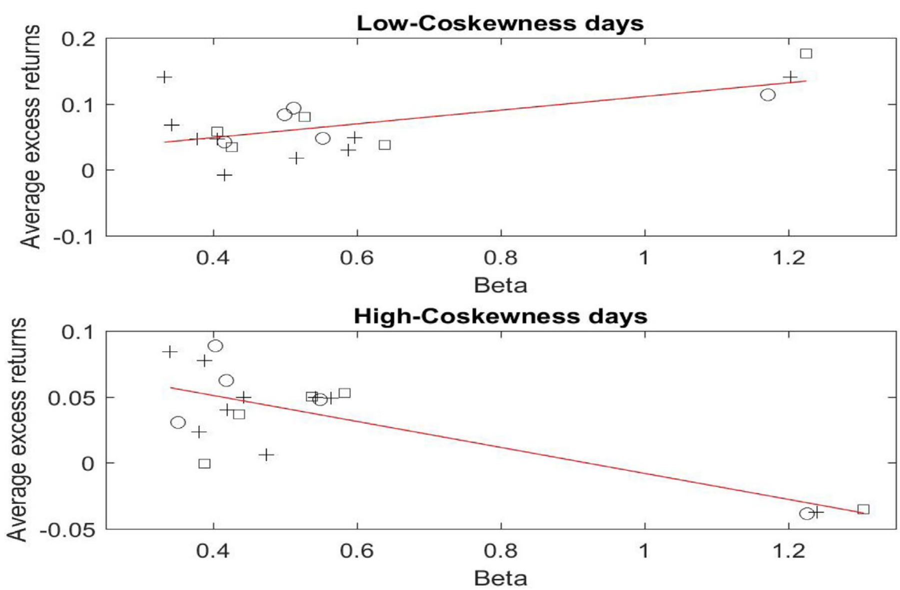


Fig. 6. Capital asset pricing model on days with high or low degree of coskewness. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ◯ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol) on high and low-coskewness days in the market. A stock coskewness is estimated using regression (5) and then the overall market coskewness on day t is the average of individual stocks' coskewness. Day t exhibits high coskewness if the overall coskewness is higher than its average over the past quarter. The first graph shows the relation between average returns and beta on days with low coskewness, while the second graph presents the relation on days with high coskewness. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns. The implied ordinary least squares estimates of the securities market line for each type of day are also plotted.

Panel B of Table 7 compares lottery characteristics between High- and Low-IT days. High-IT days have a lower beta, which is associated with lower skewness, lower coskewness, and higher MAX return. Furthermore, on High-IT days, skewness and coskewness are lower and MAX is higher compared to Low-IT days, and the differences are all significant at the 1% level. This is consistent with the notion that on High-IT days, preferences for lottery-like stocks are lower.

These results provide support for our hypothesis and are consistent with the implication of Mitton and Vorkink's (2007)) model, that low lottery preference is a characteristic of institutions that causes the SML to be upward sloping on High-IT days. In contrast,

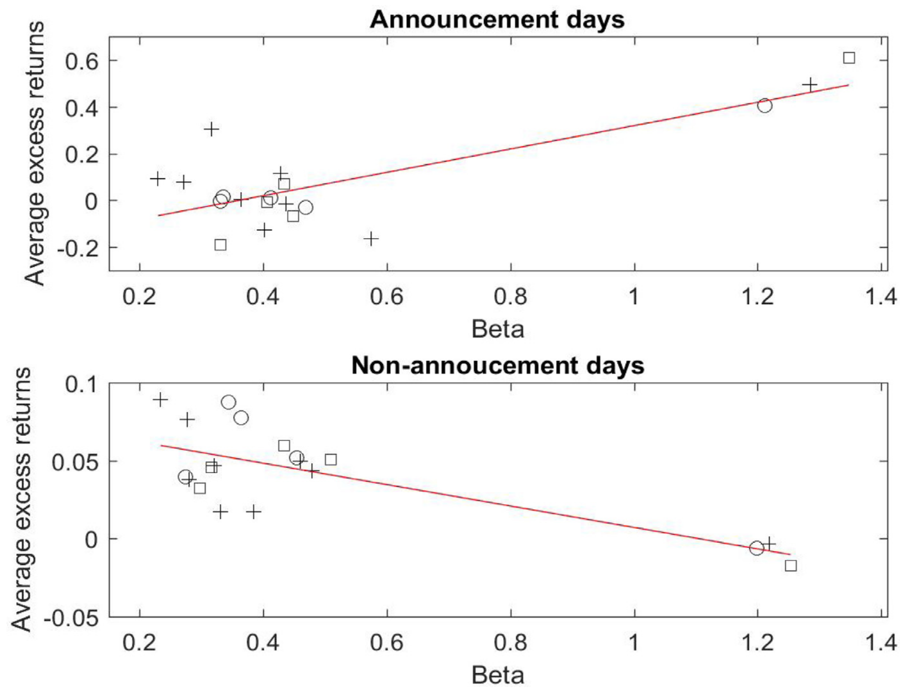


Fig. 7. Capital asset pricing model on announcement and non-announcement days. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ◯ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol) on ECB announcement and non-announcement days. The first graph shows the relation between beta and average returns on days with scheduled ECB monetary policy decision. The second graph presents a similar line on non-announcement days. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns. The implied ordinary least squares estimates of the Security Market Line for each type of day are also plotted. The sample period is between 1999 and 2011 (the ECB was formally established in 1999).

on Low-IT days when individual investors with stronger lottery-preferences are more active in the market, their trading impact on high-beta stocks is stronger, causing the downward sloping SML.

