A News-Based Model for Stock Pricing

Anatoly B. Schmidt[1)]

Finance and Risk Engineering, NYU Tandon School

A news-based model (NBM) in which stock prices are determined by three types of news is proposed. First, it is non-diversifiable macroeconomic and geopolitical news. Their impact on prices is accounted using total market return in the spirit of the CAPM. Second, it is the equity sector/industry news whose impact on prices is accounted using returns of the relevant industry ETFs. Finally, the company-specific news and its momentum are described with an optimized ARMA-GARCH model. A comparison of the accuracy of NBM and the momentum-enhanced five-factor Fama-French model for a representative list of holdings of nine major US equity sector ETFs demonstrates superiority of the former in most examples.

[1)] Email: as8098@nyu.edu

Keywords: asset pricing, equity factors, ARMA model

JEL codes: C5, G11, G12, G14

1. Introduction

Modern asset pricing models are rooted in the arbitrage pricing theory (Ross 1976). This theory proposes that expected asset returns are determined with a linear combination of various risk factors but does not specify these factors. The most prominent model in this field was offered by Fama & French (1993, 2015). Carhart (1997) enhanced it with the momentum factor. The resulting regression model (denoted here FF5M) has five Fama-French factors and the momentum factor (Fama & French 2018):

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i(R_{M,t} - R_{f,t}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + m_i UMD_t + \varepsilon_{it} \qquad (1)$$

In (1), R_{it} is the return on asset i at time t, $R_{f,t}$ is the one-month U.S. Treasury bill rate, $R_{M,t}$ is the return on the value-weight portfolio of NYSE-AMEX-NASDAQ stocks, SMB_t (small minus big) and HML_t (high minus low book-to-market ratio) are the size and value factors, RMW_t (robust minus weak) is a profitability factor, CMA_t (conservative minus aggressive) is an investment factor, and UMD_t (up minus down) is a momentum factor. The first two terms in the right-hand side of (1) represent the famous capital asset pricing model (CAPM). While the accuracy of the Fama-French models usually increases with switching from the CAPM to three factors to five factors to the FF5M, the Fama-French framework has various problems (Blitz et al. 2018). Also, portfolio sorts, on which the estimates of the FF5M factors are based, depend on several heuristic choices that can significantly affect the model's performance (Soebhag et al. 2022).

In recent years, numerous risk factor models beyond the Fama-French framework (sometimes called *smart betas*) were developed in academia and in investment institutions. Harvey et al. (2016) reviewed 314 factors published in the academic literature and questioned performance robustness of many of them (see also Arnott et al. 2019). Among the general problems related to the factor-based investing are possible exposure of the factor of interest to other factors (which leads to multicollinearity), neglecting tail behavior of performance distribution, and increased correlations among various factors in bear markets. It can be said that some of these problems constitute selection bias (Novy-Marx 2015).

Here I offer a news-based model (NBM) of stock pricing that emphasizes dynamic nature of financial markets and does not rely on heuristic choices typical for the factor models. Indeed, news do have a major impact on prices, let alone (generally transient) effect of imbalance of buy/sell order flows that is sometimes called market inelasticity (see Gabaix & Koijen 2022 and references therein). The efficient market hypothesis even states that all new information is instantly incorporated into market prices, which can prevent forecasting of future returns. In real life, however, stock prices reach their new fair values not instantly but over some time, during which investors adapt to new information by trial and error (Lo 2004). As a result, markets may be partially predictable provided that price dynamics for the chosen time period is a (weakly) stationary process, i.e. has no unit roots (see, e.g., Tsay 2005). While news generate random price shocks, it is price momentum that adds a deterministic component to the price dynamics. Investors' delayed reaction to new information is an important cause of price momentum (DeLong et al. 1990; Barberis et al. 1998). Two other effects can contribute to momentum (Cai & Schmidt 2020). First, large investment institutions with high trading volumes try to minimize implementation shortfall, i.e. impact of trading due to limited market liquidity (Perold 1988). This problem is usually addressed by splitting intended trading volume into small pieces, so-called child orders. These orders are submitted to the market according to various time schedules that may last extended periods (Johnson, 2010). Also, one can expect that new pension contributions and other regularly scheduled money allocations are often invested in recent winners, thus, prolonging their momentum.

I discern three types of news for deriving NBM. First, it is undiversifiable macroeconomic and geopolitical news. I describe their impact on stock prices using total market return in the spirit of the CAPM. Second, it is equity sector/industry news whose impact is described with returns of relevant sector/industry ETFs. Importance of the industry effects on stock pricing was recently discussed by Eshani et al. (2021) and Vyas & van Baren (2021). Finally, I describe an impact of the company-specific news using an optimized ARMA-GARCH model. Hence, price momentum within NBM is an intrinsic part of the econometric framework. Some important company-specific news with predetermined announcement times, such as earnings reports and new product releases, can be incorporated into NBM explicitly (see, e.g., Schmidt 2020, 2021 and references therein) but this feature is not included in the current model.

In this work, I compare the NBM accuracy with that of FF5M for a representative list of holdings of nine US equity sector ETFs and find superiority of the former in most instances. The description of NBM, the data used in this work, and the results of comparison with FF5M are presented in the next three Sections. The discussion concludes.

