Internet Appendix:
So Many Jumps, So Few News

July 17, 2024

A.1. Details for Computing Abnormal Returns

In this section, we elaborate on the steps for computing the abnormal returns in (2). The first step focuses on the construction of high-frequency factors – the intraday returns of portfolios sorted by various firm-level characteristics. The methodology for this construction follows Aït-Sahalia, Kalnina and Xiu (2020). The return of the market factor (MKT) is simply the value-weighted returns of all 30 DJIA components. More precisely, for an interval $[(i-1)\Delta, i\Delta]$ on trading day t, the market return is given by the following weighted average

$$r_{i\Delta}^{(\mathrm{MKT})} = \frac{\sum_{j=1}^{30} w_{j,(i-1)\Delta} r_{i\Delta}^{(j)}}{\sum_{j=1}^{30} w_{j,(i-1)\Delta}}.$$

Here, $r_{i\Delta}^{(j)}$ represents the log return of stock j in $[(i-1)\Delta, i\Delta]$, i.e., $Y_{i\Delta} - Y_{(i-1)\Delta}$, and $w_{j,(i-1)\Delta}$ is the market cap of stock j at time $(i-1)\Delta$, i.e.,

$$w_{j,(i-1)\Delta} = p_{(i-1)\Delta}^{(j)} \mathrm{shrout}_{t-1}, \tag{A.1}$$

where $p_{(i-1)\Delta}^{(j)}$ is the price of stock j at time $(i-1)\Delta$ and shrout_{t-1} is the number of shares outstanding on the last trading day $t-1$.

The construction of the rest of factors are similar; take the value factor (HML) for example. We retrieve annual data of book-to-market ratios from the Compustat database and then classify the companies into three categories at the 30th and 70th percentiles of their book-to-market ratios cross-sectionally. Companies in the top 30%, in the middle 40%, and at the bottom 30% are labeled with H, M, and L respectively. On the other hand, based on the data of market cap computed according to (A.1), we independently label a DJIA stock with B (resp. S) representing big (resp. small) firms, if its market cap falls in the top 50% (resp. bottom 50%). The HML is the return of the following long-short portfolio

$$\mathrm{HML} = \frac{1}{2}(\mathrm{SH} + \mathrm{BH}) - \frac{1}{2}\left(\mathrm{SL} + \mathrm{BL}\right),$$

where SH, BH, SL, BL are sub-portfolios constructed by companies labelled with S and H, B and H, S and L, as well as B and L, respectively. Take the return of SH portfolio for example. Its return is given by

$$r_{i\Delta}^{(\mathrm{SH})} = \frac{\sum_{j\in \mathrm{SH}} w_{j,(i-1)\Delta} r_{i\Delta}^{(j)}}{\sum_{j\in \mathrm{SH}} w_{j,(i-1)\Delta}}.$$

The outputs of the first step include the time series of all five factors at a prespecified frequency Δ.

The second step targets the estimation of the beta coefficients (continuous and jump) based on the discrete-time observations of stocks and factors obtained in the first step. Suppose each time series consists of n observations on the day t and the continuous beta β^C is a constant within the day. To estimate β^C, we introduce the notation

$$\begin{aligned}&(\Delta_i X)\, 1_{\{\|\Delta_i X\| \le u_t\}}\\ &= \left((\Delta_i X^{(1)}) 1_{\left\{\left|\Delta_i X^{(1)}\right| \le u_t^{(1)}\right\}}, (\Delta_i X^{(2)}) 1_{\left\{\left|\Delta_i X^{(2)}\right| \le u_t^{(2)}\right\}}, \cdots, (\Delta_i X^{(d)}) 1_{\left\{\left|\Delta_i X^{(d)}\right| \le u_t^{(d)}\right\}} \right),\end{aligned}$$

with $\Delta_i X^{(k)} = X_{i\Delta}^{(k)} - X_{(i-1)\Delta}^{(k)}$, for $k = 1, 2, \ldots, d$, and collect the truncated returns of the factors X and truncated returns of stock Y on day t in the following matrix $\mathcal{X}$ and vector $\mathcal{Y}$, respectively,