5. Other market-wide explanations

In this section, we test whether other competing hypotheses could explain our findings. Specifically, based on existing theories and empirical findings in the literature, we investigate the role of macroeconomic news announcements, leverage constraints, short-sale constraints, and the effect of up-market states. The general conclusion is that the IT effect cannot be explained by these market-wide explanations.

5.1. Macroeconomic announcements

Savor and Wilson (2014) find that the relation between beta and average returns is strong and positive on days when there is an interest rate announcement. This relation on non-announcement days, in contrast, is downward sloping. Thus, one possible explanation for our findings is that the IT effect could be a manifestation of the announcement effect.

Fig. 7 plots the SML on ECB scheduled monetary policy announcement days and non-announcement days between 1999 and 2011. Similar to Savor and Wilson (2014), we find a positive relation between beta and average returns on announcement days. Non-announcement days, however, exhibit the well-documented negative relation. Thus, the announcement effect is present in the Finnish market.

If the effect of institutional trading is a manifestation of the announcement effect then we should not see the upward sloping SML on High-IT days when announcement days are removed from the sample. As Fig. 8 clearly shows, the positive relation between beta and average returns is still present on High-IT days. Consequently, our results cannot be explained by macroeconomic announcements.

To formally control for announcement days, we estimate Model (4), where *Control* is replaced by *Ann*, which takes a value of one if the day has a scheduled interest rate announcement and zero otherwise. We report the regression results in Table 8. First, in Model (1), we confirm the finding of Savor and Wilson (2014) that the coefficient on the interaction term between *Ann* and beta is positive, indicating that the market risk premium is higher on announcement days than on non-announcement days. When we control for announcement days in Model (2), we observe that the coefficient on the interaction between beta and High-IT is 27 bps, which is significant at the 1% level, suggesting that the risk premium is higher on High-IT days regardless of whether the day has an announcement. We thus conclude that the announcement effect cannot explain our findings. In Appendix A, we further confirm that the IT effect cannot be explained by the U.S. FOMC announcements.[10]

5.2. Leverage-constraints

We next consider the leverage-constraints hypothesis as a possible explanation of the IT effect. Frazzini and Pedersen (2014) argue that the relation between beta and average returns is flat because investors are constrained in the leverage that they can take. They show that institutions that are less leverage-constrained can

[10] In unreported robustness tests, we confirm that the IT effect holds after controlling for other types of macroeconomic news. Following Hendershott et al. (2015), we identify days with news of the following topic codes as provided in the TRNA database: ECI (Economic Indicator), GVD (Government/Sovereign Debt), MCE (Macroeconomics). We also consider other U.S.-related macroeconomic news with the topic codes of FED (Federal Reserve) and WASH (Washington/U.S. Government news).