2. The news-based model (NBM)

The NBM for stock return $R(t)$ has the following form (index i for denoting asset i is dropped):

$$R(t) = \alpha + \beta_M R_M(t) + \beta_{SI} R_{SI}(t) + \mathrm{ARMA}(p, q, t) + \varepsilon(t) \qquad (2)$$

In (2), R_M and R_{SI} are the total market return and equity sector/industry return, respectively. White noise, $\varepsilon(t)$, is assumed to have the GARCH(1, 1) form

$$\varepsilon(t) = z(t)\sigma(t),\ \sigma^2(t) = \omega + \gamma\, \varepsilon^2(t\text{-}1) + \delta\, \sigma^2(t\text{-}1) \qquad (3)$$

where $z(t)$ is IID process with zero mean and unit variance, $\sigma(t)$ is conditional variance. While there are several popular asymmetric GARCH-type models, recent analysis indicates that the classical GARCH(1, 1) may outperform them (Dol 2021).

The ARMA model has variable numbers of autoregressive (p) and moving average (q) terms (e.g. Tsay 2005)

$$\mathrm{ARMA}(p, q, t) = \sum_{k=1}^{p} a_k r(t-k) + \sum_{k=1}^{q} b_k \varepsilon(t-k) \tag{4}$$

In fact, the entire right-hand side of the equation (2) can be treated as an ARMA model with external regressors $R_M(t)$ and $R_{SI}(t)$, and a constant α.

NBM can be optimized using the Akaike informational criterion (AIC) or the Bayesian informational criterion (BIC) that represent compromises between the goodness of fit and simplicity of the model. BIC has a stronger penalty for the number of model parameters ($p + q$) and hence yields more parsimonious models. I used the software package *rugarch* (Galanos 2019) for estimating the optimal NBMs and compared an accuracy of both AIC-based and BIC-based optimal NBMs with that of FF5M.

3. The data

I considered current holdings of nine major US SPDR equity sector ETFs: Materials (ticker XLB), Energy (XLE), Finance (XLF), Industrials (XLI), Technology (XLK), Consumer Staples (XLP), Utilities (XLU), Healthcare (XLV), and Consumer Discretionary (XLY). For each sector ETF, I chose the first 12 alphabetically ordered company tickers that traded publicly at least since 1999. Since the equity sector ETFs started trading in the end of 1998, I used the data samples for 1/1/1999 – 12/31/2021 when it was possible. Some equity sectors include industries with distinct price dynamics. Then I used the relevant industry ETFs for calculating R_{SI}(t) rather than the sector ETFs. Some industry ETFs started trading several years after 1999. Then, the starting date of the data samples was changed accordingly.

The daily adjusted closing prices were downloaded from *finance.yahoo.com.* The SPDR S&P 500 ETF (ticker SPY) was used as a proxy for the total market. The Fama-French daily factors were downloaded from the Kenneth French's website: *https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.*

3. The results

Here I provide the results for the AIC-based and BIC-based optimized NBMs, and FF5M. The NBM optimization was done within the range $0 \leq p \leq 4$ and $0 \leq q \leq 4$. I estimated the accuracy of the models using two performance measures: mean square error (MSE) and mean absolute error (MAE). Ideally, the preferred model should have smaller MSE *and* MAE. However, it's not the case most of the time. In the tables below, I indicate my preferred model and criteria on which my choice is based using the bold font. Sometimes, when the differences in the NBM MSEs and the FF5M MSEs are comparable with the differences in their MAEs in absolute values but have opposite signs, the model's choice is a 'hair-splitting'. Then both models are listed in bold. It might be tempting to use mean absolute percentage error (MAPE) that is much more sensitive to the model specifics than MSE and MAE are. However, MAPE can have extremely high values for small daily returns; moreover, it is infinity for zero returns. Therefore I dropped MAPE from the consideration. Ultimately, however, comparative model accuracy should be studied out of sample.

SPDR Materials ETF (*XLB*)

For all XLB holdings listed in Table 1, the data samples were in the range 1/1/1999 – 12/31/2021 and the XLB returns were used for calculating R_{SI}(t). Some results for the XLB holdings are typical for other equity sectors. Specifically for such long time periods, the difference between MSE and MAE for the AIC-based and BIC-based NBMs rarely exceeds the rounding error. However, these differences grow for shorter data samples (cf. the ALB MSEs and MAEs for the samples starting in 1999 and 2020 in Table 1).

Table 1. The NBM and FF5M performance statistics for the XLB holdings

Company ticker	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
ALB since 1/1/1999	**AIC NBM**	(2,0)	**3.09**	**1.16**
	BIC NBM	(1,0)	3.09	1.17
	FF5M	na	3.23	1.23
ALB since 1/1/2020	**AIC NBM**	(0,1)	**6.15**	**1.82**
	BIC NBM	(0,0)	6.19	1.88
	FF5M	na	6.50	1.88
APD	**AIC NBM**	(4,3)	**1.42**	**0.78**
	BIC NBM	(1,1)	1.43	0.78
	FF5M	na	1.77	0.90
AVY	**AIC NBM**	(4,4)	**1.92**	**0.90**
	BIC NBM	(0,0)	1.93	0.90
	FF5M	na	2.01	0.91
DD	**AIC NBM**	(4,3)	**1.97**	**0.92**
	BIC NBM	(0,1)	1.97	0.93
	FF5M	na	2.67	1.12
ECL	**AIC NBM**	(3,3)	**1.18**	**0.75**
	BIC NBM	(1,0)	**1.18**	**0.75**
	FF5M	na	1.24	0.80
EMN	**AIC NBM**	(4,3)	**1.78**	**0.92**
	BIC NBM	(1,0)	1.79	0.92
	FF5M	na	2.18	1.03
FCX	**AIC NBM**	(3,3)	**6.41**	1.82
	BIC NBM	(0,0)	6.45	1.81
	FF5M	na	7.68	2.06
FMC	**AIC NBM**	(4,3)	**2.41**	**0.97**
	BIC NBM	(0,0)	2.42	0.97
	FF5M	na	2.64	1.10
IFF	**AIC NBM**	(3,2)	**1.77**	0.82
	BIC NBM	(0,1)	**1.77**	0.82
	FF5M	na	1.81	0.82
IP	**AIC NBM**	(2,3)	**2.52**	**1.07**
	BIC NBM	(0,1)	**2.52**	**1.07**
	FF5M	na	2.82	1.48
LIN	**AIC NBM**	(4,4)	**1.27**	**0.78**
	BIC NBM	(0,1)	**1.27**	**0.78**
	FF5M	na	1.65	0.87
MLM	AIC NBM	(4,4)	2.95	1.25
	BIC NBM	(0,0)	2.97	1.25
	FF5M	na	**2.84**	1.25