$$\mathcal{X} = \begin{pmatrix} (\Delta_1 X)\, 1_{\{\|\Delta_1 X\| \le u_t\}} \\ (\Delta_2 X)\, 1_{\{\|\Delta_2 X\| \le u_t\}} \\ \vdots \\ (\Delta_n X)\, 1_{\{\|\Delta_n X\| \le u_t\}} \end{pmatrix} \text{ and } \mathcal{Y} = \begin{pmatrix} (\Delta_1 Y)\, 1_{\{|\Delta_1 Y| \le u_t'\}} \\ (\Delta_2 Y)\, 1_{\{|\Delta_2 Y| \le u_t'\}} \\ \vdots \\ (\Delta_n Y)\, 1_{\{|\Delta_n Y| \le u_t'\}} \end{pmatrix},$$

where $\Delta_i Y = Y_{i\Delta} - Y_{(i-1)\Delta}$ and u_t' represents the threshold for truncating the returns of the stock Y. A standard choice of the above thresholds u_t and u_t' is derived based on the bipower variation of the corresponding process. Take the kth dimension of X for instance. We choose $u_t^{(k)} = 3\Delta^{0.47}\sqrt{\mathrm{BV}_t^{(k)}}$, where $\mathrm{BV}_t^{(k)}$ represents the bipower variation of the kth factor on day

t.[1] Then, the OLS estimator of β^C follows

$$\hat{\beta}^C = (\mathcal{X}^\intercal \mathcal{X})^{-1} \mathcal{X}^\intercal \mathcal{Y}.$$

As to the optimal choice of observation frequency Δ for estimating β^C, we could perform a Hausman test for each subsample but since the stocks we consider are highly liquid, we simply set $\Delta = 1$mn for the daily subsamples of all stocks.

To estimate the jump beta β^J, we refer to Aït-Sahalia, Jacod and Xiu (2023) and assume β^J is a constant within each calendar year. Suppose each time series consists of N_s observations in the year s. The OLS estimator of β^J follows

$$\hat{\beta}^J = \left(\tilde{\mathcal{X}}^\intercal \tilde{\mathcal{X}}\right)^{-1} \tilde{\mathcal{X}}^\intercal \tilde{\mathcal{Y}}.$$

Here, the matrix $\tilde{\mathcal{X}}$ and the vector $\tilde{\mathcal{Y}}$ are given by

$$\tilde{\mathcal{X}} = \begin{pmatrix} (\Delta_1 X)\, 1_{\{\|\Delta_1 X\| > u_t\}} \\ (\Delta_2 X)\, 1_{\{\|\Delta_2 X\| > u_t\}} \\ \vdots \\ (\Delta_{N_s} X)\, 1_{\{\|\Delta_{N_s} X\| > u_t\}} \end{pmatrix} \text{ and } \tilde{\mathcal{Y}} = \begin{pmatrix} (\Delta_1 Y) \\ (\Delta_2 Y) \\ \vdots \\ (\Delta_{N_s} Y) \end{pmatrix},$$

where

$$\begin{aligned} &(\Delta_i X)\, 1_{\{\|\Delta_i X\| > u_t\}} \\ &= \left((\Delta_i X^{(1)}) 1_{\left\{\left|\Delta_i X^{(1)}\right| > u_t^{(1)}\right\}}, (\Delta_i X^{(2)}) 1_{\left\{\left|\Delta_i X^{(2)}\right| > u_t^{(2)}\right\}}, \cdots, (\Delta_i X^{(d)}) 1_{\left\{\left|\Delta_i X^{(d)}\right| > u_t^{(d)}\right\}} \right). \end{aligned}$$

A.2. Details for Jump Detection at High Frequency

As discussed in Section 2.3, the detection of jumps relies on the estimation of a noise-robust realized volatility RV_t, based on combining two time scales. We update the estimator of RV_t for each trading day. Given a time series of 5s returns on day t, we estimate RV_t according to the following procedure. First, we eliminate the jump components in returns: we remove returns $R_{[t_i - 5\text{s}, t_i]}$ satisfying $\left|R_{[t_i - 5\text{s}, t_i]}\right| > \hat{c} \times \widehat{RV}_{t-1} \sqrt{\Delta}$, where Δ is the annualized time interval,