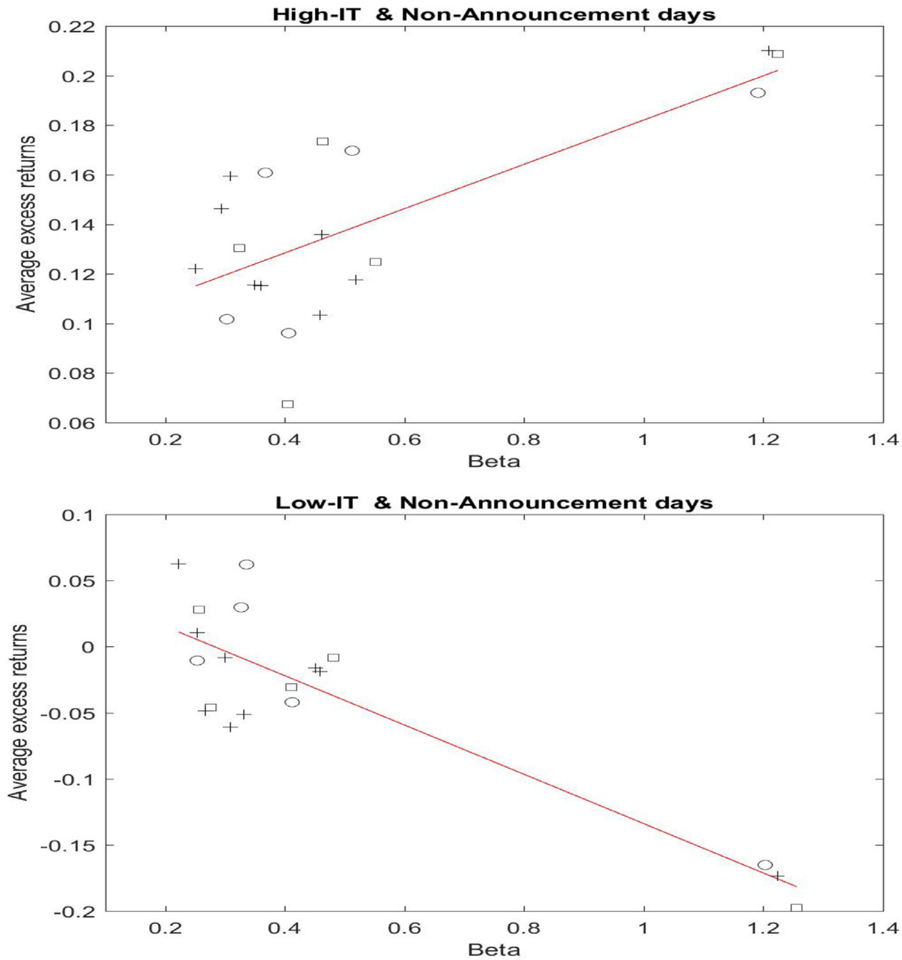


Fig. 8. Capital asset pricing model – non-announcement days. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ◯ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol) conditioned on the IT effect and non-announcement days. The first graph shows the relation between average returns and beta on days with High-IT volume but no ECB monetary policy decision. The second graph presents a similar line on days with Low-IT volume and no announcement. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns. The implied ordinary least squares estimates of the securities market line for each type of day are also plotted. The sample period is between 1999 and 2011 (the ECB was formally established in 1999).

take advantage of this flat relation by pursuing a profitable betting-against-beta strategy, which buys low and sells high-beta stocks.

If High-IT days represent times when overall leverage constraints are less binding so that institutions can employ leverage to bet against beta, stocks that they buy should have low betas while stocks they sell should have high betas. To test whether leverage constraints can explain our results, we follow Frazzini and Pedersen (2014) and employ the TED spread as a proxy for overall funding constraints in the market. Frazzini and Pedersen (2014) contend that a high TED spread indicates that leverage constraints become more binding. We tailor the TED definition to the Finnish market by defining this spread to be the difference between three-month EURIBOR (Euro interbank offered rate) and the yield on short-term Finnish government bonds. The sample period is between 1996 and 2011. For the interbank rate before 1999, we use HELIBOR (Helsinki interbank offered rate). We then define a day to have high leverage constraints if the day's TED is higher than the average over the past quarter. Models (3) and (4) in Table 8 test the effect of funding constraints by using high TED days as control in regression (4). First, Model (3) shows a negative coefficient on $beta \times TED$, consistent with the leverage constraints hypothesis that the market risk premium is lower when leverage constraints are more binding. Second, Model (4) shows that, after controlling for leverage constraints, the coefficient on $beta \times High$ is still positive and significant. This result suggests that market-wide leverage constraints hypothesis cannot explain our findings on High- versus Low-IT days.

5.3. Controlling for market returns

We examine whether the effect of High IT exists only on up-market days. Mechanically, the market risk premium is positive on days when the market return is positive and thus, the CAPM appears to hold on these up-market days. In this section, we test whether our findings are simply a manifestation of the mechanical up-market effect. We construct an *UpMarket* dummy variable that takes the value of one if the day's market return is positive or zero otherwise. We then estimate Model (4), where Control is replaced with the *UpMarket* dummy. Not surprisingly, Table 8, Model (5) shows that the market risk premium on *UpMarket* days is positive and higher than on other days.

Model (6) shows that the coefficient on $beta \times High$ is positive and significant at the 1% level, even after controlling for *UpMarket*. These results suggest the evidence of stronger beta pricing on High-IT days is distinct from the effect of up-market states. In Appendix B, we regress IT volume on its lagged volume, market returns, size, and BM factors. We find that lagged IT volume is the sole significant predictor of current IT volume, while the other controls are insignificant. These results suggest that the IT effect

Table 8
Regressions to test alternative hypotheses.

	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.0008	0.00070	0.00001	−0.00014	0.00015	0.00010
	(5.60)	(3.55)	(0.02)	(−0.61)	(0.84)	(0.37)
Beta	−0.00048	−0.00161	0.00031	−0.00107	−0.0133	−0.01369
	(−2.26)	(−5.71)	(1.15)	(−3.23)	(−52.07)	(−46.50)
HighIT		0.00025		0.00029		0.00017
		(0.87)		(1.03)		(0.66)
Beta*HighIT		**0.00266**		**0.00286**		**0.00107**
		(6.42)		**(7.06)**		**(2.99)**
Ann	−0.00097	−0.00097				
	(−1.59)	(−1.59)				
Beta*Ann	0.00249	0.00235				
	(2.79)	(2.64)				
HighTED			0.00155	0.00155		
			(5.40)	(5.43)		
Beta*HighTED			−0.00084	−0.00054		
			(−2.07)	(−1.32)		
UpMarket					0.00113	0.00112
					(4.54)	(4.48)
Beta*UpMarket					0.02485	0.02477
					(70.29)	(69.90)
R^2	0.0001	0.003	0.0001	0.004	0.23	0.24