Sometimes, the AIC-based NBMs (but not BIC-based NBMs) were over-fitted. For example, the ALB AIC-based ARMA coefficient a_2 has the p-value greater than 0.1 (see Table 2). Nevertheless, the AIC-based and BIC-based NBM coefficients α, β_M, β_{SI}, ω, γ, and δ are very close, which is typical for all securities considered in this work.

Table 2. The NBM coefficients for ALB

NBM coef.	AIC-based NBM		BIC-based NBM	
	Estimate	p-value	Estimate	p-value
α	2.13E-04	0.159	2.14E-04	0.149
a_1	-0.067	0.000	-0.068	0.000
a_2	0.018	0.112	na	na
β_M	0.453	0.000	0.453	0.000
β_{SI}	0.664	0.000	0.664	0.000
ω	0.00E+00	0.344	0.00E+00	0.328
γ	0.020	0.000	0.020	0.000
δ	0.979	0.000	0.979	0.000

On the other hand, the market betas β_M for NBM and FF5M are very different (cf. β_M in Tables 2 for NBM and Table 3 for FF5M, respectively). The NBM β_M is often lower than β_{SI}, which implies that the industry news have a higher impact on the price dynamics than the macroeconomic news. Moreover, the NBM β_M can be even negative in some instances (see data for DVN in Table 5). Another feature for the stocks considered in this work is that the NBM coefficients α are statistically insignificant more often than not. This indicates that the celebrated CAPM alpha may be determined by the industry trend rather than by the company-specifics performance. Note that the FF5M alpha, too, can be statistically insignificant (see, e.g., Table 3).

Table 3. The FF5M coefficients for ALB

FF5M coef.	Estimate	p-value
α	0.020	0.404
β_M	1.200	0.000
s	0.494	0.000
h	0.054	0.201
r	0.428	0.000
c	0.530	0.000
m	-0.072	0.006

The results in Table 1 show that NBM was more accurate than FF5M for all XLB holdings considered in this work.

SPDR Energy ETF (*XLE*)

For all XLE holdings listed in Table 4, the data samples were in the range 1/1/1999 – 12/31/2021 and the XLE returns were used for calculating R_{SI}(t).

Table 4. The NBM and FF5M performance statistics for the XLE holdings

Company ticker	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
APA	**AIC NBM**	(1,0)	**3.14**	**1.09**
	BIC NBM	(0,0)	**3.14**	**1.09**
	FF5M	na	6.02	1.70
BKR	**AIC NBM**	(4,2)	**2.82**	**1.19**
	BIC NBM	(0,0)	2.84	1.20
	FF5M	na	4.72	1.64
COP	**AIC NBM**	(4,3)	**1.07**	**0.79**
	BIC NBM	(1,0)	**1.07**	**0.79**
	FF5M	na	2.52	1.20
CVX	**AIC NBM**	(1,1)	**0.66**	**0.61**
	BIC NBM	(0,1)	**0.66**	**0.61**
	FF5M	na	1.56	0.94
DVN	**AIC NBM**	(4,3)	**2.43**	**1.14**
	BIC NBM	(0,0)	2.44	1.14
	FF5M	na	5.05	1.71
EOG	**AIC NBM**	(3,2)	**2.14**	**1.05**
	BIC NBM	(0,0)	2.14	1.06
	FF5M	na	4.48	1.64
HAL	**AIC NBM**	(3,4)	**3.27**	**1.22**
	BIC NBM	(0,0)	3.28	1.22
	FF5M	na	5.67	1.69
HES	AIC NBM	(3,2)	2.04	1.07
	BIC NBM	(0,0)	**2.04**	**1.06**
	FF5M	na	4.27	1.57
MRO	**AIC NBM**	(4,2)	**2.66**	**1.10**
	BIC NBM	(0,1)	2.66	1.11
	FF5M	na	4.88	1.57
OKE	**AIC NBM**	(4,4)	**2.74**	1.06
	BIC NBM	(1,0)	2.75	1.05
	FF5M	na	3.18	1.72
OXY	**AIC NBM**	(1,0)	**2.04**	**0.90**
	BIC NBM	(1,0)	**2.04**	**0.90**
	FF5M	na	3.91	1.31
PXD	**AIC NBM**	(4,4)	**3.41**	**1.27**
	BIC NBM	(0,0)	3.42	1.28
	FF5M	na	5.39	1.76

The NBM β_M for the XLE holdings are very small or even negative, in contrast to the FF5M β_M (cf. the data for DVN in Tables 5 and 6).