[1] Instead of truncating the increments of each factor separately, one can also truncate the $d-$dimensional increments in global based on some norm of the $d-$dimensional process and with a scalar threshold u_t. However, these two types of truncations share the same asymptotic properties and perform equally well in practice; see, e.g., Remark 6.7 of Aït-Sahalia and Jacod (2014).

i.e., $5/(6.5\times 60\times 60\times 252) = 8.48\times 10^{-7}$; $\widehat{RV}_{t-1}$ is the estimator of realized volatility RV_{t-1} on the previous trading day;[2] the constant $\hat{c}$ is another truncation cutoff parameter independent of c.[3] Denote by $\{R^{(5s)}_{t_i}\}_{t_i\in\mathcal{T}_0}$ the collection of 5s returns obtained from the aforementioned preparatory step, i.e., $R^{(5s)}_{t_i} = R_{[t_i-5s,t_i]}$, if $\left|R_{[t_i-5s,t_i]}\right| \leq \hat{c}\times\widehat{RV}_{t-1}\sqrt{\Delta}$, and $R^{(5s)}_{t_i} = 0$, otherwise; the set of time indices $\mathcal{T}_0$ is given by {09:30:05, 09:30:10, ..., 16:00:00}. Second, we construct 12 groups of 1mn returns $\{R^{(1mn)}_{t_i}\}_{t_i\in\mathcal{T}_j}$ for $j = 1, 2, \ldots, 12$. Here, the time index $\mathcal{T}_j$ is a subset of $\mathcal{T}_0$ given by $\{09{:}31{:}00+\delta_j, 09{:}32{:}00+\delta_j, \ldots, 15{:}59{:}00+\delta_j\}$, with $\delta_j = 5(j-1)$ seconds. For each $t_i \in \mathcal{T}_j$, the 1mn return $R^{(1mn)}_{t_i}$ accumulates the 5s returns between $t_i - 1$mn and t_i and is computed according to

$$R^{(1mn)}_{t_i} = \sum_{k=0}^{11} R^{(5s)}_{t_{i-k}}.$$

Third, we follow the "first-best estimator" approach proposed in Zhang, Mykland and Aït-Sahalia (2005) to estimate the realized volatility RV_t by

$$\widehat{RV}_t = \sqrt{\frac{252}{12}\left(\sum_{j=1}^{12}\sum_{t_i\in\mathcal{T}_j}\left(R^{(1mn)}_{t_i}\right)^2 - \sum_{t_i\in\mathcal{T}_0}\left(R^{(5s)}_{t_i}\right)^2\right)},$$

which is consistent and unbiased according to Zhang, Mykland and Aït-Sahalia (2005).

Following the above procedure, for each trading day, we estimate the realized volatility for each constituent company in DJIA individually and obtain 30 estimators altogether.

[2] If day t is the first trading day in the sample, we replace $\widehat{RV}_{t-1}$ by an arbitrary level of realized volatility RV_{arb}, e.g., $RV_{\mathrm{arb}} = 0.4$.

[3] Instead of considering various choices of $\hat{c}$, we simply fix the constant $\hat{c}$ as 6 in practice since it plays a similar role as the constant c in terms of tuning the final threshold $c\times\widehat{RV}_t\sqrt{\Delta}$ for jump detection. Indeed, with the choice of a small (resp. large) constant $\hat{c}$, one would exclude more (resp. fewer) returns from the time series and thus obtain a small (resp. large) estimator of RV_t through the construction of quadratic variation of returns. Then, a small (resp. large) RV_t leads to a small (resp. large) threshold of $c\times\widehat{RV}_t\sqrt{\Delta}$ consequently. Owing to the overlapping in the roles of constants c and $\hat{c}$, we fix $\hat{c}$ and leave all the degrees of freedom to c.