Note: This table reports results for the following pooled cross-sectional regression:
$R_{i,t+1} = \gamma_0 + \gamma_1 \hat{\beta}_{i,t} + \gamma_2 High_{t+1} + \gamma_3 \times \hat{\beta}_{i,t} \times High_{t+1} + \gamma_4 \times Control_{t+1} + \gamma_5 \times \hat{\beta}_{i,t} \times Control_{t+1} + u_{i,t+1}$
where $R_{i,t+1}$ is the excess return on test assets (five beta-sorted portfolios, nine value-weighted size-BM portfolios, and five industry portfolios); $Control = \{Ann, HighTED, HighDisagree, HighCoskewness, PositiveRm\}$; *Ann* equals one if the day is an ECB announcement day or zero otherwise; *HighTED* – a proxy for leverage constraints in the market – equals one if the day is High TED day or zero otherwise; *HighDisagree* equals one if the day has high market disagreement or zero otherwise; *HighCoskewness* equals one if the day has high market coskewness or zero otherwise. In regressions (5) and (6), "UpMarket" equals one if the day's market return is positive or zero otherwise. Due to the mechanical effect of positive market returns on the risk premium (i.e., the UpMarket variable picks up days when market risk premium must be positive and the CAPM should work by selection), we see a large coefficient (and associated *t*-statistic) on beta*UpMarket. Thus, tests in Models (5) and (6) are conservative. *t*-statistics, reported in parentheses, are computed using clustered standard errors by trading day.

Table 9
Daily excess returns on High-IT buying and High-IT selling days.

	Fama–MacBeth regression			Pooled regression				
Type of day	Intercept	Beta	Avg. R^2	Intercept	Beta	High	Beta*High	R^2
Panel A: high- versus Low-IT buying days								
High	0.000070	0.000827	0.11	0.000500	−0.001201	0.000016	**0.002548**	0.002
	(0.51)	(3.17)		(2.57)	(−4.37)	(0.06)	**(6.29)**	
Low	0.000314	−0.000926	0.13					
	(1.50)	(−2.77)						
High – Low	−0.000243	**0.001753**						
	(−1.06)	**(4.10)**						
Panel B: high- versus Low-IT selling days								
High	0.000349	0.000156	0.11	0.000078	−0.000121	0.000933	**0.000151**	0.001
	(2.24)	(0.61)		(0.40)	(−0.44)	(3.27)	**(0.37)**	
Low	0.000035	−0.000255	0.13					
	(0.18)	(−0.75)						
High – Low	0.000313	**0.000411**						
	(1.38)	**(0.96)**						

Note: This table reports estimates from Fama–MacBeth regressions of daily excess returns on betas for various test portfolios on high- and low-IT buying and selling days (rather than total IT volume as in previous tables). Test assets are five beta-sorted, nine size-BM, and five industry portfolios. Estimates are computed for days with high institutional trading (High-IT days or High) and other days (Low-IT days or Low). Day *t* is a High-IT day when IT buy (sell) volume (scaled by total market volume) at *t* is greater than its average over the past quarter. The difference in the coefficient between High- and Low-IT days is reported in the last row of each panel. The right-hand side panel reports estimates from the pooled regression of excess returns on betas, High-IT day dummy, and interaction between beta and High-IT (Beta*High). Panel A shows the estimate for buying volume (scaled by total market volume) while Panel B shows the results for selling volume (scaled by total market volume). *t*-statistics, which are computed using Newey–West standard errors with five lags, are reported in parentheses. For pooled regressions, standard errors are clustered by trading day. This table shows that the difference in the coefficient on beta between High- and Low-IT days is driven by the buying side of institutional trading.

cannot be explained by the predictability of market returns from IT volume.[11]

5.4. Short-sale constraints

High-IT days could be times when short-sale constraints are less binding, which makes it less costly for institutions to short sell and, in turn, helps cause the stock price to be more efficient. In fact, Nagel (2005) employs institutional holdings as a proxy for short-sale constraints and finds that prices of stocks with low institutional ownership do not quickly reflect cash-flow news. Although our data do not allow us to identify short sales, under this hypothesis, we should observe that the IT effect is stronger on the selling side of institutional trading. Consequently, we separately examine the effects of High-IT buying and High-IT selling.

Panel A of Table 9 reports regression results for High- versus Low-IT buying days. On High-IT days, the coefficient on beta is 8.3 bps with a *t*-statistic of 3.2. The picture reverses on Low-IT buy-

[11] As an additional robustness test for the mechanical effect of up- and down-market days, we assess the relation between institutional trading activity and the slope of the SML on subsamples of up- and down-market days. We confirm that our main findings hold for both subsamples.