Table 5. The NBM coefficients for DVN

NBM coef.	AIC-based NBM		BIC-based NBM	
	Estimate	p-value	Estimate	p-value
α	-2.25E-04	0.116	2.31E-04	0.112
a_1	0.047	0.000	na	na
a_2	0.068	0.000	na	na
a_3	-0.978	0.000	na	na
a_4	-0.022	0.080	na	na
b_1	-0.066	0.000	na	na
b_2	-0.057	0.000	na	na
b_3	0.980	0.000	na	na
β_M	-0.247	0.000	-0.246	0.000
β_{SI}	1.276	0.000	1.276	0.000
ω	1.00E-06	0.040	1.00E-06	0.039
γ	0.039	0.000	0.039	0.000
δ	0.956	0.000	0.956	0.000

Table 6. The FF5M coefficients for DVN

FF5M coef.	Estimate	p-value
α	-0.003	0.920
β_M	1.140	0.000
s	0.577	0.000
h	0.687	0.000
r	0.523	0.000
c	0.133	0.121
m	-0.036	0.280

For all XLE holdings considered in this work, NBM was more accurate than FF5M.

SPDR Finance ETF (*XLF*)

The Finance sector includes some industries whose ETFs are better fits for the XLF holdings than XLF itself. Specifically, the SPDR Insurance ETF (KIE) better reflects price dynamics of the Aflac Inc. (AFL) and Allstate Corporation (ALL); Invesco KBW Property & Casualty Insurance ETF

(KBWP) fits Arthur J. Gallagher & Co. (AJG), Aon plc (AON), and Chubb Limited (CB); American Express Company (AXP) is related to iShares U.S. Financial Services ETF (IYG). The rest of the companies listed in Table 7 (6 out of 12) are well covered by XLF.

Table 7. The NBM and FF5M performance statistics for the XLF holdings

Company ticker	Industry ETF	Sample starting date	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
AFL	KIE	12/1/2005	**AIC NBM**	(0,0)	**2.21**	**0.67**
			BIC NBM	(0,0)	**2.21**	**0.67**
			FF5M	na	2.51	0.79
AIG	XLF	1/1/1999	**AIC NBM**	(0,0)	8.37	**0.98**
			BIC NBM	(0,0)	8.37	**0.98**
			FF5M	na	**8.31**	1.11
AJG	KBWP	1/1/2011	**AIC NBM**	(2,2)	**0.69**	**0.62**
			BIC NBM	(0,0)	**0.69**	**0.62**
			FF5M	na	0.74	0.65
ALL	KIE	12/1/2005	**AIC NBM**	(4,3)	**1.25**	0.61
			BIC NBM	(0,0)	1.26	0.60
			FF5M	na	1.67	0.75
AON	KBWP	1/1/2011	**AIC NBM**	(2,3)	**1.04**	**0.67**
			BIC NBM	(0,0)	1.05	0.67
			FF5M	na	1.05	0.69
AXP	IYG	7/1/2000	**AIC NBM**	(2,3)	**1.76**	**0.78**
			BIC NBM	(0,0)	1.77	0.78
			FF5M	na	1.88	0.87
BAC	XLF	1/1/1999	AIC NBM	(3,4)	2.04	0.77
			BIC NBM	(0,0)	**2.04**	**0.76**
			FF5M	na	2.70	1.00
BEN	XLF	1/1/1999	**AIC NBM**	(3,1)	**1.97**	**0.95**
			BIC NBM	(0,0)	1.97	0.96
			FF5M	na	1.99	0.97
BK	XLF	1/1/1999	**AIC NBM**	(3,3)	**1.92**	**0.82**
			BIC NBM	(0,0)	1.93	0.82
			FF5M	na	2.36	0.93
BLK	XLF	10/1/1999	**AIC NBM**	(3,3)	2.74	**0.99**
			BIC NBM	(0,0)	2.76	1.00
			FF5M	na	**2.68**	1.05
C	XLF	1/1/1999	**AIC NBM**	(4,1)	**3.02**	**0.79**
			BIC NBM	(0,0)	3.04	0.79
			FF5M	na	3.38	0.99
CB	KBWP	1/1/2011	**AIC NBM**	(3,4)	**0.76**	**0.61**
			BIC NBM	(0,1)	0.77	0.61
			FF5M	na	0.89	0.67

The AIG was the only case of 'hair-splitting' for the XLF holdings considered here: the NBM MAE is 13% lower than the FF5M MAE while the FF5M MSE is about one percent lower than the NBM MSE. For the other 11 XLF holdings, the NBM was more accurate.

SPDR Industrials ETF (*XLI*)

The Industrials sector includes several industries whose ETFs are better fits for their holdings than XLI (see Table 8).

Table 8. The NBM and FF5M performance statistics for the XLI holdings

Company ticker	Industry ETF	Sample starting date	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
ALK	JETS	5/1/2015	**AIC NBM**	(3,2)	**1.66**	**0.93**
			BIC NBM	(0,0)	1.68	0.97
			FF5M	na	3.78	1.51
AME	XLI	1/1/1999	**AIC NBM**	(2,1)	1.76	**0.86**
			BIC NBM	(0,0)	1.77	0.87
			FF5M	na	**1.73**	0.89
AOS	FXR	6/1/2007	AIC NBM	(3,3)	2.32	0.96
			BIC NBM	(0,0)	2.34	0.97
			FF5M	na	**2.04**	0.97
BA	XLI	1/1/1999	**AIC NBM**	(3,3)	**2.67**	**1.05**
			BIC NBM	(0,1)	2.68	1.06
			FF5M	na	2.93	1.18
CAT	XLI	1/1/1999	**AIC NBM**	(3,4)	**1.96**	**1.00**
			BIC NBM	(0,1)	1.96	1.01
			FF5M	na	2.20	1.11
CHRW	FTXR	9/30/2016	**AIC NBM**	(3,3)	1.84	**9.48**
			BIC NBM	(0,0)	1.88	9.83
			FF5M	na	**1.82**	9.90
CMI	XLI	1/1/1999	**AIC NBM**	(3,4)	**3.01**	**1.10**
			BIC NBM	(0,0)	3.03	1.11
			FF5M	na	3.22	1.20
CPRT	ONLN	11/1/2018	**AIC NBM**	(3,4)	1.91	**0.94**
			BIC NBM	(0,0)	1.91	0.96
			FF5M	na	**1.84**	1.01
CSX	FTXR	9/30/2016	AIC NBM	(3,4)	1.93	0.98
			BIC NBM	(0,0)	1.96	0.98
			FF5M	na	**1.88**	**0.96**
CTAS	XLI	1/1/1999	AIC NBM	(2,3)	2.27	0.85
			BIC NBM	(1,0)	2.28	0.85
			FF5M	na	**2.25**	**0.84**
DE	XLI	1/1/1999	**AIC NBM**	(1,1)	**2.49**	**1.09**
			BIC NBM	(0,0)	**2.49**	**1.09**
			FF5M	na	2.64	1.15
DOV	XLI	1/1/1999	**AIC NBM**	(2,3)	**1.44**	**0.87**
			BIC NBM	(0,0)	1.45	0.87
			FF5M	na	1.54	0.92