Company name	Abbreviation	Sample period
Alcoa	AA	9/10/2003–9/20/2013
Apple Inc.	AAPL	3/19/2015–12/31/2018
American International Group	AIG	4/8/2004–9/19/2008
American Express	AXP	9/10/2003–12/31/2018
Boeing	BA	9/10/2003–12/31/2018
Bank of America	BAC	2/19/2008–9/20/2013
Citigroup	C	9/10/2003–6/5/2009
Caterpillar Inc.	CAT	9/10/2003–12/31/2018
Cisco Systems	CSCO	6/8/2009–12/31/2018
Chevron Corporation	CVX	2/19/2008–12/31/2018
DuPont/DowDuPont	DD	trades with ticker DD during 9/10/2003–8/31/2017, and switches to DWDP during 9/1/2017–12/31/2018 after the merger of DowChemical Company with DuPont
Walt Disney Company	DIS	9/10/2003–12/31/2018
Kodak	EK	9/10/2003–4/7/2004
General Electric	GE	9/10/2003–6/25/2018
General Motors	GM	9/10/2003–6/5/2009
Goldman Sachs	GS	9/10/2003–12/31/2018
The Home Depot	HD	9/10/2003–12/31/2018
Honeywell	HON	9/10/2003–2/15/2008
Hewlett-Packard	HPQ	9/10/2003–9/20/2013
IBM	IBM	9/10/2003–12/31/2018
Intel	INTC	9/10/2003–12/31/2018
International Paper	IP	9/10/2003–4/7/2004
Johnson & Johnson	JNJ	9/10/2003–12/31/2018
J.P. Morgan Chase	JPM	9/10/2003–12/31/2018
Kraft Foods Inc.	KFT	9/22/2008–9/21/2012
The Coca-Cola Company	KO	9/10/2003–12/31/2018
McDonald's	MCD	9/10/2003–12/31/2018
3M	MMM	9/10/2003–12/31/2018
Altria Group	MO	9/10/2003–2/15/2008
Merck & Co.	MRK	9/10/2003–12/31/2018
Microsoft	MSFT	9/10/2003–12/31/2018
Nike, Inc.	NKE	9/23/2013–12/31/2018
Pfizer	PFE	4/8/2004–12/31/2018
Procter & Gamble	PG	9/10/2003–12/31/2018
SBC Communications/AT&T	SBC	trades with ticker SBC during 9/10/2003–11/30/2005, and swiches to T during 12/1/2005–3/18/2015 after acquisition
AT&T Corporation	T	of AT&T Corporation 9/10/2003–4/7/2004
The Travelers Companies	TRV	6/8/2009–12/31/2018
UnitedHealth Group	UNH	9/24/2012–12/31/2018
United Technologies	UTX	9/10/2003–12/31/2018
Visa Inc.	V	9/23/2013–12/31/2018
Verizon Communications	VZ	4/8/2004–12/31/2018
Walgreens Boots Alliance	WBA	6/26/2018–12/31/2018
Walmart	WMT	9/10/2003–12/31/2018
ExxonMobil	XOM	9/10/2003–12/31/2018

Table A.1: DJIA constituent companies during the sample period

Note: We exclude Kodak (EK), International Paper (IP), AT&T Corporation (T), and Walgreens Boots Alliance (WBA) when analyzing individual stocks due to their limited overlap as DJIA constituents (less than one year) with the sample period 9/10/2003–12/31/2018.