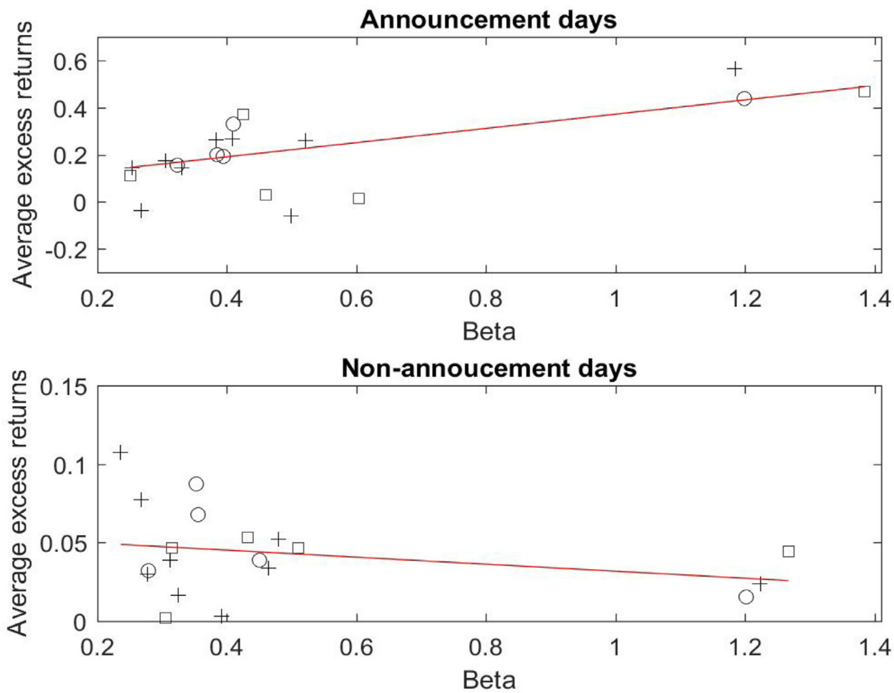


Fig. A1. Capital asset pricing model on FOMC announcement and non-announcement days. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ◯ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol) on FOMC scheduled announcement versus non-announcement days. The first graph shows the relation between beta and average returns on days with FOMC monetary policy decision. The second graph presents similar line on non-announcement days. The implied ordinary least squares estimates of the Security Market Line for each type of day are also plotted. The sample period is between 1997 and 2011. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns.

ing days on which the coefficient is negative and significant. The difference between the two days is 17.5 bps with an associated *t*-statistic greater than four. The pooled regression shows a consistent result that the coefficient on $beta \times High$ is positive and significant. These results indicate that the IT effect is stronger on the buying side.

Panel B reports results for High- versus Low-IT selling days. In contrast to the buying side, the selling side of IT exhibits an insignificant CAPM relation. On High-IT selling days, the beta in the Fama–MacBeth regression is small and insignificant. Similarly, the pooled regression shows that the coefficient on the interaction term $beta \times High$ is insignificant even at the 10% level.

In short, Table 9 shows that the IT effect is driven by the buying side, rather than the selling side, of institutional trading. While these results are inconsistent with the prediction of the short-sale constraints hypothesis, they are actually intuitive. When institutions buy, they tend to rely more on the analysis of risk and return and could be using the standard capital asset pricing model. On the other hand, they are less likely to sell a stock because of its risk level.

6. Conclusion

By employing a comprehensive and unique data set of daily institutional trades, this study is one of the first to directly link the effect of institutional trading to the relation between beta and average returns. We find that this relation is strong and positive on High-IT days. The difference in market risk premium between High- and Low-IT days is also positive and significant. These findings hold for various test portfolios as well as individual stocks, and are not driven by a specific sub-sample period, up-market states, the January effect, or the turn-of-month effect. We also demonstrate that our results cannot be explained by alternative hypotheses such as short-sale constraints, liquidity, small market problems, macroeconomic announcements, leverage-constraints, market disagreement, or market-wide lottery characteristics. Our findings are generally consistent with the notion advocated in the intermediary-based asset pricing literature that financial institutions are the marginal investors, who drive the positive relation between beta and average returns when they trade actively.

Our study is the first to establish a connection between trading activity and the validity of the CAPM. There is a considerable body of literature documenting the heterogeneity in trading behavior of various investor types in the equity market (e.g., Barber and Odean, 2013). In the literature on flat-beta puzzles, however, traditional tests of the CAPM have not been able to account for this long-standing consensus regarding the different preferences between institutions and individual investors. Moreover, despite the advance of intermediary-based asset pricing theory, existing tests still assume that the marginal investor is the household. Using a unique data set that covers the entire market, we are able to disentangle the trading impact of institutions (who are ex-ante more likely to be traditional investors) from individual investors in the test of the CAPM. We offer new evidence on the tale of two days that cannot be explained by existing market-wide explanations for the failure of the CAPM. On the whole, our findings suggest that institutions are key to explaining the flat-beta puzzle.