Specifically, the U.S. Global Jets ETF (JETS) is related to the airline companies, such as Alaska Air Group (ALK); First Trust Industrials/Producer Durables Fund (FXR) – to Specialty Industrial Machinery, such as A. O. Smith Corporation (AOS); First Trust Nasdaq Transportation ETF (FTXR) – to C.H. Robinson Worldwide (CHRW) and CSX Corporation (CSX).

While Copart Inc. (CPRT) that currently provides online auctions and vehicle remarketing services is an XLI holding, its price dynamics is better described with the ProShares Online Retail ETF (ONLN). Still, CPRT is a case of 'hair-splitting': while the FF5M MSE is 4% lower than the NBM SME, the FF5M MAE is 7% higher than the MAE of the AIC-based NBM. There are other two instances of 'hair-splitting' among 12 XLI holdings listed in Table 10 (AME and CHRW) where the differences between MSEs and MAEs in NBM and FF5M have opposite signs and are comparable in their absolute values. FF5M was more accurate for three XLI holdings: AOS, CSX, and CTAS.

SPDR Technology ETF (*XLK*)

The Technology sector has several industries with distinct price dynamics described with the relevant ETFs. Specifically, Adobe (ADBE), Autodesk (ADSK), and Cadence Design Systems (CDNS) are related to iShares Expanded Tech-Software Sector ETF (IGV); Applied Materials (AMAT) and Advanced Micro Devices (AMD) – to VanEck Semiconductor ETF (SMH); ANSYS (ANSS) and Amphenol Corporation (APH) – to Vanguard Information Technology Index Fund (VGT). However, Automatic Data Processing (ADP) that provides human capital management solutions fits better with iShares U.S. Financial Services ETF (IYG). Using these industry ETFs for calculating R_{SI}(t) yields the NBMs that are more accurate than FF5Ms for 9 of 12 XLK holdings considered in this work (see Table 9). FF5M is more accurate for ADP while APH and CTSH represent the instances of 'hair-splitting'.

Table 9. The NBM and FF5M performance statistics for the XLK holdings

Company ticker	Industry ETF	Sample starting date	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
AAPL	XLK	1/1/1999	**AIC NBM**	(3,2)	**4.03**	**1.15**
			BIC NBM	(0,0)	4.03	1.16
			FF5M	na	4.38	1.33
ADBE	IGV	8/1/2001	**AIC NBM**	(2,0)	**2.36**	**0.88**
			BIC NBM	(0,1)	**2.36**	**0.88**
			FF5M	na	2.85	1.05
ADI	XLK	1/1/1999	**AIC NBM**	(3,4)	**3.58**	**1.15**
			BIC NBM	(0,1)	3.59	1.16
			FF5M	na	3.98	1.27
ADP	IYG	7/1/2000	AIC NBM	(3,3)	1.29	7.31
			BIC NBM	(0,1)	1.29	7.39
			FF5M	na	**1.25**	7.32
ADSK	IGV	8/1/2001	**AIC NBM**	(1,0)	3.27	1.17
			BIC NBM	(0,0)	3.28	1.16
			FF5M	na	3.51	1.25
AMAT	SMH	7/1/2000	**AIC NBM**	(2,3)	**1.81**	**1.01**
			BIC NBM	(1,0)	**1.81**	**1.01**
			FF5M	na	3.76	1.43
AMD	SMH	7/1/2000	**AIC NBM**	(4,4)	**9.41**	**2.00**
			BIC NBM	(0,0)	9.45	2.00
			FF5M	na	10.6	2.17
ANSS	IGV	8/1/2001	**AIC NBM**	(4,4)	**2.90**	**1.01**
			BIC NBM	(0,0)	2.91	1.01
			FF5M	na	2.92	1.04
APH	VGT	2/1/2004	**AIC NBM**	(3,3)	1.51	**0.81**
			BIC NBM	(0,0)	1.52	0.81
			FF5M	na	**1.49**	0.84
CDNS	IGV	8/1/2001	**AIC NBM**	(3,4)	**3.48**	1.04
			BIC NBM	(0,1)	3.49	1.03
			FF5M	na	3.63	1.11
CSCO	XLK	1/1/1999	**AIC NBM**	(3,0)	**2.39**	**0.95**
			BIC NBM	(0,0)	**2.39**	**0.95**
			FF5M	na	2.80	1.07
CTSH	VGT	2/1/2004	**AIC NBM**	(4,4)	**2.70**	**1.03**
			BIC NBM	(0,0)	2.72	1.04
			FF5M	na	**2.70**	1.05

Consumer Staples Sector SPDR ETF (*XLP*)

For all XLP holdings considered here (see Table 10), the data samples were in the range 1/1/1999 – 12/31/2021 and the XLP returns were used for calculating R_{SI}(t). For the XLP holdings considered here, the NBM was generally more accurate than the FF5M. The only exclusion was the Archer-Daniels-Midland Company (ADM). Also, the Estée Lauder Companies Inc. (EL) was an instance of 'hair-splitting'.