Category	Announcement	Observations	Frequency	Release time	Source
GDP and fiscal policy	GDP growth	192	monthly[a]	8:30	BEA
	Core CPI	192	monthly	8:30	BLS
	Core PPI	192	monthly	8:30	BLS
	Budget	192	monthly	14:00	CBO
	Tax rate	15	occasional[b]	–	–
	Index of leading indicator	192	monthly	10:00	CB
Labor market	Unemployment rate	192	monthly	8:30	BLS
	Initial jobless claims	836	weekly	8:30	DOL
	Labor force participation rate	96	monthly[c]	8:30	BLS
	Minimum wage	3	occasional[d]	–	DOL
	Nonfarm payroll	192	monthly	8:30	BLS
	Nonfarm productivity	128	twice per quarter[e]	8:30	BLS
Trade	Balance of trade	192	monthly	8:30	BEA
	Current account	64	quarterly	8:30	BEA
	Foreign direct investment	62	quarterly	8:30	BEA
Interest rate and monetary policy	Target federal funds rate	126	every six weeks	around 14:15	FED
	Money supply	833	weekly	16:30	FED
Consumers	Personal income	192	monthly	8:30	BEA
	Personal savings	192	monthly	8:30	BEA
	Consumer credit	192	monthly	15:00	FED
	Consumer spending	192	monthly	8:30	BEA
	Consumer sentiment index	383	biweekly[f]	around 10:00	UMich
	Consumer confidence index	192	monthly	10:00	CB
Housing market	Housing index	191	monthly	13:00/10:00[g]	NAHB
	Building permits	191	monthly	8:30	USCB
	Housing starts	190	monthly	8:30	USCB
	New home sales	190	monthly	10:00	BEA
Business	Capacity utilization	192	monthly	9:15	FED
	Business inventories	192	monthly	8:30/10:00	BEA
	Corporate profits	105	twice per quarter[h]	8:30	BEA
	Industrial production	192	monthly	9:15	FED
	Retail sales	192	monthly	8:30	USCB
	Durable goods orders	229	once/twice per month[i]	8:30/10:00	USCB
	Construction spending	191	monthly	10:00	USCB
	Factory orders	191	monthly	10:00	USCB

Table A.2: Macro news announcements

Note: In the Column "Source", BEA, BLS, CBO, CB, DOL, FED, UMich, NAHB, USCB are abbreviations of Bureau of Economic Analysis, Bureau of Labor Statistics, Congressional Budget Office, Conference Board, Department of Labor, Federal Reserve, University of Michigan, National Association of Home Builders, and U.S. Census Bureau, respectively. TRNA experienced an outage in recording some macroeconomic news in 2017 into the historical database. We complete the data using the release dates of macroeconomic announcements obtained from Bloomberg.

[a] For each quarter, there is one advance, one preliminary, and one final estimate of the GDP of the last quarter.

[b] We consider 15 major decisions on tax rate – Jobs and Growth Tax Relief Reconciliation Act of 2003, Tax Increase Prevention and Reconciliation Act of 2005, Pension Protection Act of 2006, Tax Relief and Health Care Act of 2006, Economic Stimulus Act of 2008, American Recovery and Reinvestment Act of 2009, Affordable Care Act, Small Business Jobs Act of 2010, Tax Relief, Unemployment Insurance Reauthorization, and Job Creation Act of 2010, Middle Class Tax Relief and Job Creation Act of 2012, American Taxpayer Relief Act of 2012, Consolidated Appropriations Act, 2016, Tax Cuts and Jobs Act of 2017, and Bipartisan Budget Act of 2018. The release times are the times of the last pass by the House or Senate (depending on which goes last) before sending to the president.

[c] The release of labor force participation rate starts from 2011.

[d] We consider three changes of minimum wage in 2007, 2008, and 2009, respectively.

[e] For each quarter, there is a preliminary and a final announcement of the nonfarm productivity of the last quarter.

[f] For each month, there is a preliminary and a final announcement of the consumer sentiment index of the last month.

[g] The release time switched from 13:00 to 10:00 on June 2010.

[h] For each quarter, there is a preliminary estimate and a follow-up revision of the corporate profits of the last quarter.

[i] The news of durable goods orders was monthly released at 8:30 until November 2015. Starting from December 2015, USCB releases two announcements at the beginning and end of each month, respectively. The former revises the estimate for the month before last month and was released at 10:00. The latter provides a preliminary estimate of the durable goods orders for the last month and is released at 8:30.