Appendix A. High-IT vs. FOMC announcements

This section examines whether the IT effect in Finland could be driven by days of scheduled FOMC announcements in the U.S. First, Fig. A1 establishes the effect of FOMC announcements on the relation between beta and returns in the Finnish market. On days with FOMC announcements, the SML is upward sloping whereas on non-announcement days, the SML is downward sloping. We then test whether the IT effect in Finland could be driven by FOMC

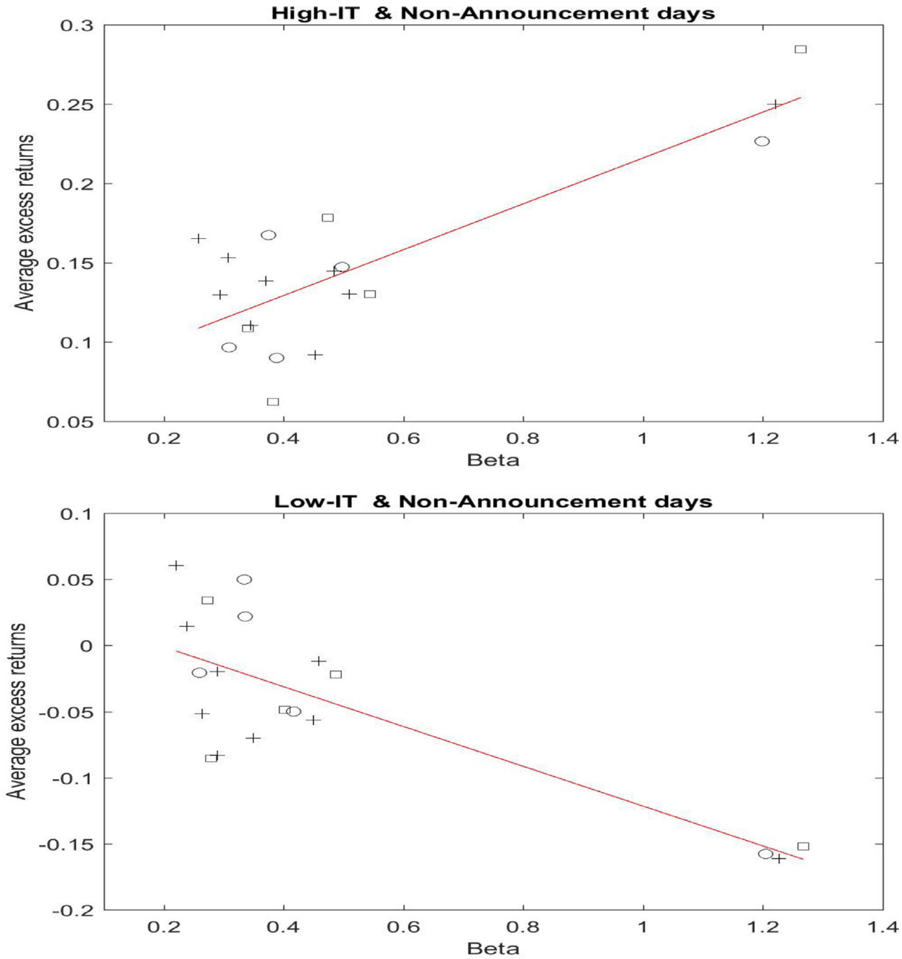


Fig. A2. Capital asset pricing model – non-FOMC-announcement days. This figure plots average daily excess returns (in %) against market betas for five value-weighted beta portfolios (denoted with ◯ symbol), nine value-weighted size-BM portfolios (denoted with + symbol), and five value-weighted industry portfolios (denoted with □ symbol) on High- and Low-IT days accounting for the FOMC announcement effect. The first graph shows the relation between excess returns and beta on days with High-IT volume but no FOMC monetary policy decision, while the second graph presents the relation on days with Low-IT volume and no announcement. The implied ordinary least squares estimates of the securities market line for each type of day are also plotted. The sample period is between 1997 and 2011. Beta of each portfolio is estimated from the full-sample regression of daily excess portfolio returns on excess market returns. The purpose of this figure is to show that, even when the day has no macroeconomic announcement, the difference in the risk premium between High- and Low-IT days remains robust. Consequently, the effect of institutional trading is not a manifestation of announcement effects.

announcement days. Replicating the methodology in the main text for Fig. 8, we examine the SML on days of High-IT and non-FOMC announcement. We find that the High-IT effect in Finland is still strong even though the day does not have an announcement. These results confirm that our IT effects are distinct from the announcement effect.

Appendix B. Other robustness tests

In this section, we provide an extra test for whether institutional trading has an effect on the relation between beta and expected returns by performing a two-stage regression. In the first stage, we regress aggregated IT volume ($ITVol_t$) on its lags over the past five days, as well as market-level controls such as (contemporaneous and lagged) market returns, value, and size factors. We then obtain the predicted value from this regression and compute High-IT dummy as usual. In the second stage, we run the regression (3) and report the estimates in Table A1.