Table 10. The NBM and FF5M performance statistics for the XLP holdings

Company ticker	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
ADM	AIC NBM	(4,2)	2.60	1.12
	BIC NBM	(0,0)	2.61	1.11
	FF5M	na	**2.54**	**1.10**
CAG	**AIC NBM**	(0,1)	**1.81**	**0.84**
	BIC NBM	(0,0)	**1.81**	**0.84**
	FF5M	na	1.95	0.92
CHD	**AIC NBM**	(3,1)	**1.88**	**0.86**
	BIC NBM	(1,0)	**1.88**	**0.86**
	FF5M	na	1.94	0.92
CL	**AIC NBM**	(2,4)	**1.21**	**0.65**
	BIC NBM	(0,1)	1.22	0.65
	FF5M	na	1.41	0.78
CLX	**AIC NBM**	(0,1)	**1.78**	**0.79**
	BIC NBM	(0,1)	**1.78**	**0.79**
	FF5M	na	1.95	0.85
COST	**AIC NBM**	(0,2)	2.16	**0.90**
	BIC NBM	(0,0)	2.17	0.90
	FF5M	na	**2.11**	0.92
CPB	**AIC NBM**	(2,3)	**1.62**	**0.81**
	BIC NBM	(0,0)	**1.62**	**0.81**
	FF5M	na	1.79	0.89
EL	**AIC NBM**	(1,1)	2.54	**1.03**
	BIC NBM	(0,0)	2.55	1.03
	FF5M	na	**2.52**	1.04
GIS	**AIC NBM**	(4,4)	**1.04**	**0.71**
	BIC NBM	(0,1)	**1.04**	**0.71**
	FF5M	na	1.14	0.77
HRL	**AIC NBM**	(2,3)	**1.65**	**0.85**
	BIC NBM	(0,1)	1.66	0.85
	FF5M	na	1.76	0.93
HSY	**AIC NBM**	(3,3)	**1.55**	**0.77**
	BIC NBM	(0,1)	**1.55**	**0.77**
	FF5M	na	1.67	0.86
K	**AIC NBM**	(4,3)	**1.40**	**0.72**
	BIC NBM	(0,1)	**1.41**	**0.71**
	FF5M	na	1.56	0.82

Utilities Sector SPDR ETF (*XLU*)

The Utilities sector ETF is another example of a homogeneous sector (on par with XLB, XLE, and XLP) when the single sector ETF could be used for calculating $R_{SI}(t)$. For all XLU holdings considered in this work, NBM was more accurate than FF5M (see Table 11).

Table 11. The NBM and FF5M performance statistics for the XLU holdings

Company ticker	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
AEE	**AIC NBM**	(2,3)	**0.66**	**0.58**
	BIC NBM	(0,0)	**0.66**	**0.58**
	FF5M	na	1.21	0.90
AEP	**AIC NBM**	(3,3)	**0.94**	**0.58**
	BIC NBM	(0,1)	**0.94**	**0.58**
	FF5M	na	1.71	0.92
AES	**AIC NBM**	(3,3)	**7.78**	**1.37**
	BIC NBM	(0,0)	7.81	1.37
	FF5M	na	8.12	1.48
ATO	**AIC NBM**	(2,1)	**1.24**	**0.75**
	BIC NBM	(0,1)	**1.24**	**0.75**
	FF5M	na	1.46	0.90
CMS	**AIC NBM**	(4,4)	**1.98**	**0.65**
	BIC NBM	(0,1)	1.99	0.65
	FF5M	na	2.58	1.06
CNP	**AIC NBM**	(0,1)	**3.42**	**0.84**
	BIC NBM	(0,1)	**3.42**	**0.84**
	FF5M	na	4.08	1.08
D	**AIC NBM**	(1,1)	**0.75**	**0.59**
	BIC NBM	(0,0)	**0.75**	**0.59**
	FF5M	na	1.32	0.87
DTE	AIC NBM	(4,4)	0.78	0.55
	BIC NBM	(0,0)	**0.78**	**0.54**
	FF5M	na	1.27	0.86
DUK	**AIC NBM**	(2,1)	**1.06**	**0.57**
	BIC NBM	(1,0)	**1.06**	**0.57**
	FF5M	na	1.75	0.96
ED	**AIC NBM**	(0,0)	**0.62**	**0.51**
	BIC NBM	(0,0)	**0.62**	**0.51**
	FF5M	na	1.10	0.83
EIX	**AIC NBM**	(4,3)	**3.00**	**0.79**
	BIC NBM	(1,0)	**3.00**	**0.79**
	FF5M	na	3.67	1.11
ES	**AIC NBM**	(0,3)	**0.94**	**0.62**
	BIC NBM	(0,0)	**0.94**	**0.62**
	FF5M	na	1.46	0.93

Healthcare Sector SPDR ETF (*XLV*)

Two industries within the healthcare sector XLV have distinct price dynamics: medical devices and biotechnology. Hence, two relevant ETFs were used for calculating $R_{SI}(t)$: iShares U.S. Medical Devices ETF (IHI) and iShares Biotechnology ETF (IBB). The NBM was more accurate than FF5M for all XLV holdings considered in this work (see Table 12). Amgen (AMGN) and Bristol-Myers Squibb Company (BMY) illustrate the instances when the NBM MSE and the FF5M MSE are equal, so MAE becomes the criterion for the model's choice.