Variable	Estimator (z-stat)	Variable	Estimator (z-stat)	Variable	Estimator (z-stat)	Variable	Estimator (z-stat)
Panel A: Frequency = 15s							
Jump char.		Liquidity		Transaction impact		Cost	
AbsJSize	0.0988	Immediacy	−0.0560	Vol_all	0.1138	QSpread	−0.0148
	(95.3)		(−47.4)		(109.7)		(−12.0)
JSign	0.0058	Breadth	0.1433	Vol_avg	0.0126	ESpread_D	−0.0057
	(2.72)		(141.7)		(11.6)		(−4.8)
		BSize_avg	0.0069	Vol_max	0.0432	ESpread_S	−0.0057
			(6.50)		(41.0)		(−4.8)
		ASize_avg	0.0122	Turnover	0.1139	AbsRSpread_D	0.0349
			(11.5)		(109.8)		(32.9)
		BSize_all	0.0527	Lambda	−0.0802	AbsRSpread_S	0.0349
			(51.4)		(−68.8)		(32.9)
		ASize_all	0.0560	Autocov	0.0173		
			(54.7)		(15.5)		
		AbsLOBImb	0.0021	AbsTImb	−0.0580		
			(1.99)		(−53.5)		
				TQImb	0.0528		
					(49.8)		
Panel B: Frequency = 30s							
Variable	Estimator (z-stat)	Variable	Estimator (z-stat)	Variable	Estimator (z-stat)	Variable	Estimator (z-stat)
Jump char.		Liquidity		Transaction impact		Cost	
AbsJSize	0.1079	Immediacy	−0.0631	Vol_all	0.1382	QSpread	−0.0174
	(70.9)		(−38.3)		(89.1)		(−9.4)
JSign	0.0041	Breadth	0.1677	Vol_avg	0.0125	ESpread_D	−0.0095
	(1.31)		(111.7)		(7.8)		(−5.4)
		BSize_avg	0.0135	Vol_max	0.0453	ESpread_S	−0.0095
			(8.7)		(29.0)		(−5.4)
		ASize_avg	0.0192	Turnover	0.1384	AbsRSpread_D	0.0232
			(12.4)		(89.2)		(14.8)
		BSize_all	0.0612	Lambda	−0.1057	AbsRSpread_S	0.0232
			(40.5)		(−62.8)		(14.8)
		ASize_all	0.0654	Autocov	0.0191		
			(43.5)		(11.9)		
		AbsLOBImb	0.0005	AbsTImb	−0.0589		
			(0.34)		(−37.0)		
				TQImb	0.0585		
					(36.3)		
Panel C: Frequency = 60s							
Variable	Estimator (z-stat)	Variable	Estimator (z-stat)	Variable	Estimator (z-stat)	Variable	Estimator (z-stat)
Jump char.		Liquidity		Transaction impact		Cost	
AbsJSize	0.1146	Immediacy	−0.0921	Vol_all	0.1660	QSpread	−0.0167
	(52.7)		(−36.6)		(74.0)		(−6.2)
JSign	0.0072	Breadth	0.1883	Vol_avg	0.0195	ESpread_D	−0.0060
	(1.63)		(86.7)		(8.5)		(−2.33)
		BSize_avg	0.0208	Vol_max	0.0516	ESpread_S	−0.0060
			(9.4)		(23.1)		(−2.33)
		ASize_avg	0.0233	Turnover	0.1662	AbsRSpread_D	0.0137
			(10.6)		(74.1)		(6.1)
		BSize_all	0.0750	Lambda	−0.1281	AbsRSpread_S	0.0137
			(34.4)		(−53.8)		(6.1)
		ASize_all	0.0746	Autocov	0.0200		
			(34.4)		(8.7)		
		AbsLOBImb	−0.0031	AbsTImb	−0.0608		
			(−1.36)		(−26.7)		
				TQImb	0.0632		
					(27.0)		

Table A.3: Estimation results of panel regression (9) under various frequencies

Note: Except for the change of sampling frequency, all the settings for producing each panel are the same as those for Panel A of Table 8.