Panel A1 of Table A1 reports the estimation results for the first-stage regression. We can see that past aggregate IT volume is the strongest predictor of current $ITVol_t$, while the coefficients on market returns, size, and book-to-market factors are not significant. In untabulated tests, we also control for lagged returns on the test portfolios and find that these variables are not significant predictors of $ITVol_t$. These findings are consistent with the theoretical argument of Hong and Stein (2007) that investors' preferences are a simple function of their own priors and past trading volume.

The second-stage regression in Panel A2 shows that the coefficient on the interaction between beta and predicted High-IT remains positive and statistically significant at the 1% level. Since High-IT is highly predictable using lagged IT volume, these results suggest that the IT effect can have an impact on the relation between expected stock returns and stock betas. Furthermore, since investors' preferences are persistent, these findings are also consistent with the notion that High-IT (Low-IT) days exhibit stronger preferences of institutional (individual) investors (Hong and Stein, 2007).

For completeness, in Panel B of Table A1, we re-run regression (3) using lagged High-IT dummy. We obtain consistent results that the coefficient on the interaction between lagged High-IT and beta is positive and statistically significant. These results point to the same evidence that high institutional trading helps makes the beta risk correctly priced on the following day.

Table A1
Predicted IT, Lagged IT, and the CAPM.

Intercept	$ITVol_{t-1}$	$ITVol_{t-2}$	$ITVol_{t-3}$	$ITVol_{t-4}$	$ITVol_{t-5}$	$R_{M,\,t}$	$R_{M,t-1}$	
Panel A1: first-stage regression (dependent variable: $ITVol_t$)								
0.097	0.386	0.230	0.209	0.0517	0.118	8.455	−3.157	
(1.66)	(6.22)	(7.11)	(5.65)	(1.51)	(3.54)	(1.65)	(−0.74)	
	$R_{M,t-2}$	HML_t	HML_{t-1}	HML_{t-2}	SMB_t	SMB_{t-1}	SMB_{t-2}	R^2
	−1.972	−1.906	0.768	1.308	0.424	−8.419	−5.703	0.98
	(−0.44)	(−0.41)	(0.15)	(0.26)	(0.06)	(−1.52)	(−0.93)	
Panel A2: second-stage regression (dependent variables: 18 portfolios)								
Intercept		Beta		$\widehat{High}$		$\widehat{High} \times Beta$		R^2
0.000618		−0.000620		−0.000181		**0.001120**		0.0002
(3.53)		(−2.09)		(−0.69)		**(2.54)**		
Panel B: regression with lagged High-IT								
Intercept		Beta		*LaggedHigh*		*LaggedHigh* × *Beta*		R^2
−0.004799		−0.005704		0.000245		**0.000589**		0.002
(−5.83)		(−3.53)		(1.75)		**(2.48)**		

Note: The table reports two-stage regression results to examine the effect of IT on the relation between beta and expected returns. The first-stage regression regresses institutional trading volume ($ITVol_t$) on its lags as well as market-level measures such as market returns, BM, and size factors, where $ITVol_t$ is the aggregate daily trading volume of all institutions scaled by the total market trading volume. We then obtain the predicted IT volume and construct the High-IT dummy variable as before using the predicted volume. In the second stage, we run the pooled cross-sectional regression (3) using the predicted High-IT dummy. Test assets are excess returns on nine value-weighted size-BM portfolios, five beta-sorted portfolios, and five industry portfolios. *t*-statistics, whose standard errors are corrected for clustering by trading day, are reported in parentheses. In Panel B, we re-run one-stage regression (3) using the lagged High-IT dummy instead of contemporaneous High-IT. This table shows that the coefficient on $\widehat{High} \times beta$ is positive and statistically significant even after we use either predicted IT or lagged IT. These results suggest that institutional trading could have an impact on the relation between expected returns and beta. The evidence is also consistent with the notion that investors' preferences, which are persistent, could affect the future CAPM relation.