Table 12. The NBM and FF5M performance statistics for the XLV holdings

Company ticker	Industry ETF	Sample starting date	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
A	IHI	6/1/2006	**AIC NBM**	**(4,3)**	**1.58**	**0.88**
			BIC NBM	(0,1)	1.59	0.89
			FF5M	na	1.61	0.93
ABC	XLV	1/1/1999	**AIC NBM**	(4,4)	**3.21**	**1.05**
			BIC NBM	(0,1)	3.22	1.05
			FF5M	na	3.31	1.11
ABMD	IHI	6/1/2006	**AIC NBM**	(4,4)	**6.86**	1.60
			BIC NBM	(0,0)	6.88	1.59
			FF5M	na	6.92	1.63
ABT	IHI	6/1/2006	**AIC NBM**	(4,3)	**0.98**	**0.74**
			BIC NBM	(0,0)	0.98	0.75
			FF5M	na	1.13	0.79
AMGN	XLV	1/1/1999	**AIC NBM**	(0,1)	2.95	**1.01**
			BIC NBM	(0,0)	2.96	1.01
			FF5M	na	2.95	1.16
BAX	IHI	6/1/2006	**AIC NBM**	(2,4)	**1.18**	**0.74**
			BIC NBM	(0,0)	**1.18**	**0.74**
			FF5M	na	1.33	0.83
BDX	XLV	1/1/1999	**AIC NBM**	(2,1)	**2.00**	**0.83**
			BIC NBM	(1,1)	**2.00**	**0.83**
			FF5M	na	2.02	0.90
BIIB	IBB	3/1/2001	**AIC NBM**	(4,1)	**4.56**	**1.16**
			BIC NBM	(0,1)	**4.56**	**1.16**
			FF5M	na	5.66	1.36
BIO	IHI	6/1/2006	**AIC NBM**	(4,2)	**2.05**	**0.87**
			BIC NBM	(0,1)	2.06	0.89
			FF5M	na	2.12	0.94
BMY	XLV	1/1/1999	**AIC NBM**	(2,0)	2.34	**0.92**
			BIC NBM	(0,0)	2.34	0.93
			FF5M	na	2.34	1.02
BSX	IHI	6/1/2006	**AIC NBM**	(3,0)	**2.33**	**1.02**
			BIC NBM	(0,1)	**2.33**	**1.02**
			FF5M	na	2.95	1.16
CAH	XLV	1/1/1999	**AIC NBM**	(1,1)	**2.67**	**1.01**
			BIC NBM	(0,0)	2.68	1.01
			FF5M	na	2.70	1.06

Consumer Discretionary Sector SPDR ETF (*XLY*)

The consumer discretionary sector includes several industries with distinct price dynamics. The following relevant ETF returns were used for calculating R_{SI}(t): SPDR Retail ETF (XRT), Invesco Dynamic Leisure and Entertainment ETF (PEJ), SPDR Homebuilders ETF (XHB), First Trust Dow Jones Internet Index Fund (FDN), and First Trust S-Network Future Vehicles & Technology ETF (CARZ). NBM was more accurate than FF5M for all XLY holdings considered in this work except the Ford Motors Company (F) (see Table 13).

Table 13. The NBM and FF5M performance statistics for the XLY holdings

Company ticker	Industry ETF	Sample starting date	Model	ARMA (p,q)	MSE*10^4	MAE*10^2
AAP	XLY	12/1/2001	**AIC NBM**	(3,3)	**3.15**	**1.16**
			BIC NBM	(0,0)	3.16	1.17
			FF5M	na	3.21	1.20
AMZN	FDN	7/1/2006	**AIC NBM**	(0,2)	**3.06**	**1.03**
			BIC NBM	(0,0)	3.07	1.03
			FF5M	na	3.58	1.18
AZO	XLY	1/1/1999	**AIC NBM**	(0,1)	**2.31**	**1.01**
			BIC NBM	(0,1)	**2.31**	**1.01**
			FF5M	na	2.44	1.05
BBWI	XRT	7/1/2006	**AIC NBM**	(3,3)	**4.79**	**1.20**
			BIC NBM	(0,0)	4.79	1.21
			FF5M	na	5.20	1.41
BBY	XRT	7/1/2006	**AIC NBM**	(4,3)	**4.01**	**1.25**
			BIC NBM	(0,0)	**4.01**	**1.25**
			FF5M	na	4.43	1.37
BKNG	FDN	7/1/2006	**AIC NBM**	(3,4)	**3.99**	**1.13**
			BIC NBM	(0,0)	4.00	1.13
			FF5M	na	4.07	1.23
CCL	PEJ	7/1/2006	**AIC NBM**	(2,3)	**4.33**	**1.14**
			BIC NBM	(0,0)	4.34	1.14
			FF5M	na	4.44	1.30
DHI	XHB	2/6/2006	**AIC NBM**	(2,3)	**2.55**	**1.04**
			BIC NBM	(0,1)	2.56	1.04
			FF5M	na	5.39	1.68
DLTR	XRT	7/1/2006	**AIC NBM**	(1,0)	**3.06**	**1.13**
			BIC NBM	(1,0)	**3.06**	**1.13**
			FF5M	na	3.25	1.18
DRI	PEJ	7/1/2006	**AIC NBM**	(3,3)	**3.01**	**1.10**
			BIC NBM	(0,0)	3.01	1.11
			FF5M	na	3.37	1.19
EBAY	FDN	7/1/2006	**AIC NBM**	(3,4)	**2.59**	**1.06**
			BIC NBM	(0,1)	2.60	1.06
			FF5M	na	2.80	1.10
F	CARZ	6/1/2011	AIC NBM	(3,2)	2.19	0.98
			BIC NBM	(0,0)	2.20	0.97
			FF5M	na	**2.05**	**0.94**

4. Discussion

The news-based model (NBM) of stock pricing offered in this work has an advantage over other heuristic asset pricing models in that it reflects dynamic nature of financial markets and does not depend on a subjective choice of risk factors and portfolio sorts. NBM, too, has some implementation uncertainties. First, its accuracy depends on the presence of the relevant industry ETFs. NBM seems to be more accurate for those companies whose price dynamics is similar to that of their equity sectors. The examples include holdings of such sectors as XLB, XLE, XLP, and XLU (see Table 14). A more general problem may be that some companies have products related to several industries or changed their business directions in recent times.