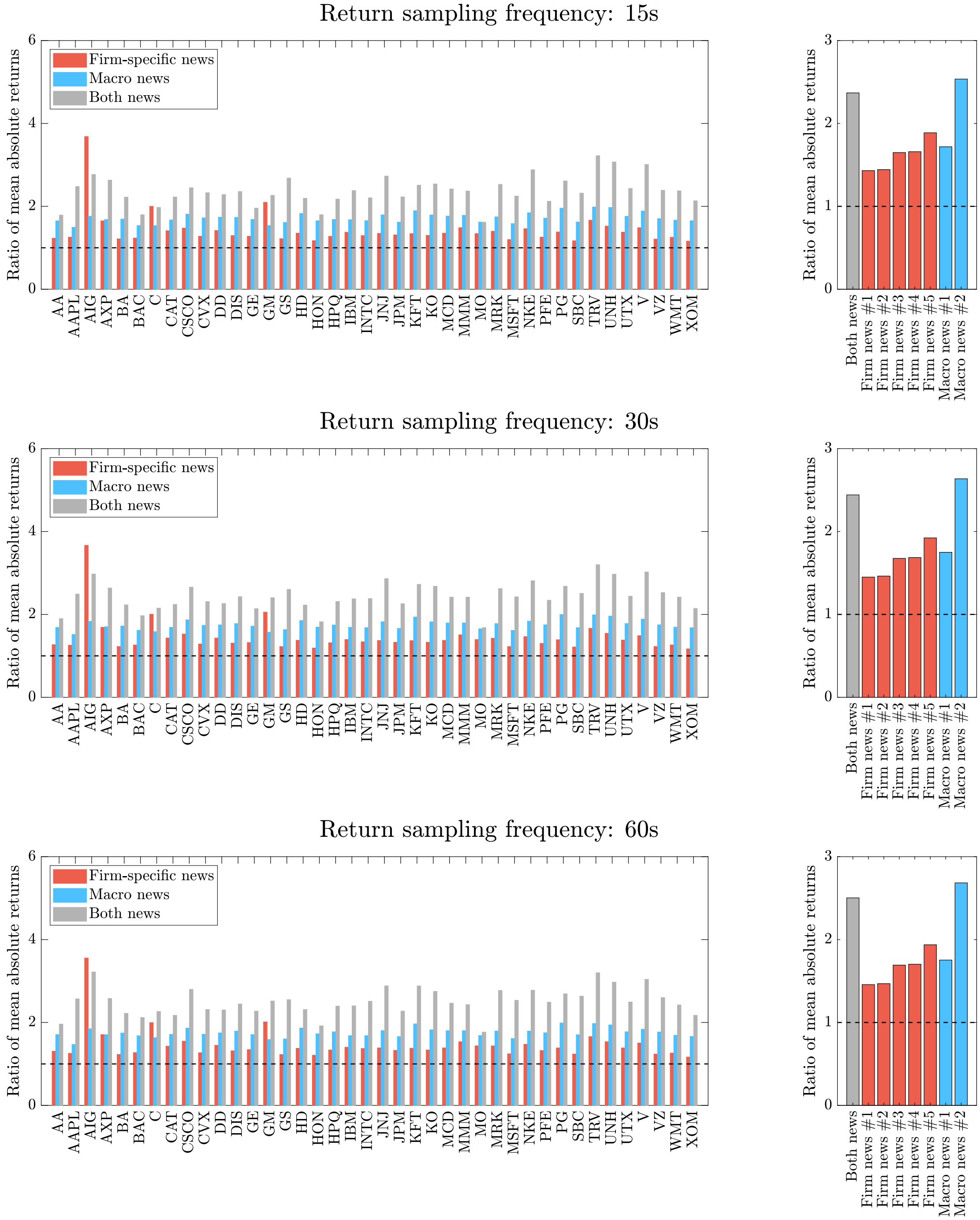

Figure A.1: Ratio of mean absolute returns at various return sampling frequencies

Note: The two panels on the first, second, and last rows correspond to the top two panels of Figure 2 for intraday returns under the return sampling frequencies of 15s, 30s, and 60s, respectively. Except for the choice of frequency, all other settings for producing this figure are the same as for those for producing the top two panels of Figure 2.