References

Adrian, T., Etula, E., Muir, T., 2014. Financial intermediaries and the cross-section of asset returns. Journal of Finance 69 (6), 2557–2596.
Ang, A., Chen, J., Xing, Y., 2006. Downside risk. Review of Financial Studies 19, 1191–1239.
Bali, T.G., Cakici, N., Whitelaw, R., 2011. Maxing out: stocks as lotteries and the cross-section of expected returns. J. Financ. Econ. 99, 427–446.
Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21, 785–818.
Barber, B.M., Odean, T., 2013. Chapter 22 – The Behavior of Individual Investors. In: Handbook of the Economics of Finance, 2, Part B. Elsevier, pp. 1533–1570. doi:10.1016/B978-0-44-459406-8.00022-6.
Boehmer, E., Kelly, E.K., 2009. Institutional investors and the information efficiency of prices. Rev. Financ. Stud. 22 (9), 3563–3594.
Chordia, T., Goyal, A., Jegadeesh, N., 2016. Buyers versus sellers: who initiates trades and when? J. Financ. Quant. Anal. 51 (5), 1467–1490.
Cronqvist, H., Siegel, S., 2014. The genetics of investment biases. J. Financ. Econ. 113 (2), 215–234.
Cronqvist, H., Siegel, S., Yu, F., 2015. Value versus growth investing: why do different investors have different styles. J. Financ. Econ. 117, 333–349.
Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. J. Financ. Econ. 60, 1673–1712.
Dzieliński, M., Hasseltoft, H., 2017. News Tone Dispersion and Investor Disagreement. Available at SSRN: https://ssrn.com/abstract=2192532 or https://dx.doi.org/10.2139/ssrn.2192532.
Fama, E., French, K.R., 1992. The cross section of expected stock returns. J. Financ. Econ. 47, 427–465.
Fama, E., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56.
Fama, E.F., MacBeth, J.D., 1973. Risk, return and equilibrium: empirical tests. J. Polit. Econ. 81, 607–636.
Foster, F.D., Gallagher, D.R., Looi, A., 2011. Institutional trading and share returns. J. Bank. Financ. 35 (12), 3383–3399.
Frazzini, A., Pedersen, L.H., 2014. Betting against beta. J. Financ. Econ. 111, 1–25.
Griffin, J.M., Kelly, P.J., Nardari, F., 2010. Do market efficiency measures yield correct inferences? a comparison of developed and emerging markets. Rev. Financ. Stud. 23, 3225–3277.
Grinblatt, M., Keloharju, M., 2000. The investment behavior and performance of various investor types: a study of Finland's unique data set. J. Financ. Econ. 55, 43–67.
Grinblatt, M., Keloharju, M., 2001a. How distance, language, and culture influence stock holdings and trades. J. Financ. 56 (3), 1053–1073.
Grinblatt, M., Keloharju, M., 2001b. What makes investors trade? J. Financ. 56, 589–616.
Grinblatt, M., Keloharju, M., 2009. Sensation seeking, overconfidence, and trading activity. J. Financ. 64 (2), 549–578.
Grinblatt, M., Keloharju, M., Linnainmaa, J.T., 2012. IQ, trading behavior, and performance. J. Financ. Econ. 104 (2), 339–362.
Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing. J. Financ. 34, 465–477.
He, Z., Kelly, B., Manela, A., 2017. Intermediary asset pricing: new evidence from many asset classes. J. Financ. Econ. 126 (1), 1–35.
He, Z., Krishnamurthy, A., 2013. Intermediary asset pricing. Am. Econ. Rev. 103 (2), 732–770.
Hendershott, T., Livdan, D., Schurhoff, N., 2015. Are institutions informed about news? J. Financ. Econ. 117 (2), 249–287.
Hong, H., Sraer, D., 2016. Speculative betas. J. Financ. 71 (5), 2095–2144.
Hong, H., Stein, J.C., 2007. Disagreement and the stock market. J. Econ. Perspect. 21 (2), 109–128.
Ince, O.S., Porter, R.B., 2006. Individual equity return data from Thomson datastream: handle with care!. J. Financ. Res. 29, 463–479.
Itzkowitz, J., Itzkowitz, J., Rothbort, S., 2015. ABCS of trading: behavioral biases affect stock turnover and value. Rev. Financ. 20 (2), 663–692.
Kumar, A., 2009. Who gambles in the stock market? J. Financ. 64, 1889–1933.
Lesmond, D.A., Ogden, J.P., Trzcinka, C.A., 1999. A new estimate of transaction costs. Rev. Financ. Stud. 12 (5), 1113–1141.
Leung, H., Rose, A., Westerholm, P.J., 2014. Systematic trading behavior and the cross-section of stock returns on the OMXH. Rev. Financ. 18 (6), 2325–2374.
Lewellen, J., Nagel, S., Shanken, J., 2010. A skeptical appraisal of asset pricing tests. J. Financ. Econ. 96, 175–194.
Mitton, T., Vorkink, K., 2007. Equilibrium underdiversification and the preference for skewness. Rev. Financ. Stud. 20 (4), 1255–1288.
Nagel, S., 2005. Short sales, institutional investors and the cross-section of stock returns. J. Financ. Econ. 78, 277–309.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Savor, P., Wilson, M., 2014. Asset pricing: a tale of two days. J. Financ. Econ. 113, 171–201.
Sias, R.W., Starks, L.T., 1997. Institutions and individuals at the turn-of-the-year. J. Financ. 52 (4), 1543–1562.
Sikes, S.A., 2014. The turn-of-the-year effect and tax-loss-selling by institutional investors. J. Account. Econ. 57, 22–42.
Siriwardane, E., 2016. Concentrated Capital Losses and the Pricing of Corporate Credit Risk. Harvard Business School Finance Working Paper No. 16-007. Available at SSRN: https://ssrn.com/abstract=2584043 or https://dx.doi.org/10.2139/ssrn.2584043.