Table 14. Average NBM SMEs and FF5M SMEs for the equity sector holdings

Sector	XLB		XLE		XLF		XLI		XLK		XLP		XLU		XLV		XLY		All sectors	
MSE	NBM	FF5M	NBM	FF5M	NBM	FF5M	NBM	FF5M	NBM	FF5M	NBM	FF5M	NBM	FF5M	NBM	FF5M	NBM	FF5M	NBM	FF5M
Average	2.39	2.67	2.37	4.30	2.31	2.51	2.11	2.32	3.23	3.66	1.77	1.86	1.93	2.48	2.73	2.92	3.25	3.69	2.45	2.93
p-value	0.004		0.000		0.013		0.249		0.028		0.005		0.000		0.067		0.081		0.000	

Also, estimation of the optimal ARMA terms in NBM may represent an over-fitting challenge. While the AIC-based NBM is somewhat more accurate in-sample than the BIC-based model, some of its coefficients can be statistically insignificant. Specifically, the coefficients of the AIC-based NBM for 16 out of 108 securities considered in this work (i.e. 14.8%) had some p-values exceeding 0.05 while those of the BIC-based NBM had none. Nevertheless, the results obtained in this work show that NBM is generally superior to FF5M: it was more accurate for 94 securities (i.e. in 87.0%) and less accurate only in 7 occurrences (6.5%). The remaining 7 securities (6.5%) represent the instances of 'hair-splitting' for which the preferred model can be based on the choice of the performance measure (MSE or MAE) and/or on the size of their differences.

References

Arnott R., C. R. Harvey, V. Kalesnik, and J. Linnainmaa (2019). "Alice's Adventures in Factorland: Three Blunders That Plague Factor Investing." *The Journal of Portfolio Management* 45 (4), 18-36.

Barberis N., A. Shleifer, and R. Vishny (1998). "A model of investor sentiment," *Journal of Financial Economics* 49, 307–343.

Blitz D., M. X. Hanauer, M. Vidojevic, and P. van Vliet (2018). " Five Concerns with the Five-Factor Model." *The Journal of Portfolio Management* 44(4), 71-78.

Cai H. and A. B. Schmidt (2020). "What's So Special about the Time Series Momentum?" Journal of Investment Strategies 9 (2), 33 – 43.

Carhart M. (1997). "On Persistence in Mutual Fund Performance." *Journal of Finance* 52, 57-82.

DeLong J. B, A. Shleifer, L. H. Summers, and R. J. Waldmann (1990). "Positive feedback investment strategies and destabilizing rational speculation." *Journal of Finance* 45, 379–395.

Dol M. (2021). "Comparison of the GARCH, EGARCH, GJR-GARCH and TGARCH model in times of crisis for the S&P500, NASDAQ and Dow-Jones." Erasmus School of Economics. Available online: https://thesis.eur.nl/pub/59759/Thesis-Misha-Dol-final-version.pdf

Ehsani S., C.R. Harvey, and F. Li, Feifei (2021). "Is Sector-neutrality in Factor Investing a Mistake?" Available online: https://ssrn.com/abstract=3959116

Fama F. and K. R. French (1993) "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33, 3–56.

Fama E. F. and K. R. French (2015). "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 16, 1–22.

Fama E. F. and K.R. French (2018). "Choosing factors." *Journal of Financial Economics* 128 (2), 234–252.

Gabaix X and K.S.J. Koijen (2021). "In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis." NBER working paper 28967.

Galanos A. (2019). Package '*rugarch*'. Available online:

https://cran.r-project.org/web/packages/rugarch/rugarch.pdf

Harvey C., Y. Liu, and H. Zhu (2016). ". . . and the Cross-Section of Expected Returns." *Review* *of Financial Studies* 29 (1), 5–68.

Johnson B. (2010). *Algorithmic trading and DMA: An introduction to direct access trading strategies*. 4Myeloma Press.

Lo A. W. (2004). “Adaptive market hypothesis: Market efficiency from evolutionary perspective.” *Journal of Portfolio Management* 30, 15 – 29.

Novy-Marx R. (2015). “Backtesting strategies based on multiple signals.” NBER working paper #21329.

Ross S.A. (1976). The Arbitrage Theory of Capital Pricing. *Journal of Economic Theory* 13, 341 – 360.

Perold A. F. (1988). “The implementation shortfall: Paper versus reality.” *Journal of Portfolio Management* 24, 4–9.

Schmidt A. B. (2020). “Impact of Earnings Announcements for Dow Jones Index Stocks.” Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3537660.

Schmidt A. B. (2021). “*Modern Equity Investing Strategies*.” World Scientific.

Soebhag A., B. van Vliet, and P. Verwijmeren (2022). “Non-Standard Errors in Asset Pricing: Mind Your Sorts.” Available online: https://ssrn.com/abstract=4136672

Tsay R. S. (2005). *Analysis of financial time series.* Wiley.

Vyas K. and M. van Baren (2021). “Should Equity Factors Be Betting on Industries?” *The Journal of Portfolio Management* 48 (1) 73-92.